

Komputeryzacja testów w fizjoterapii

Computerization of testing in physical therapy

Numer DOI: 10.2478/v10109-011-0013-7

Emilia Mikołajewska¹, Dariusz Mikołajewski²

¹ Klinika Rehabilitacji, 10 Wojskowy Szpital Kliniczny z Polikliniką SP ZOZ w Bydgoszczy
Rehabilitation Clinic, 10th Military Polyclinical Hospital in Bydgoszcz

² Katedra Informatyki Stosowanej, Wydział Fizyki, Astronomii i Informatyki Stosowanej Uniwersytetu Mikołaja Kopernika w Toruniu
Department of Informatics, Faculty of Physics, Astronomy and Informatics, Nicolaus Copernicus University in Torun

Streszczenie:

Rosnące wykorzystanie komputerowego testowania adaptacyjnego (ang. *computerized adaptive tests* – CAT) do oceny wyników fizjoterapii stymuluje szereg pytań w zakresie m.in. wdrażania tej metody i interpretacji jej wyników. CAT pozwala, w oparciu o zbiór (bank) pytań mierzyć, m.in. wpływ poszczególnych schorzeń na funkcjonowanie i samopoczucie pacjenta, zapewniając szybkie, wiarygodne i dokładne wartości szacunkowe oraz efektywne gromadzenie danych. Szersze wykorzystanie CAT może poprawić użyteczność i ułatwić interpretację kliniczną tych indywidualnie dopasowanych do pacjenta badań, wspomagając specjalistów medycznych w codziennej praktyce klinicznej. Co więcej – może być stosowane zarówno w fizjoterapii szpitalnej, jak również ambulatoryjnej i domowej.

Słowa kluczowe: rehabilitacja, fizjoterapia, komputerowe testowanie adaptacyjne.

Abstract:

Increased use of computerized adaptive tests (CATs) to generate outcome measures during physical therapy has stimulated questions concerning CAT implementation and score interpretation. CAT allows, based on item bank, to measure e.g. the impact of diseases on functioning and well-being, producing quick, valid and precise estimates and efficient outcomes collection. Wider application of CAT may improve usefulness and clinical interpretation of these individually tailored outcome measures and assist clinicians during clinical practice. What more, it can be applied both in hospital and outpatient physical therapy settings.

Key words: rehabilitation, physical therapy, computerized adaptive testing.

Wprowadzenie

Dokładne i efektywne narzędzia pomiarowe są w fizjoterapii niezbędne z kilku zasadniczych przyczyn:

- zwiększają wiarygodność, dokładność i porównywalność wyników badań diagnostycznych,
- skracają czas wykonania badania,
- podwyższają efektywność terapii,
- zmniejszają ryzyko wdrożenia niewłaściwego postępowania terapeutycznego,
- pozwalają wdrożyć wskazówki i procedury postępowania w poszczególnych przypadkach chorobowych,
- pozwalają na natychmiastowe przetwarzanie, obrazowanie, porównanie danych, np. z poprzednimi wynikami danego pacjenta lub z normą oraz przesłanie ich do innych specjalistów i włączenie do dokumentacji medycznej pacjenta,
- pozwalają w praktyce klinicznej implementować zasady Medycyny Opartej na Faktach (ang. *Evidence Based Medicine* – EBM) [1-6].

Przegląd literatury w oparciu o bazę PubMed [7] (ryc. 1) wykazał, że liczba artykułów poświęconych komputero-

Introduction

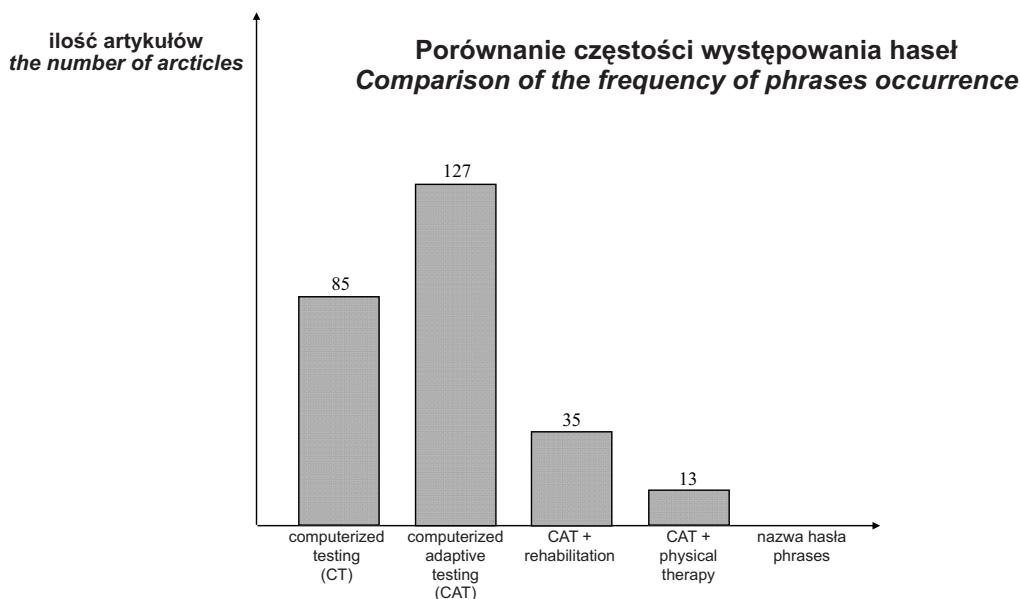
Precise and effective measuring tools are required in physiotherapy for several fundamental reasons:

- they increase validity, precision and comparability of outcomes of diagnostic tests,
- reduce test duration,
- improve effectiveness of therapy,
- reduce risk of introduction of inadequate therapeutic procedures,
- enable tailoring procedures for individual clinical cases,
- allow for immediate processing, imaging, comparison of data (f.e. With previous outcomes of a given patient or reference norms) and adding it to patient's medical history or sending it to other specialists,
- allow to implement EBM (Evidence Based Medicine) in clinical practice [1-6].

PubMed based review of literature [7] (Fig. 1) showed that the number of papers regarding CAT (Computerized Adaptive Testing) at present surpasses by 49.4% the number of papers concerning the older solution, i.e. the CT (Comput-

wemu testowaniu adaptacyjnemu (CAT) już obecnie przewyższa o 49,4% liczbę artykułów poświęconych starszemu rozwiązaniu, czyli testowaniu komputerowemu (ang. *computerized testing* – CT). Natomiast liczba artykułów poświęconych wykorzystaniu CAT w rehabilitacji i fizjoterapii stanowią odpowiednio 27,6% i 10,2% wszystkich artykułów poświęconych CAT. Z jednej strony może to świadczyć o dużej popularności badań nad wykorzystaniem CAT w rehabilitacji, a z drugiej – o niedocenianiu znaczenia fizjoterapii w tym zakresie.

erized Testing). What is interesting, the number of papers on CAT use in rehabilitation and physiotherapy is respectively 27.6% and 10.2% of all of the papers on the use of CAT. On one hand, this may stand for high popularity of the CAT based studies in rehabilitation, and on the other hand, it may mean that physiotherapy has been underestimated in that particular regard.



