

*Piotr Brzyski, Tomasz Knururowski, Beata Tobiasz-Adamczyk*

TRAFNOŚĆ I RZETELNOŚĆ  
KWESTIONARIUSZA OCENY OGÓLNEGO STANU ZDROWIA SF-20  
W POPULACJI OSÓB W PODESZŁYM WIEKU

Zakład Socjologii Medycyny  
Katedra Epidemiologii i Medycyny Zapobiegawczej  
Collegium Medicum, Uniwersytet Jagielloński w Krakowie  
Kierownik Katedry: Beata Tobiasz-Adamczyk

*W pracy omówiono trafność i rzetelność testu SF-20 i jego przydatność dla oceny w warunkach polskich stanu zdrowia osób w podeszłym wieku. Analizowano wywiady zebrane wśród 529 losowo wybranych mieszkańców Krakowa w podeszłym wieku. Ocenę trafności teoretycznej testu dokonano za pomocą eksploracyjnej analizy czynnikowej metodą głównych składowych. Ocenę rzetelności testu określono poprzez zbadanie wewnętrznej spójności testu i analizę wartości współczynnika alpha Cronbacha.*

*Słowa kluczowe: trafność, rzetelność, test SF-20, osoby w podeszłym wieku*

*Key words: validity, reliability, SF-20 test, elderly*

WSTĘP

Coraz częstsze posługiwanie się, tak w działalności klinicznej jak i w medycznych badaniach populacyjnych, wielowymiarowymi skalami zdrowia nasuwa pytanie o poprawność metodyczną ich stosowania. Istotnym problemem staje się ocena użyteczności wskaźników szeroko rozumianego stanu zdrowia, szczególnie tych, które wyykają się tradycyjnej medycznej diagnozie, takich jak zdolność pełnienia ról społecznych, zakres interakcji społecznych czy poziom satysfakcji z życia (1, 2).

Medyczny model zdrowia koncentruje się na wskaźnikach negatywnych, co sprowadza się do postawienia diagnozy przez lekarza, posługującego się w większości obiektywnymi kryteriami ustalonymi na podstawie wieloletnich doświadczeń medycyny. Definicja zdrowia WHO (3, 4) stawia wyzwania dotyczące pomiaru każdego z jego wymiarów. Stosunkowo prosty jest pomiar wymiaru fizycznego, nadal niedoskonałe są jednak metody mierzące pozostałe wymiary: stan funkcjonalny oraz zdrowie psychiczne i społeczne.

Próba kompleksowego podejścia do zdrowia stało się stworzenie skal, składających się z kilkunastu/kilkudziesięciu pytań, oceniających różne aspekty stanu zdrowia, pozwalających stwierdzić nie tylko fakt występowania pewnego zjawiska lub jego brak, ale także uzyskać dane o natężeniu zarówno ujemnych (dolegliwości i stany chorobowe oraz ich

konsekwencje) jak i dodatnich cech stanu zdrowia, (np. zdolność wykorzystania możliwości, jakie daje stan pełnego zdrowia lub zdolność do normalnego funkcjonowania pomimo ograniczeń wynikających z odczuwanych dolegliwości). Takie rozwiązanie zrodziło jednak kolejny problem: czy stworzona skala mierzy to zjawisko, które ma mierzyć – wskazując na różne aspekty trafności skali oraz potrzebę ustalenia, stopnia dokładności, z jaką skala oddaje naturalne różnicowanie badanego zjawiska tj. określenia rzetelności skali.

Większość stosowanych obecnie w medycynie skal powstała w określonym kręgu kulturowym, przy użyciu pojęć, znaczeń i symboli charakterystycznych dla języka, w którym została ona stworzona. Bezkrytyczne zastosowanie skali pomiarowej w innym środowisku może dać zafałszowane wyniki, których źródłem mogą być różne (w zależności od warunków kulturowych) wskaźniki, używane do oceny tych samych zjawisk lub różne znaczenia tych samych symboli i wartości. Wskaźnik będący w jednej kulturze trafnym wyznacznikiem posiadania lub nasilenia pewnej cechy, w innej może być zupełnie niezrozumiały lub nawet posiadać przeciwne znaczenie. Konieczne zatem staje się określenie, jakie walory narzędzia spełniającego dobrze swoją rolę w warunkach, w których zostało stworzone, zostają zachowane przy jego adaptacji do zastosowania w nowych okolicznościach i jakich zabiegów adaptacyjnych wymaga jego zastosowanie w innym środowisku.