Ryc. 1. Wyniki wyszukiwania w bazie pubmed (U.S. National Library of Medicine) [7]
Fig. 1. Results of investigation of the PubMed database [7]

Informatyzacja medycyny, a w szczególności fizjoterapii i rehabilitacji, jest już faktem. Pozwala ona nie tylko na praktyczne implementowanie zasad EBM w praktyce klinicznej, ale również na automatyczne tworzenie dokumentacji medycznej, zestawień i baz danych o schorzeniach. Przyczynia się to do wzrostu ilości nagromadzonej wiedzy, dając możliwość wyciągnięcia z niej wniosków za pomocą programów do eksploracji wiedzy z danych (ang. *data mining*), również na potrzeby medycznych programów ekspertowych czy programów wspierających podejmowanie decyzji (tzw. „drugiej opinii”), również na podstawie niepełnych danych, zwiększając prawdopodobieństwo sukcesu terapeutycznego.

Przedstawione na ryc. 2 z informatyzowane narzędzia diagnostyczne wykorzystywane w fizjoterapii obejmują przede wszystkim:

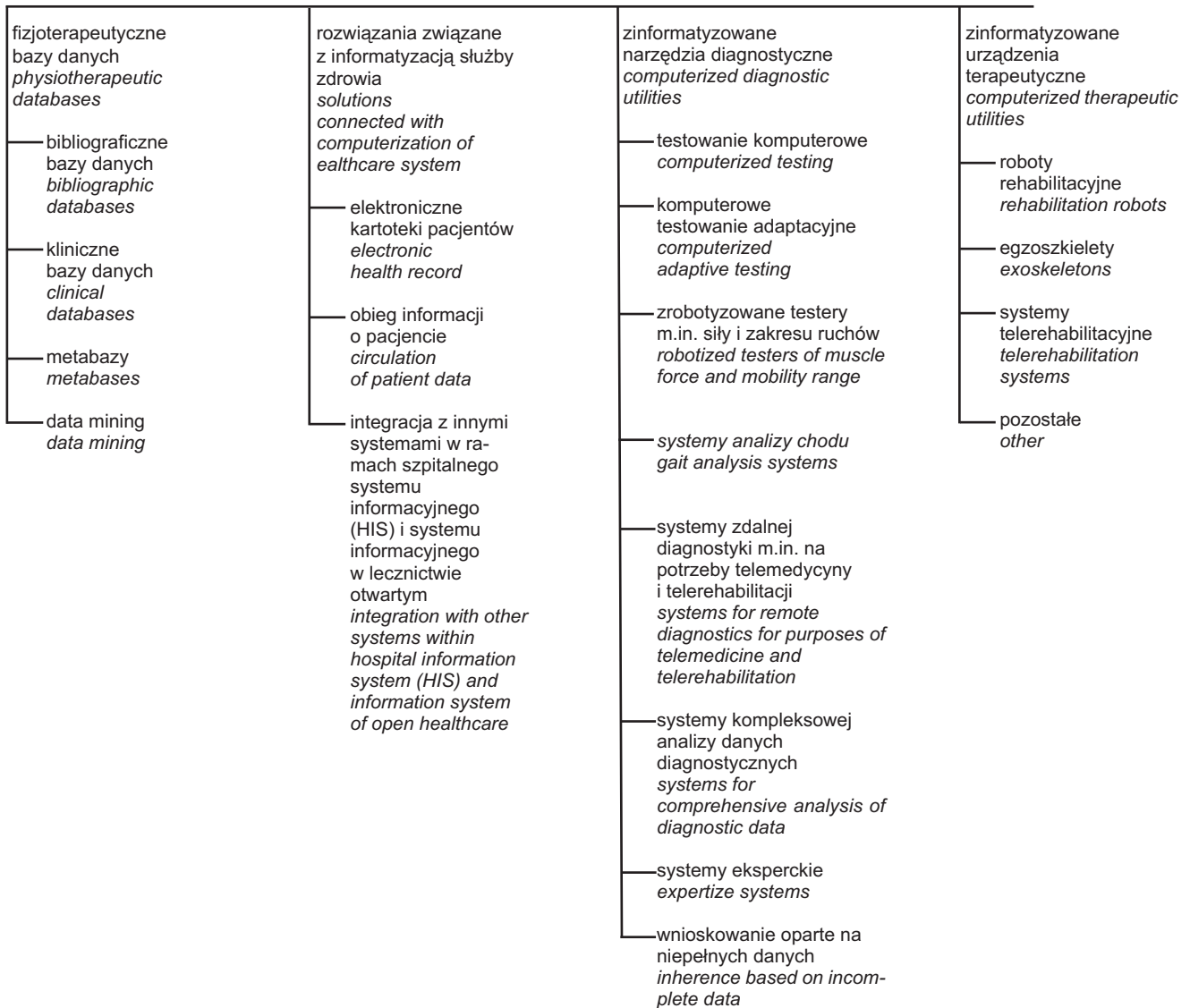
- omówione poniżej testowanie komputerowe (CT),
- komputerowe testowanie adaptacyjne (CAT), będące przedmiotem niniejszego artykułu,
- zrobotyzowane testery (siły, zakresów ruchów, czasów reakcji itp.), umożliwiające przeprowadzenie różnego rodzaju pomiarów i ich serii (w tym izokinetycznych, izotonicznych, izometrycznych, rozszerzonych o wykorzystanie biofeedbacku itp.) oraz generowanie różnych form raportów, dostosowanych do potrzeb klinicznych, co istotnie skraca czas przygotowania pacjenta i przeprowadzenia niezbędnych badań,
- systemy analizy chodu, od prostych, lecz efektywnych wykorzystujących internetowe arkusze kalkulacyjne [8] aż po połączone systemy kinematyczne/kinetyczne [8],
- systemy zdalnej diagnostyki, np. potrzeby telemedycyny, telerehabilitacji itp. o zróżnicowanym przeznaczeniu: od monitorowania (w czasie rzeczywistym lub

Computerization of medicine and especially of physiotherapy and rehabilitation has become a fact. It allows not only for implementation of EBM rules in clinical practice, but also for automatic creation of medical documentation and its comparison. It also generates increasing amounts of gathered data, which can be analyzed by means of data mining software for facilitation of medical interpretation or used by medical expertise and decision-aiding programs (so called “second opinion”) to increase the success rate of therapeutic actions.