Określenie przydatności narzędzia pomiarowego w nowych warunkach wymaga zbadania, w możliwie szerokim zakresie, jego trafności i rzetelności. Rzetelność jest miarą tego, w jakim stopniu wynik testu oddaje rzeczywistą wartość badanej cechy, a w jakim pomiar ten jest zakłócony przez błąd pomiaru pochodzący z różnych źródeł (5, 6). Trafność testu jest zagadnieniem o wiele bardziej złożonym, z tego względu, iż wyróżniamy 3 aspekty trafności: trafność kryterialną, treściową i teoretyczną. Trafność kryterialna najczęściej jest rozumiana jako zbieżność wyników testu z innymi kryteriami oceny danego zjawiska, zarówno testowymi jak i nietestowymi (np. z diagnozą lekarską). Trafność treściowa (wewnętrzna) oznacza fakt, że test składa się z pozycji stanowiących reprezentatywną próbę dla uniwersum zmiennych, które w jakimś stopniu opisują mierzone zjawisko. Najistotniejszym aspektem trafności jest jednak trafność teoretyczna, czyli związek narzędzia pomiarowego z konstruktem teoretycznym (zmienną teoretyczną), którą ten test ma mierzyć (niektórzy badacze utożsamiają trafność treściową z trafnością teoretyczną) (5).

Kwestionariusz Oceny Ogólnego Stanu Zdrowia SF-20 jest szeroko stosowanym w krajach Europy Zachodniej testem o uznanej i wysokiej trafności i rzetelności, zarówno w populacji ogólnej jak i w populacji osób w podeszłym wieku. Brak w literaturze polskiej prac na temat rzetelności i trafności testów, służących do oceny ogólnego stanu zdrowia osób starszych, zrodził potrzebę przeprowadzenia badań, których wyniki zostaną poniżej przedstawione.

#### CEL PRACY

Celem pracy jest ocena trafności i rzetelności testu SF-20 w populacji osób w podeszłym wieku. Uzasadnieniem podjęcia pracy jest obserwowany wzrost populacji osób w starszym wieku objętych opieką geriatryczną, gdzie narzędzie to znajduje zastosowanie, zwłaszcza w kontekście oceny zapotrzebowania na świadczenia medyczne i opiekuńcze osób w tej kategorii wiekowej.

## MATERIAŁ I METODY

Ocenę trafności i rzetelności testu SF-20 przeprowadzono w badaniu obejmującym grupę 529 mieszkańców Krakowa w wieku 65–85 lat, mieszkających w swoim środowisku. Średni wiek badanych wynosił 72,7 lat; 59,5% badanych stanowiły kobiety, 21,9% respondentów posiadało wykształcenie pomaturalne i wyższe.

Polska wersja testu utworzona została na podstawie przekładu na język polski dokonanego niezależnie przez dwóch tłumaczy, w oparciu o kwestionariusz wykorzystywany w holenderskich badaniach dotyczących ludzi starych (projekt Nestor, prowadzony przez Northern Center for Health Care Research, University of Groningen), których rozwinięciem były badania prowadzone w ramach sieci Central European Network. Polska wersja kwestionariusza została następnie przetłumaczona z powrotem na język angielski, w celu określenia zgodności polskiej wersji testu z wersją stosowaną w badaniach Nestor.

Test SF-20 składa się ze skal oceniających zdrowie psychiczne (Mental Health – MH) i fizyczne (Physical Functioning – PF) oraz percepcję zdrowia (Health Perception – HP) i ograniczenia w pełnieniu ról społecznych (Role Functioning – RF), a także pytań dotyczących relacji społecznych (Social Functioning – SF) oraz odczuwania bólu (Pain Perception – PP) (7). Treść pytań i ich podział na skale oraz możliwe odpowiedzi przedstawia tabela I.