Presented in Fig. 2, computerized diagnostic utilities used in physiotherapy, most of all include:

- computerized testing discussed below (CT),
- computerized adaptive testing (CAT), which is the subject of this paper,
- robotized testers (of strength, mobility range, response time etc.) for performing various kinds of measurements (isokinetic, isotonic, isometric, biofeedback etc.) and generating various forms of reports tailored for clinical purposes and needs, which significantly reduces time needed to prepare the patient and conduct necessary diagnostics,
- systems of gait analysis – from simple, yet effective solutions using online spread sheets [8] to sophisticated combined kinetic/kinematic systems [8],
- remote diagnostic systems (telemedicine, telerehabilitation etc.) of various use: from monitoring (in real time or not) execution of exercises at the patient’s home, through reception and interpretation of data from simple diagnostic tests by physiotherapist (automated measurement or videoed), to advanced telerehabilitation systems [9] or rehabilitation robots [10-12].

Wykorzystanie rozwiązań informatycznych w fizjoterapii
Application of IT solutions in physiotherapy



Ryc. 2. Zastosowania informatyki w fizjoterapii
Fig. 2. IT applications in physical therapy

nie) przebiegu ćwiczeń w domu pacjenta, poprzez odbiór i interpretację przez fizjoterapeutę danych z prostych testów diagnostycznych (o pomiarze zautomatyzowanym lub filmowanych) aż po złożone systemy telerehabilitacyjne [9] czy roboty rehabilitacyjne [10-12],

- systemy kompleksowej analizy danych diagnostycznych, w tym *data mining* (eksploracji danych), znajdujące ukryte powiązania w ramach dużych zbiorów danych klinicznych – szacuje się, że klasyczne formy wyszukiwania wiedzy w danych pozwalają na uzyskanie ok. 15% zawartej w nich informacji, stąd powiedzenie „*data rich, information poor*”, czyli „bogaty w dane, ubogi w wiedzę”,
- systemy eksperckie, w tym systemy tzw. „drugiej opinii”, stanowiące pomost pomiędzy wiedzą teoretyczną, wiedzą pozyskaną z ww. systemów *data mining* oraz praktyką kliniczną poprzez wspomaganie decyzji klinicystów, w tym diagnostycznych – systemy te, co bardzo ważne, najczęściej podają wiedzę „wprost” jako podpowiedź diagnostyczną do natychmiastowego wykorzystania, a nie jako dane wymagające dodatkowej interpretacji,

- systems of comprehensive analysis of diagnostic data including data mining (exploration) finding hidden correlations within large clinical data sets – it has been estimated that the conventional forms of exploration of data allow to extract approximately 15% of information the data contains, hence the saying “data rich, information poor”,
- expertise systems, including so called “second opinion” systems, which are a bridge connecting theoretical knowledge, knowledge obtained from the aforementioned data mining systems and the clinical practice in order to aid clinicians in the decision taking process – what is hugely important, these systems most often deliver knowledge directly as diagnostic advice for immediate use and not as data requiring additional interpretation,
- inference based on incomplete data – this process uses all available data providing for interference, deductive database structure, affiliation function and making correction to obtained outcome (f.e. Indicating special cases) to minimize risk of decisions taken for particular clinical cases, where ordinary generalization would not be sufficient.

- wnioskowanie oparte na niepełnych danych – wykorzystujące wszystkie dostępne dane w celu uwzględnienia interferencji, budowy baz reguł, funkcji przynależności oraz wprowadzania korekt do otrzymanych wyników (np. wskazując na przypadki szczególne), a przez to minimalizacji ryzyka decyzji podejmowanej w konkretnych przypadkach klinicznych, gdzie zwykła generalizacja doświadczenia może być niewystarczająca.

Komputerowe testy adaptacyjne w fizjoterapii

Standardowe testy stosowane w fizjoterapii, w tym klinimetryczne, składają się zwykle z zestawu na sztywno zadeklarowanych pytań (punktów do sprawdzenia). Każdorazowo wymagane jest zatem wyczerpanie kompletu pytań, bez względu na problem czy schorzenie pacjenta. Wczesne próby informatyzacji badań fizjoterapeutycznych obejmowały zatem głównie testowanie komputerowe (CT), związane z administrowaniem konwencjonalnych testów (zestawów pytań lub zadań do wykonania) poprzez komputer. Umożliwiało to przede wszystkim szybsze i automatyczne przetwarzanie uzyskanych w ten sposób danych. Przy obecnym poziomie informatyki możliwe jest jednak tworzenie testów automatycznie dostosowujących się do użytkownika. Dostosowanie możliwe jest zarówno w ilości i rodzaju pytań, jak i w obszarze zwiększania szczegółowości pytań w miarę uzyskiwania oczekiwanych odpowiedzi na kolejne pytania, co zwiększa rozdzielczość testu przy występowaniu zbliżonych objawów u poszczególnych pacjentów. Tworzone są więc badania zindywidualizowane, umożliwiające dostosowanie ich do konkretnego pacjenta. Zwiększona rozdzielczość takiego badania wpływa korzystnie nie tylko na sam przebieg terapii, ale i na motywację pacjenta, który zauważa odzwierciedlenie w wyniku testu nawet najmniejszych istotnych klinicznie zmian w swoim stanie funkcjonalnym.

Komputerowe testy adaptacyjne (ang. *computerized adaptive testing* – CAT) jako efektywne narzędzie mogą stanowić znaczące wsparcie fizjoterapeuty i całego zespołu wielospecjalistycznego w terapii poszczególnych schorzeń. Umożliwiają one prowadzenie badań w czasie rzeczywistym z możliwie najmniejszą liczbą punktów/pytań skrojonych indywidualnie pod kątem oceny poziomu funkcjonalnego określonego pacjenta. Ocena może być zwykle prowadzona w czterech obszarach: czynności złożonych, kończyn górnych, kończyn dolnych i tułowia. Tzw. przegląd pacjenta umożliwia odzwierciedlenie istotnych zmian stanu funkcjonalnego poprzez test o liczbie punktów/pytań zredukowanych nawet do 4-5 [13, 14].

Zastosowania kliniczne CAT

Idea CAT wywodzi się z psychologii i pedagogiki [15, 16], jednak stopniowo przechodzi do innych dziedzin nauki, w tym medycyny. Na chwilę obecną z informatyzowane testy w fizjoterapii znajdują się dopiero na początku swojego rozwoju, jednak już znalazły zastosowanie w badaniu równowagi [17], kończyny górnej [18], kończyny dolnej [19, 20] oraz schorzeniach takich jak zapalenie stawów [21]. Badania te szczegółowo przedstawiono w tabeli 1.