W celu zbadania trafności teoretycznej testu poszczególne skale wchodzące w jego skład zostały poddane analizie czynnikowej metodą głównych składowych – oczekiwano, że zmienne tworzące skalę będą korelowały na poziomie conajmniej 0,7 (niektórzy badacze dopuszczają wartość 0,6) z pierwszą główną składową, reprezentującą teoretyczny konstrukt, który mierzy skala. Następnie dokonano analizy rzetelności metodą badania wewnętrznej spójności skali – wymaga się, aby korelacje poszczególnych zmiennych z sumarycznym wynikiem skali były wyższe niż 0,4 – kryterium Kline'a (8) oraz aby badane skale charakteryzowały się wartością współczynnika alpha Cronbacha wyższą niż 0,7 – warunek ten nosi nazwę kryterium Nunnally'ego. Jakkolwiek w literaturze można znaleźć skale o wartości alfa większej od 0,5, jednak niektórzy badacze uważają je za narzędzia wątpliwej jakości (9, 10).

## WYNIKI

Analiza głównych składowych, dotycząca skali oceniającej zdrowie psychiczne (MH), wyodrębniła jedną składową o wartości własnej większej od 1 (tab. II). Koreluje ona na poziomie od 0,76 do 0,84 z czterema z pięciu zmiennych wchodzących w skład skali. Wyjątkiem jest zmienna MH4 – „Przez jaki okres czasu w ciągu ostatniego miesiąca czuł/a się Pan/i szczęśliwy/a?” – wykazująca z tym czynnikiem korelację  $r$  równą 0,54 i znacznie silniej korelująca z drugą składową wyodrębnioną przez procedurę ( $r=0,84$ ). Procent wariancji wyjaśnianej przez drugi czynnik – 16% – jest jednak wyraźnie niższy niż przez pierwszy (58%).

Wszystkie pytania, z wyjątkiem MH4, spełniają kryterium Kline'a. Wartość współczynnika alfa Cronbacha dla skali wynosi 0,81. Wspomniana zmienna zaniża rzetelność skali: po jej usunięciu wartość współczynnika alfa dla skali skonstruowanej z pozostałych zmiennych jest wyższa niż w przypadku oryginalnej skali (tab. II).

Jeśli analizie głównych składowych poddamy skalę dotyczącą ograniczeń w aktywności fizycznej (PF) otrzymamy dwie składowe o wartościach własnych większych od 1: z pierw-

Tabela I. Treść pytań tworzących skale wchodzące w skład testu SF-20.  
Table I. Questions and scales of SF-20 test.

Nr pyt.	Treść pytania
MH1	Jak często w ciągu ostatniego miesiąca czuł/a się Pan/i bardzo zdenerwowany/a?
MH2	Przez jaki okres czasu w ciągu ostatniego miesiąca czuł/a Pan/i wewnętrzny spokój?
MH3	Przez jaki okres czasu w ciągu ostatniego miesiąca czuł/a się Pan/i przygnębiony/a i był/a w złym nastroju?
MH4	Przez jaki okres czasu w ciągu ostatniego miesiąca czuł/a się Pan/i szczęśliwy/a?
MH5	Jak często w ciągu ostatniego miesiąca czuł/a się Pan/i tak przygnębiony/a, że nic nie mogło Pana/i rozweselić?
	Czy stan Pana/i zdrowia ograniczył w ostatnich czterech tygodniach możliwość wykonywania następujących czynności:
PF1	Czynności wymagających użycia siły (jak np. podnoszenie ciężkich przedmiotów, bieganie lub wykonywanie ćwiczeń siłowych)?
PF2	Czynności umiarkowanie ciężkich (jak np. przesuwanie stołu, niesienie zakupów)?
PF3	Chodzenie pod górę lub wchodzenie po schodach?
PF4	Zginanie się, podnoszenie przedmiotów lub schyłanie się?
PF5	Chodzenie bez odpoczynku?
PF6	Jedzenie, ubieranie się, kąpiel lub skorzystanie z toalety?
HP1	Czy może Pan/i powiedzieć, że ogólnie stan Pana/i zdrowia jest:
HP2	Czuję się trochę chory/a.
HP3	Jestem tak zdrowy/a jak nikt inny.
HP4	Moje zdrowie jest doskonałe.
HP5	Ostatnio czułem/am się źle.
PP	Czy w ciągu ostatnich 4 tygodni odczuwał/a Pan/i ból?
RF1	Czy stan Pana/i zdrowia uniemożliwia podjęcie pracy zawodowej lub wykonywanie prac domowych (niemożność podjęcia i wykonania pracy)?
RF2	Czy z powodu stanu zdrowia był/a Pan/i niezdolny/a do wykonania pewnego rodzaju lub ilości pracy lub zajęć domowych (ograniczenia w wykonywaniu pewnych zajęć)?
SF	Jak często w ciągu ostatniego miesiąca, ze względu na zdrowie ograniczał/a Pan/i swoje życie towarzyskie (np. wizyty u przyjaciół, bliskich, sąsiadów)?

#### Możliwe odpowiedzi na poszczególne pytania:

**MH1 – MH5, SF:** 1. cały czas, 2. większość czasu, 3. sporo czasu, 4. przez jakiś czas, 5. przez krótki okres czasu, 6. w ogóle nie.

**PF1 – PF6:** .....1. ograniczenia od ponad 3 miesiące, 2. ograniczenia od 3 miesięcy lub krócej, 3. żadnych ograniczeń.

**HP1:** .....1. doskonały, 2. bardzo dobry, 3. dobry, 4. średni, 5. słaby.

**HP2 – HP5:** .....1. zdecydowanie tak, 2. tak, 3. nie jestem pewien/na, 4. nie, 5. zdecydowanie nie.

**PP:** .....1. żadnego bólu, 2. bardzo słaby, 3. słaby, 4. umiarkowany, 5. silny.

**RF1 – RF2:** .....1. ograniczenia od ponad 3 miesięcy, 2. ograniczenia od 3 miesięcy lub krócej, 3. żadnych ograniczeń.

Pytania **MH1, MH3, MH5, HP1, HP3, HP4** zostały przekodowane tak, aby wyższa ranga odpowiedzi odpowiadała lepszemu zdrowiu.

Tabela II. Macierz nierotowanych głównych składowych oraz wyniki analizy wewnętrznej spójności dla skali dotyczącej zdrowia psychicznego.

Table II. Matrix of no rotated principal components and results of internal consistency analysis for mental health scale.

Zmienna	Składowa		Korelacja pytanie – skala	Alfa po usunięciu pytania
	1-sza*	2-ga*		
MH2	0,84	-0,12	0,70	0,73
MH3	0,84	-0,09	0,69	0,73
MH1	0,78	-0,14	0,62	0,76
MH5	0,76	-0,22	0,59	0,77
MH4	<b>0,54</b>	0,84	<b>0,39</b>	<b>0,83</b>

\* kolumny zawierają korelacje zmiennych wchodzących w skład skali z głównymi składowymi wyodrębnionymi przez procedurę głównych składowych. We wszystkich tabelach dotyczących analizy głównych składowych pokazano zmienne posortowane wg wartości ładunków czynnikowych na pierwszej głównej składowej, pogrubiono wartości niespełniające odpowiednich kryteriów, pogrubioną kursywą zaznaczono wyniki spełniające łagodniejsze kryteria lub wyniki posiadające inne niepożądane cechy, pomimo spełnienia wspomnianych wcześniej kryteriów.

szą z nich wysoko korelują cztery z sześciu zmiennych należących do skali, osiągając korelacje na poziomie od 0,76 do 0,81 (tab. III). Dwie pozostałe zmienne posiadają korelacje zbliżone wartościami zarówno z pierwszą jak i z drugą składową, przy czym tylko jedna z nich – PF6 – „Czy stan Pana/i zdrowia ograniczył w ostatnich czterech tygodniach możliwość wykonywania następujących czynności: jedzenie, ubieranie się, kąpiel lub skorzystanie z toalety?” – posiada z pierwszą główną składową korelację wyższą niż 0,6. Wariancja zestawu zmiennych tłumaczona przez poszczególne składowe wynosi odpowiednio 52 i 17%.