Efektywność pomiarów CAT powoduje, że rozwija się cała grupa pomiarów CAT-based functional status (CAT-FS). Trwają również badania nad możliwościami przetwarzania i interpretacji wyników takich badań, żeby optymalizować ich przydatność dla klinicystów w codziennej praktyce i zwiększyć efektywność fizjoterapii [13, 22-24]. CAT daje również możliwość podwyższenia efektywności samooceny wyników leczenia (ang. *patient-reported outcome* – PRO) w codziennej praktyce klinicznej. Metodyka konstruowania i wdrożenia CAT została zawarta w kolejnych etapach tego procesu (tab. 2):

Computerized Adaptive Testing in physiotherapy

Standard tests used in physiotherapy, including the clinimetric tests, often consist of a set of rigidly declared questions (points to check). Each time a test is taken, it is necessary to complete the set of questions without regard to the patient's problem or disease. The early attempts at computerization of psychotherapeutic tests mainly concerned computerized testing (CT) related to administration of the conventional tests (sets of questions or tasks to complete) with the use of computers. Most of all, it allowed quicker and automated processing of data gathered in this manner. However, at present the advancement of informatics enables development of tests that automatically adapt themselves to the user. Adaptation is possible in terms of the quantity and type of questions as well as of increasing the level of detail along with getting the expected answers to successive questions, which boosts the test's resolution in the case of occurrence of similar symptoms in particular subjects. This results in individualized testing that can be tailored to suit a particular patient. Increased resolution of such test has a positive impact not only on the course of the therapy itself, but also on the motivation of the patient, who is able to notice even the smallest clinically significant changes in his functional status in the outcome of the test.

Computerized Adaptive Tests (CAT) as an effective tool may become a significant aid to the work of physiotherapists and other specialists in the therapy of particular diseases. CAT enables real time testing with the least possible number of questions tailored to assess the functional status of a given patient. The assessment is usually conducted at four levels: compound activities, upper limbs, lower limbs and the trunk. So called patient's review reflects the significant changes of the functional status through a test reduced to even 4 or 5 questions/tasks [13, 14].

Clinical application of CAT

The idea of CAT derives from psychology and pedagogics [15, 16], however, gradually it is being transferred to other fields of science, including medicine. At present, the computerized tests in physiotherapy have just began their development, yet they have found use in examinations of balance [17], upper limbs [18], lower limbs [19, 20] and arthritis [21]. These examinations are presented in Table 1.

The effectiveness of CAT measurements cause development of an entire group – CAT-based functional status (CAT-FS). There are also studies on possibilities of processing and interpretation of scores of such tests, so as to optimize their usability for clinicians in their everyday practice and increase the effectiveness of physiotherapy [13, 22-24]. CAT also provides a way of improving the effectiveness of self-evaluation of therapy outcomes (PRO – patient-reported outcome) in everyday clinical practice. Methodology of construction and introduction of CAT is pointed out below (Tab. 2):

Tabela 1. Wybrane badania z zakresu zastosowania CAT w rehabilitacji i fizjoterapii
Table 1. Selected studies concerning CAT application In rehabilitation and physical therapy

Badanie Study	Omówienie wyników Discussion of findings
Badanie równowagi I-Ping i wsp. 2010 [17] Balance test I-Ping et al 2010 [17]	Badanie wykonywało 5 oceniających w 6 szpitalach na 764 pacjentach po udarze, zarówno przebywających w szpitalu, jak i w opiece ambulatoryjnej i domowej. Wykonanie testu Balance computerized adaptive testing (Balance CAT) zajmuje średnio jedynie 83 sekundy tj. 18% czasu trwania Berg Balance Scale (BBS) przy porównywalnej dokładności. Należy jednak zaznaczyć, że dostosowanie BBS do wymagań CAT wymagało usunięcia 9 pytań spośród 41. <i>Test conducted on 764 patients after stroke from 6 hospitals, hospitalized or under ambulatory or home care. Performance of Balance CAT takes 83 sec. on average, which is approx. 18% of Berg Balance Scale (BBS) duration with similar accuracy. It should be underlined that adaptation of BBS to CAT standards required elimination of 9 among 41 questions.</i>
Badanie kończyny górnej Hart i wsp. 2010 [18] Upper limb CAT test Hart et al. 2010 [18]	Analizie poddano wyniki badań CAT w zakresie funkcji kończyny górnej (ramienia) u 30 987 pacjentów w 518 szpitalach w USA. Średni czas trwania testu CAT składającego się z 6 pytań wyniósł mniej niż 2 minuty. Proponowany test CAT udowodnił swoją efektywność, dokładność i wrażliwość. <i>Upper limb (arm) CAT scores of 30987 patients from 518 hospitals in USA. Mean duration of CAT test consisting of 6 questions was less than 2 minutes. Suggested CAT test proved to be effective, accurate and sensitive.</i>
Badanie kończyny dolnej Hart i wsp. 2008 [19] Lower limb test Hart et al. 2008 [19]	Badanie kohortowe w zakresie funkcji kończyny dolnej (biodra) u 9 158 pacjentów w 257 szpitalach w USA. Proponowany test CAT składający się z 7 pytań udowodnił swoją efektywność i powtarzalność oraz sprawdza się w warunkach klinicznych. Wskazano jednak potrzebę większego zróżnicowania pytań w teście CAT. <i>Cohort study of the lower limb function (hip) in 9158 patients from 257 hospitals in USA. Suggested CAT test consisting of 7 questions proved to be effective and repetitive in clinical conditions. A necessity of higher differentiation of CAT questions was indicated.</i>
Badanie kończyny dolnej Hart i wsp. 2008 [20] Lower limb test Hart et al. 2008 [20]	Analizie poddano wyniki badań CAT w zakresie funkcji kończyny dolnej (kolana) u 21 896 pacjentów w 291 szpitalach w USA. Proponowany test CAT składający się z 7 pytań udowodnił swoją efektywność i powtarzalność oraz sprawdza się w warunkach klinicznych. <i>Analysis of lower limb function (knee) CAT scores of 21896 patients from 291 hospitals in USA. Suggested test consisting of 7 questions proved to be effective and repetitive in clinical conditions.</i>
Zapalenie stawów Fries i wsp. 2009 [21] Arthritis Fries et al. 2009 [21]	Analizie poddano test CAT Patient-Reported Outcomes Measurement Information (PROMIS) opracowany na bazie Health Assessment Questionnaire Disability Index (HAQ). W trakcie prac zbiór pytań zmniejszono z 1865 do 124, a następnie, po procedurze kalibracji z udziałem 21 000 pacjentów, zwiększono do 199 w celu zwiększenia dokładności. W porównaniu czterech różnych narzędzi diagnostycznych test oparty na CAT okazał się najlepszy. <i>Analysis of CAT PROMIS (Patient-Reported Outcomes Measurement Information designed on the basis of Health Assessment Questionnaire Disability Index (HAQ)). The set of questions was reduced from 1865 to 124 and following calibration increased to 199 to improve accuracy. CAT test prove to be the best in a comparison of 4 different diagnostic utilities.</i>

- budowa uporządkowanego zbioru (banku) pytań,
- przebadanie stworzonego zbioru pytań na próbie pacjentów (kalibracja),
- selekcja pytań,
- sprawdzenie, czy wszystkie pytania podlegają wpływowi określonej cechy dominującej,
- kalibracja zbioru pytań do modelu wyniku zadania testowego (ang. *item response theory model* – IRT model) [25],
- ocena równoważności pytań w podgrupach,
- budowa banku pytań po kalibracji,
- skonstruowanie CAT,
- badanie pilotażowe CAT [14].