Korelacje pytań tworzących skalę z jej sumarycznym wynikiem są wyższe niż 0,4 a wartość alfa Cronbacha dla skali wynosi 0,81. W przypadku zmiennych PF1 – „Czy stan Pana/i zdrowia ograniczył w ostatnich czterech tygodniach możliwość wykonywania czynności wy-

Tabela III. Macierz nierotowanych głównych składowych oraz wyniki analizy wewnętrznej spójności dla skali dotyczącej zdrowia fizycznego.

Table III. Matrix of no rotated principal components for physical functioning scale and results of internal consistency analysis.

Zmienna	Składowa		Korelacja pytanie – skala	Alfa po usunięciu pytania
	1-sza	2-ga		
PF2	0,81	0,11	0,69	0,75
PF5	0,78	-0,32	0,63	0,77
PF3	0,76	0,34	0,64	0,76
PF4	0,76	-0,13	0,61	0,77
PF6	<b>0,62</b>	-0,58	0,46	<b>0,80</b>
PF1	<b>0,54</b>	0,66	<b>0,41</b>	<b>0,81</b>

magających użycia siły? (jak np. podnoszenie ciężkich przedmiotów, bieganie lub wykonywanie ćwiczeń siłowych).” i PF6, zmiana wartości alfa po usunięciu tych zmiennych ze skali jest tak nieznaczna, że sugeruje rozważenie rezygnacji z którejsz z tych zmiennych lub zastąpienie zmienną silniej skorelowaną z pierwszą główną składową, reprezentującą teoretyczny konstrukt, który ma mierzyć skala (tab. III).

W przypadku skali zawierającej pytania dotyczące percepcji zdrowia (HP) z pierwszą składową wyodrębnioną przez procedurę analizy głównych składowych (tylko ona posiada wartość własną większą od 1) wszystkie zmienne korelują na poziomie od 0,66 do 0,81 (tylko zmienna HP5 – „Ostatnio czułem/am się źle” – posiada korelację ze składową niższą niż 0,7). Składowa ta tłumaczy 57% całkowitej wariancji zestawu zmiennych (tab. IV).

Table IV. Macierz nierotowanych głównych składowych oraz wyniki analizy wewnętrznej spójności dla skali dotyczącej percepcji zdrowia.

Table IV. Matrix of no rotated principal components for health perception scale and results of internal consistency analysis.

Zmienna	Składowa		Korelacja pytanie – skala	Alfa po usunięciu pytania
	1-sza	2-ga		
HP2	0,81	0,28	0,67	0,70
HP1	0,78	0,26	0,64	0,73
HP3	0,77	-0,49	0,57	0,74
HP4	0,74	-0,54	0,54	0,75
HP5	<b>0,66</b>	0,52	0,51	<b>0,77</b>

Wszystkie pytania wchodzące w skład skali spełniają kryterium Kline’a, a rzetelność skali mierzona alfą Cronbacha wynosi 0,78. Analizując wartość współczynnika alfa po usunięciu poszczególnych pytań okazuje się, że nie wzrasta ona po usunięciu ze skali żadnego z pytań, co potwierdza poprawną konstrukcję tej skali i akceptowalną rzetelność tej części testu. Jedynie mała zmiana wartości alfa po usunięciu zmiennej HP5 sugeruje poszukiwanie zmiennej, która mogłaby ją zastąpić w konstrukcji tej skali i podnieść jej rzetelność (tab. IV).