Należy przy tym pamiętać, że w pełni adaptacyjna implementacja CAT wymaga:

- ustalenia zasad wyboru pierwszego pytania dla danego badanego pacjenta,
- szacowania poziomu wyników badanego pacjenta na podstawie jego wyników przy poprzednich pytaniach testu,
- ustalenia zasad wyboru kolejnego pytania z banku pytań na podstawie oszacowanego poziomu wyników,
- ustalenia zasad zakończenia procedury CAT po osiągnięciu przez pacjenta określonego wyniku,
- ustalenia sposobu obrazowania i wnioskowania na wymienionej wcześniej podstawie.

- creation of an ordered set (bank) of questions,
- examining of the created set of questions on a population of patients (calibration),
- selection of questions,
- checking whether all of the questions are under impact of the dominant characteristic,
- calibration of the set of question to the item response theory model – IRT model [25],
- assessment of equivalence of questions in subgroups,
- recreation of the question bank after calibration,
- creation of CAT,
- CAT pilot examination [14].

It should also be underlined that entirely adaptive implementation of CAT requires the following:

- to work out the rules of selection of the first question for a given patient,
- estimation of the score level of the examined patient on the basis of their scores from the previous questions of the test,
- to work out rules of selection of the next question from the question bank on the basis of the estimated score level,
- to work out the rules of completion of the CAT procedure after the patient has achieved the assumed result,
- to work out the mode of imaging and inference basing on the afore-mentioned points.

Tabela 2. Wybrane etapy konstruowania CAT (jeden z możliwych wariantów działania) [21, 26, 27]
Table 2. Selected stages of CAT preparation (one of possible modes of operation) [21, 26, 27]

Nazwa etapu Stage name	Element Element	Charakterystyka elementu Element characteristics
<p>Założenie wyjściowe: dla określonej cechy mierzonej za pomocą testu da się wprowadzić taką zmienną ukrytą (jedną lub więcej), że na jej podstawie da się jednoznacznie ustalić wektor badanych parametrów (wynik testu) dla danego pacjenta. <i>Initial assumption: for a given trait measured by the test it is possible to introduce a hidden variable (one or more), so it is possible to define the vector of the examined parameters (score) for a given patient.</i></p>		
<p>Budowa uporządkowanego zbioru (banku) pytań – w celu otrzymania jednowymiarowego modelu monotonicznego teorii odpowiadania na pozycje testowe (ang. Item Response Theory – IRT) <i>Creation of an ordered set (bank) of items (questions) – in order to obtain a single dimension monotone theoretic model of response to items (Item Response Theory – IRT)</i></p>	<p>Liczba pytań (pozycji testowych, ang. items) w banku <i>Quantity of items in the question bank</i></p>	<p>Liczba pytań zależy silnie od celu i rodzaju testu, liczby możliwych odpowiedzi na jedno pytanie itp., dla dużych testów przyjmuje wartość z zakresu 100-200 <i>Number of items strongly depends on goal and type of test, number of possible answers to a given question etc.; 100-200 items assumed quantity for big tests</i></p>
	<p>Charakterystyka pytań w banku <i>Questions characteristics</i></p>	<p>Idealny test zawiera dużą liczbę silnie różniących pytań dla każdego możliwego wyniku (od najniższego do najwyższego) testu <i>Ideal test contains a large number of strongly differentiated questions for every possible test score (from lowest to highest)</i></p>
	<p>Analiza statystyczna banku pytań (parametrów tym dla pojedynczej osoby) <i>Statistical analysis of the question bank (parameters for individual patient)</i></p>	<p>Dobór parametrów technicznych CAT, w tym zmiennej ukrytej oraz krzywej charakterystycznej pozycji testowej przekształcającej ją w wynik testu, decydujący m.in. o trudności i zdolności rozdzielczej testu <i>Selection of CAT parameters including hidden variable and characteristic curve of test position transforming it into test score, which determines the difficulty and resolution of the test</i></p>
<p>Kalibracja <i>Calibration</i></p>	<p>Oszacowanie parametrów modelu <i>Estimation of model's parameters</i></p>	<p>Badanie za pomocą testu n pacjentów losowo wybranych z populacji <i>Examination of a number of patients randomly chosen from a population</i></p>
	<p>Analiza wyników i strojenie modelu <i>Score analysis and model tuning</i></p>	<p>Złożone zadanie obliczeniowe, wymagające wykorzystania specjalistycznego oprogramowania <i>Complex computing task requiring special software</i></p>
<p>Konstrukcja testu: dobór pytań <i>Test construction: Selection of question/items</i></p>	<p>Zakres i dokładność pomiaru badanej cechy <i>Range and accuracy of measurement of a given trait</i></p>	<p>Określenie zakresu, w jakim ma być mierzona badana cecha oraz dokładności tego pomiaru, tj. określenie tzw. docelowej funkcji informacji testu <i>Defining range, within which the tested trait is to be measured and accuracy of that measurement i.e. defining the target function of test information</i></p>
	<p>Dobór pytań <i>Selection of questions</i></p>	<p>Optymalizacja: wybór takiego wariantu testu z jak najmniejszą liczbą pytań, dla którego funkcja informacji testu jest jeszcze nie mniejsza od docelowej funkcji informacji testu <i>Optimization: selection of a variation of the test with a least possible number of questions, for which the information function is still not lower from the target test information function</i></p>

c.d. tab. 2 ze strony 64

<p>Ustalenie zasad automatycznego doboru pytań podczas testu oraz zakończenia testu <i>Setting out of rules for automatic selection of test items and test completion</i></p>	<p>Osiągnięcie, przy danych (nieznanych wcześniej) odpowiedziach pacjenta, wyniku (wartości cechy badanej) z zadanego zakresu i o zadanej dokładności <i>Obtaining, with given (previously unknown) patient's answers, of a score (value of tested trait) from a given range and accuracy</i></p>	<p>Określenie wartości startowej danej cechy badanej (np. rozpoczęcie od wartości średniej dla danej populacji) i wymaganej dokładności, a następnie, począwszy od wartości startowej, dynamiczny dobór kolejnego pytania dopasowanego do poziomu cechy z poprzednich pytań w sposób zwiększający dokładność pomiaru (tj. dający najwyższą wartość funkcji informacji w danym punkcie). <i>Determining of the initial value of a given trait (f.e. starting from the mean value for a given population) and required accuracy and then beginning from the zero value, a dynamic selection of the next question adapted to the level of trait from previous questions in a way that increases accuracy of measurement (i.e. giving the highest value of the information function at a given point).</i></p>
	<p>Zakończenie testu <i>Test completion</i></p>	<p>Po osiągnięciu określonej założonej dokładności pomiar jest kończony i określany jest wynik. <i>Having achieved target accuracy measurement is complete and its score is defined.</i></p>

Zgodnie z ideą CAT prawidłowo kontynuowane badanie powinno zawierać kolejne pytania odpowiadające aktualnemu oszacowanemu poziomowi wyników pacjenta oraz uszczegóławiać go (tzn. każde kolejne pytanie powinno dostarczać maksimum informacji). Pozwala to osiągnąć zakończenie testu (diagnozę) już po 30-40% standardowej liczby pytań stosowanej w badaniu konwencjonalnym.

Pełna wykonalność ww. badania przy odpowiednio dużym zbiorze możliwych pytań diagnostycznych właściwych dla dużych grup pacjentów jest możliwa tylko przy pełnym wsparciu informatycznym procedur CAT. Oznacza to jednocześnie, że każdy pacjent może mieć różną liczbę pytań w badaniu, przy niezmięnionej jego wiarygodności.

Słabym punktem CAT jest wrażliwość na stosowane metody estymacji oraz wielkości prób, na podstawie których zostały one oszacowane [26]. Jedyną metodą na wyeliminowanie tego problemu jest dołożenie należytej staranności na etapie definiowania i konstruowania testu CAT.

Podsumowanie

Szersze wykorzystanie CAT w fizjoterapii wymaga rozstrzygnięcia szeregu kwestii.

Testy CAT (oparte na IRT) dają, w porównaniu z tzw. klasyczną teorią testów, pełniejszy opis własności hipotetycznej cechy determinującej odpowiedzi uzyskane w teście [22], co przekłada się nie tylko na wzrost wiarygodności, ale również optymalizację procesu badania, w tym skrócenie czasu, liczby pytań, zmniejszenie kosztów itp. Możliwe jest również porównanie wyników z dwóch testów badających tę samą cechę lub nawet badanie, tzw. zróżnicowania funkcjonowania pozycji testowych (ang. *differential item functioning* – DIF), tj. stronniczości dla wybranych podgrup testowych [27]. Będą to prawdopodobnie jedne z głównych przyczyn stopniowego wypierania tradycyjnych testów (w tym testów klinimetrycznych stosowanych w fizjoterapii) przez testy CAT. Wpisuje się to nie tylko w rozwój narzędzi diagnostycznych, ale również w paradygmat EBM i informatyzację służby zdrowia i integrację jej systemów informatycznych.

Following the CAT concept, a continued test should include successive questions reflecting the actual estimated level of patient's scores and provide new details (that is, each subsequent questions should deliver maximum of information). It allows to achieve completion of the test (diagnosis) only after 30-40% of the standard number of questions used in the conventional tests.

Full feasibility of the afore-mentioned test with a suitably large set of possible diagnostic questions specific for large populations of patients is only possible with full IT support of the CAT procedures. At the same time, this means that each patient may get a different number of questions in the test, yet its reliability remains unchanged.

CAT's weak spot is its vulnerability to the applied methods of estimation and size of samples, upon which the estimates have been made [26]. The only way to eliminate this handicap is taking adequate care at the stage of defining and designing of a CAT test.

Summary

Broader use of CAT in physiotherapy requires settling of a number of matters presented below.

CAT tests give (based on IRT), in comparison with the conventional theory of tests, a more complete description of a hypothetical of a feature determining answers obtained in the test [22], which has reflection not only in the increase of reliability, but in optimization of the examination process resulting in shortened duration, lower number of questions, reduction of costs etc. It is also possible to compare scores of two tests that examine the same characteristic or even examining differential item functioning (DIF) that is the bias for selected test subgroups [27]. These will be probably the main reasons for gradual displacement of conventional tests (including clinimetric tests used in physiotherapy) by CAT. This inscribes itself not only in the development of diagnostic utilities, but also in the EBM paradigm and computerization of the health care system and integration of its computer systems.

Ze względu na stopień złożoności testów CAT, w tym od strony informatycznej i statystycznej, wydaje się słuszne zapewnienie ich wdrażania we współpracy klinicystów z informatykami medycznymi i statystykami. Należy w tym miejscu przypomnieć o konieczności dochowania należytej staranności na etapie definiowania i konstruowania testu CAT. Utrudnia to rzecz jasna dokonywanie samodzielnych zmian i prowadzenie badań w kierunku rozwoju testów CAT, niemniej jednak możliwe jest tworzenie w tym obszarze – badawczych zespołów interdyscyplinarnych, podobnych do zespołów tworzonych do modelowania procesów neurofizjologicznych, w których uczestniczy autor niniejszej pracy [28].

W celu zapewnienia pełnej użyteczności testów CAT, w tym dla pacjentów administrujących je samodzielnie, należy dołożyć starań w celu obudowania omawianych testów przyjaznymi interfejsami użytkownika, ułatwiającymi nie tylko wypełnienie testu, ale również jego analizę i interpretację, np. poprzez odpowiednie zobrazowanie na tle poprzednich wyników czy średniej w populacji. Automatyczne przesyłanie i wprowadzanie wyników testu do właściwych baz danych na obecnym etapie informatyzacji służby zdrowia stanowi jedynie kwestię wyboru rozwiązań technicznych w tym zakresie.

Pomimo udowodnionej (w wybranych zastosowaniach – przykłady w tab. 2) przydatności klinicznej testów CAT, niezwykle ważna pozostaje kwestia klinicznej interpretacji ich wyników [22-24, 29]. Kluczowa wydaje się potrzeba uzgodnienia minimalnej istotnej klinicznie zmiany wyniku (wyników) dla każdego ze stosowanych testów CAT, w tym uzgodnienie wymienionych wcześniej. wartości na forum międzynarodowym, co zapewni porównywalność wyników badań. Wypracowanie wskazówek i rekomendacji klinicznych w tym zakresie pozwoli na jednolitą interpretację wyników CAT oraz efektywne porównywanie ich z tradycyjnymi odpowiednikami.