W przypadku skali dotyczącej pełnienia ról społecznych obie zmienne wchodzące w jej skład wykazują korelację  $r=0,87$  z jedyną składową główną o wartości własnej większej od 1. Korelacje tych zmiennych z sumarycznym wynikiem skali wynoszą 0,51 (obie zmienne spełniają kryterium Kline’a) a rzetelność skali mierzona wartością współczynnika alfa Cronbacha jest równa 0,67.

Analiza głównych składowych przeprowadzona dla wszystkich zmiennych wchodzących w skład testu wyodrębniła 4 główne składowe posiadające wartości własne większe od 1 (tab. V). Pierwsza z wydzielonych składowych charakteryzuje się najwyższymi korelacjami ze zmiennymi należącymi do skali zdrowia fizycznego (PF), lecz nie są to korelacje jednakowo silne (od 0,40 do 0,72). Ze składową tą – którą możemy interpretować jako fizyczny wymiar zdrowia – na podobnym poziomie jak większość zmiennych ze skali zdrowia fizycznego, korelują także zmienne należące do skali percepcji zdrowia (HP – korelacje na poziomie 0,57 do 0,70), pełnienia ról społecznych (RF – korelacje od 0,61 do 0,69) oraz odczuwanie bólu (PP –  $r=-0,58$ ). Również zmienna mierząca funkcjonowanie w społe-

czeństwie (SF) wykazuje wyższą korelację z pierwszą składową niż z pozostałymi, lecz znacznie niższą od wcześniej wspomnianych zmiennych ( $r=0,42$ ).

Opisana powyżej struktura czynnikowa sugeruje, że ocena różnych wymiarów zdrowia u osób starszych zdominowana jest przez ich sprawność fizyczną i odczuwanie bólu (czyli percepcję fizycznych dolegliwości). Także istniejące ograniczenia w codziennej aktywności w różnych sferach życia (zarówno w pracy zawodowej, czynnościach domowych, jak i w życiu towarzyskim) są w głównej mierze konsekwencjami niedomagań w fizycznym wymiarze zdrowia.

Drugą składową można interpretować jako psychiczny wymiar zdrowia, gdyż najwyżej korelują z nią zmienne należące do skali dotyczącej tej sfery zdrowia człowieka (MH – korelacje na poziomie od 0,61 do 0,74). Wyjątkiem jest tutaj zmienna MH4 – „Przez jaki okres czasu w ciągu ostatniego miesiąca czuł/a się Pan/i szczęśliwy/a?” – ( $r=0,31$ ), korelująca na nieco wyższym poziomie z fizyczną składową zdrowia ( $r=0,42$ ). Fakt ten sugeruje, że zmienna ta, choć jest związana ze zdrowiem psychicznym człowieka, w przypadku osób starszych mierzy nieco inny jego wymiar, a zdolność odczuwania szczęścia, w rozumieniu osób starszych, może być nie tylko oznaką zdrowia psychicznego, ale przede wszystkim miarą jakości życia warunkowanej stanem zdrowia, a w szczególności sprawnością fizyczną.

Trzecia wyodrębniona składowa główna wysoko koreluje tylko z tymi zmiennymi należącymi do skali dotyczącej ograniczeń w aktywności fizycznej (PF), które wykazywały najniższe korelacje z pierwszą główną składową – są to zmienne PF5 – „Czy stan Pana/i zdrowia ograniczył w ostatnich czterech tygodniach możliwość wykonywania następujących czynności: chodzenie bez odpoczynku?” – i PF6 – „Czy stan Pana/i zdrowia ograniczył w ostatnich czterech tygodniach możliwość wykonywania następujących czynności: jedzenie, ubieranie się, kąpiel lub skorzystanie z toalety?”.

Czwarta główna składowa wykazuje najwyższe korelacje ze zmiennymi: HP5 – „Ostatnio czułem/am się źle” – oraz PP – „Czy w ciągu ostatnich 4 tygodni odczuwał/a Pan/i ból?” sugerując, że w przypadku osób starszych odczuwanie bólu i związane z tym ogólne samopoczucie, aczkolwiek związane z fizyczną sferą zdrowia, odgrywa bardzo ważną rolę w ocenie ogólnego stanu zdrowia i może aspirować do rangi odrębnego, bardzo istotnego jego wymiaru (tab. V).