Należy zaznaczyć, że omówione zastosowania nie są jedynymi możliwymi. Cały czas trwają prace nad rozszerzeniem zastosowań CAT na inne schorzenia, wszystkie grupy wiekowe [30] oraz na zastosowania w rehabilitacji i fizjoterapii ambulatoryjnej i domowej [31]. Niezwykle ciekawe są również badania nad integracją rozwiązań i ich rezultatami w codziennej praktyce klinicznej [32]. Wskazują one, że możliwe i efektywne jest automatyczne gromadzenie i przetwarzanie danych diagnostycznych (w tym z CAT) oraz integrowanie ich z elektroniczną kartoteką pacjenta (ang. *Electronic Health Record* – EHR) i innymi systemami informatycznymi w służbie zdrowia. Oprócz poprawienia i przyspieszenia obiegu informacji pozwala to automatycznie tworzyć podstawę do całościowej analizy i oceny wyników terapii (w klasyfikacji według schorzeń, grup wiekowych, stanu wyjściowego pacjenta, stopnia poprawy, oddziałów szpitalnych i placówek służby zdrowia, itp.) oraz oceny jakości usług i satysfakcji pacjenta. Wyzwaniem jest odpowiednie przygotowanie pacjentów, personelu medycznego i technicznego oraz infrastruktury technicznej i oprogramowania, żeby informatyzacja nie stanowiła utrudnienia procesu terapeutycznego, a wspomagała go również przez pozbawienie personelu medycznego obciążeń związanych z wypełnianiem dużej ilości formularzy i pozwoliła zaoszczędzony czas przeznaczyć na efektywną terapię. Wymaga to stworzenia rozwiązań CAT bezpośrednio pod kierunkiem samych terapeutów, aby spełnione zostały wymagania końcowego użytkownika.

Jak wiele zostało tu do zrobienia pokazuje model opracowany przez Villamzara i wsp. [33]. Pozwala on ocenić obecne możliwości placówki fizjoterapeutycznej w zakresie ilości i rodzaju przeprowadzanych zabiegów przy określonej liczbie zasobów sprzętowych i ludzkich oraz przebiegu procedur terapeutycznych, a także przeprowadzić szereg symulacji w tym zakresie, uwzględniających rekonfigurację w zakresie liczby jednostek sprzętowych, ich rozmieszczenia, zwiększenia/zmniejszenia liczby fizjoterapeutów czy zmiany rozkładu

Due to high degree of statistical and informatic complexity of CAT, it seems reasonable to ensure CAT introduction as collaboration of clinicians with medical IT specialists and statisticians. It should be underlined that it is necessary to take due care at the stage of defining and designing a CAT test. This, of course, makes it more difficult to introduce individual changes and conduct studies on the development of CAT. Nevertheless, it is still possible to create interdisciplinary teams similar to the teams created for modelling of neurophysiological processes, in one of which the author herself participates [28].

In order to ensure full feasibility of CAT tests, also for patients who administer the tests themselves, efforts should be made to encase the discussed tests with friendly user interfaces, which will make it easier to fill in the test, analyze and interpret it through adequate imaging against previous scores or the mean value in the population. At the present stage of computerization of the health care system, automatic transmission and coding of the test scores to proper database is only a matter of choosing one of many available technical solutions.

Despite confirmed and verified clinical usability (examples in table 2) of CAT, the matter of clinical interpretation of CAT scores is still of utmost importance [22-24, 29]. The need for an official and international accord on the minimum clinically significant change of score (scores) for each of the applied CAT tests, which will ensure comparability of the scores, is crucial for the use CAT in physiotherapy. Elaboration of guidelines and clinical recommendations in that field will allow uniform interpretation of CAT scores and their effective comparison with conventional equivalents.

It should be emphasized that the discussed applications are not the only ones possible. Studies on extending the use of CAT for other diseases, age groups [30] and use in rehabilitation and ambulatory or home physiotherapy are being conducted at present [31]. Also the studies on integration of the above-mentioned solutions and their results with daily clinical practice are very interesting [32]. They indicate that automatic gathering, processing of diagnostic data (including CAT) and its integration with the patient's electronic health record (EHR) and other informatic system in health care is possible and effective. Apart from improvement and acceleration of circulation of information, it allows to create automatically the basis for holistic analysis and evaluation of therapeutic results (classified according to disease, age group, patient initial status, degree of improvement, hospital and other health care institutions departments) as well as rating of the quality of service and patients' satisfaction. What is true challenge, is the proper preparation of patients, medical and technical personnel as well as technical and software infrastructure, so that computerization does not become an obstacle of the therapeutic process and helps the personnel through elimination of burdens of filling in numerous forms to save time that can be spent on therapeutic actions. It enforces designing CAT solutions in collaboration with the therapists, so that the end user requirements are met.

The amount of work that is left to be done can be illustrated on a model elaborated by Villamzara et al. [33]. This model allows to evaluate present capacity of a psychotherapeutic institution in terms of quantity and type of performed procedures with regard to available number of human and equipment resources and type of performed procedures, and also conduct a variety of simulations in this range, which take into account reconfiguration of the number of equipment units, their placement, increasing/decreasing the number of psychotherapeutic personnel or changing the time schedule of particular procedures. It enables optimization of the following elements: usage of equipment, personnel or duration of pa-

czasowego poszczególnych procedur. Umożliwia to zoptymalizowanie parametrów w pożądanym kierunku: wykorzystania sprzętu, fachowego personelu czy też skrócenia pobytu pacjenta. Należy jednak przy tym uwzględnić również czas niezbędny na przygotowanie pacjenta (por. wodolecznictwo czy krioterapia), personelu (zapoznanie z kartą pacjenta itp.) czy miejsca i aparatury do poszczególnych zabiegów. Kryterium nadrzędnym, bez względu na wyliczenia nawet najdoskonalszych programów komputerowych, musi pozostać efektywność fizjoterapii oraz jej bezpieczeństwo dla pacjenta i personelu. W wielu bowiem sytuacjach wykorzystanie technik informatycznych pozwala na przyspieszenie analiz oraz uzyskanie szeregu wyników dla różnych założeń w celu dalszego porównania, ale decyzję musi podjąć człowiek.

tients' stay. However, one should also take into consideration the time necessary for preparation of patients (comp. hydrotherapy or cryotherapy), personnel (becoming acquainted with patient's medical record etc.), place and equipment for particular procedures. However, regardless of the calculations of even the most advanced software, the effectiveness of therapy and patients and personnel safety remain a superior criterion. Often, the use of computer techniques enables acceleration of analysis and obtaining of various results for further comparison, yet the decision has to be taken by the human.

Piśmiennictwo References

- [1] Mikołajewska E. *Dominujące trendy we współczesnej rehabilitacji*. Niepełnosprawność i Rehabilitacja, 2010, 1, 87-102.
- [2] Mikołajewska E., Mikołajewski D. *EBM w fizjoterapii – wykorzystanie zasobów internetowych*. Rehabilitacja Praktyczna, 2008, 4, 50-52.
- [3] Mikołajewska E. *Medycyna oparta na faktach w fizjoterapii*. Valetudinaria, 2007, 2, 88-91.
- [4] Oostendorp R. A. B., Nijhuis – van der Sanden M. W. G., Heerkens Y. F. i wsp. *Rehabilitacja medyczna i fizjoterapia oparte na wiarygodnych i aktualnych publikacjach – ocena krytyczna*. Rehabilitacja Medyczna, 2008, 1, 9-15.
- [5] Płaszewski M. *Praktyka oparta na dowodach – zasady i kierunki rozwoju Evidence Based Medicine w fizjoterapii*. Rehabilitacja Medyczna, 2006, 10, 1, 1-8.
- [6] Sherrington C., Herbert R. D., Maher C. G. i wsp. *PE-Dro: A database of randomized trials and systematic reviews in physiotherapy*. Manual. Ther., 2000, 5, 223-226.
- [7] MEDLINE/PubMed (U.S. National Library of Medicine) <http://www.ncbi.nlm.nih.gov/pubmed> – data pobrania 20.06.2011 r.
- [8] Mikołajewska E. *Analiza chodu pacjentów po udarze mózgu – rozwiązanie własne*. Udar Mózgu Problemy Interdyscyplinarne, 2010, 1-2, 20-26.
- [9] Mikołajewska E., Mikołajewski D. *Telerehabilitacja*. Rehabilitacja w Praktyce, 2011, 1, 64-67.
- [10] Mikołajewski D., Mikołajewska E. *Roboty rehabilitacyjne*. Rehabilitacja w Praktyce, 2010, 4, 49-53.
- [11] Mikołajewska E. *Lokomat jako element nowoczesnej reedukacji chodu*. Praktyczna Fizjoterapia i Rehabilitacja, 2010, 10, 15-18.
- [12] Mikołajewski D., Mikołajewska E. *Nowoczesne rozwiązania informatyczne dla osób niepełnosprawnych*. Niepełnosprawność i Rehabilitacja. 2011, 2, 93-113.
- [13] Chakravarty E. F., Bjorner J. B., Fries J. F. *Improving patient reported outcomes using item response theory and computerized adaptive testing*. J. Rheumatol., 2007, 34, 6, 1426-1431.
- [14] Walker J., Böhnke J. R., Cerny T. i wsp. *Development of symptom assessments utilising item response theory and computer – adaptive testing – a practical method based on a systematic review*. Crit. Rev. Oncol. Hematol., 2010, 73, 1, 47-67.
- [15] Van der Linden W. J., Glas C. A. W. (red.) *Computerized adaptive testing: Theory and practice*. Kluwer Academic Publishers, Norwell-Dordrecht 2000.
- [16] Wainer H., Green B. F., Flaugher R. (red.) *Computerized adaptive testing: a primer*. Mallory International, Aylesbeare 2000.
- [17] I-Ping H., Jyun-Hong Ch., Chun-Hou W. i wsp. *Development of a computerized adaptive test for assessing balance function in patients with stroke*. Phys. Ther., 2010, 90, 1336-1344.
- [18] Hart D. L., Wang. Y.-C., Cook K. F. i wsp. *A computerized adaptive test for patients with shoulder impairments produced responsive measures of function*. Phys. Ther., 2010, 90, 928-938.
- [19] Hart D. L., Wang Y. C., Stratford P. W. i wsp. *A computerized adaptive test for patients with hip impairments produced valid and responsive measures of function*. Arch. Phys. Med. Rehabil., 2008, 89, 11, 2129-2139.
- [20] Hart D. L., Wang Y. C., Stratford P. W. i wsp. *Computerized adaptive test for patients with knee impairments produced valid and responsive measures of function*. J. Clin. Epidemiol., 2008, 61, 11, 1113-1124.
- [21] Fries J. F., Cella D., Rose M. i wsp. *Progress in assessing physical function in arthritis: PROMIS short forms and computerized adaptive testing*. J. Rheumatol., 2009, 36, 9, 2061-2066.
- [22] Wang Y. C., Hart D. L., Stratford P. W. i wsp. *Clinical interpretation of computerized adaptive test-generated outcome measures in patients with knee impairments*. Arch. Phys. Med. Rehabil., 2009, 90, 8, 1340-1348.
- [23] Wang Y. C., Hart D. L., Stratford P. W. i wsp. *Clinical interpretation of computerized adaptive test outcome measures in patients with foot/ankle impairments*. J. Orthop. Sports Phys. Ther., 2009, 39, 10, 753-764.
- [24] Wang Y. C., Hart D. L., Werneke M. i wsp. *Clinical interpretation of outcome measures generated from a lumbar computerized adaptive test*. Phys Ther., 2010, 90, 9, 1323-1335.
- [25] Konarski R. *Model cechy latantnej w analizie psychometrycznej testów i pozycji testowych*, [w:] B. Niemierko, H. Szaleniec (red.) *Diagnostyka edukacyjna. Standardy wymagań i normy testowe w diagnostyce edukacyjnej*. PODE, Kraków 2004.
- [26] Kondratek B. *Klasyczna teoria testów a teoria odpowiedzi na pozycje testowe. Teoretyczne i empiryczne porównanie rozwiązań dla pozycji ocenianych dychotomicznie*. Praca magisterska, Uniwersytet Śląski, Katowice 2007.
- [27] Rose M., Bjorner J. B., Becker J. i wsp. *Evaluation of a preliminary physical function item bank supported the expected advantages of the Patient-Reported Outcomes Measurement Information System (PROMIS)*. J. Clin. Epidemiol., 2008, 61, 1, 17-33.
- [28] Duch W., Nowak W., Meller J. i wsp. *Consciousness and attention in autism spectrum disorders*. Proceedings of Cracow Grid Workshop, 2010, 202-211, 2011.

- [29] Wang Y. C., Hart D. L., Stratford P. W. i wsp. *Clinical interpretation of a lower-extremity functional scale-derived computerized adaptive test*. Phys. Ther., 2009, 89, 9, 957-968.
- [30] Haley S. M., Fragala-Pinkham M. A., Dumas H. M. i wsp. *Evaluation of an item bank for a computerized adaptive test of activity in children with cerebral palsy*. Phys. Ther., 2009, 89, 6, 589-600.
- [31] Jette A. M., Haley S. M., Tao W. i wsp. *Prospective evaluation of the AM-PAC-CAT in outpatient rehabilitation settings*. Phys. Ther., 2007, 87, 4, 385-398.
- [32] Deutscher D., Hart D. L., Dickstein R. i wsp. *Implementing an integrated electronic outcomes and electronic health record process to create a foundation for clinical practice improvement*. Phys. Ther., 2008, 88, 2, 270-285.
- [33] Villamizar J. R., Coelli F. C., Pereira W. C. A. i wsp. *Discrete-event computer simulation methods in the optimisation of a physiotherapy clinic*. Physiotherapy, 2011, 97, 1, 71-77.

Adres do korespondencji:
Address for correspondence:

Emilia Mikołajewska
Klinika Rehabilitacji
10 Wojskowy Szpital Kliniczny z Polikliniką SP ZOZ
ul. Powstańców Warszawy 5
85-681 Bydgoszcz
e-mail: e.mikolajewska@wp.pl
www: e.mikolajewska.prv.pl

Wpłynęło/Submitted: VII 2011
Zatwierdzono/Accepted: VI 2011