## DYSKUSJA

Przeprowadzona analiza potwierdziła, że rzetelność i wewnętrzna spójność skal wchodzących w skład testu SF-20 jest wystarczająca do zastosowania ich do oceny ogólnego stanu zdrowia osób starszych w warunkach polskich. Rzetelność badanych skal jest wprawdzie niższa niż uzyskana przez twórców testu: dla skali zdrowia psychicznego alfa Cronbacha wynosi 0,81 wobec 0,88 uzyskanego przez autorów testu, dla skali zdrowia fizycznego alfa równa jest 0,81 wobec 0,86, dla skali percepcji zdrowia alfa równa się 0,78 wobec 0,87 i dla skali pełnienia ról społecznych alfa wynosi 0,67 wobec 0,81 otrzymanego przez twórców testu (7). Wszystkie skale, z wyjątkiem skali dotyczącej pełnienia ról, spełniają kryterium Nunnally'ego wymagające, aby narzędzie uznać za dobrą skalę, gdy wartość alfa Cronbacha przekracza 0,7 (9), należy jednak zauważyć, że niższa rzetelność wspomnianej skali może wynikać z małej ilości pytań wchodzących w jej skład. Należy także zaznaczyć, że twórcy testu uzyskali powyższe wyniki w populacji nieograniczającej się do grupy osób w podeszłym wieku.



Tabela V. Macierz nierotowanych głównych składowych dla testu SF-20.

Table V. Matrix of no rotated principal components for SF-20 test.

		Składowa			
		1	2	3	4
MH1	Jak często w ciągu ostatniego miesiąca czuł/a się Pan/i zdenerwowany/a?	0,42	<b>0,65</b>	-0,20	
MH2	Przez jaki okres czasu w ciągu ostatniego miesiąca czuł/a Pan/i wewnętrzny spokój?	0,43	<b>0,74</b>		
MH3	Przez jaki okres czasu w ciągu ostatniego miesiąca czuł/a się Pan/i przygnębiony/a i był/a w złym nastroju?	0,55	<b>0,61</b>		0,17
MH4	Przez jaki okres czasu w ciągu ostatniego miesiąca czuł/a się Pan/i szczęśliwy/a?	<b>0,42</b>	0,31	-0,10	0,30
MH5	Jak często w ciągu ostatniego miesiąca czuł/a się Pan/i tak przygnębiony/a, że nic nie mogło Pana/i rozweselić?	0,41	<b>0,64</b>	0,17	0,13
	Czy stan Pana/i zdrowia ograniczył w ostatnich czterech tygodniach możliwość wykonywania następujących czynności:				
PF1	Czynności wymagających użycia siły (jak np. podnoszenie ciężkich przedmiotów, bieganie lub wykonywanie ćwiczeń siłowych)?	<b>0,60</b>	-0,23	-0,28	
PF2	Czynności umiarkowanie ciężkich (jak np. przesuwanie stołu, niesienie zakupów)?	<b>0,72</b>	-0,17	0,31	0,15
PF3	Chodzenie pod górę lub wchodzenie po schodach?	<b>0,71</b>	-0,28	0,10	0,15
PF4	Zginanie się, podnoszenie przedmiotów lub schylanie się?	<b>0,62</b>	-0,18	0,42	
PF5	Chodzenie bez odpoczynku?	0,56	-0,13	<b>0,60</b>	0,13
PF6	Jedzenie, ubieranie się, kąpiel lub skorzystanie z toalety?	0,41		<b>0,59</b>	
HP1	Czy może Pan/i powiedzieć, że ogólnie stan Pana/i zdrowia jest:	<b>0,68</b>			-0,23
HP2	Czuję się trochę chory/a.	<b>0,70</b>	-0,11	-0,17	-0,25
HP3	Jestem tak zdrowy/a jak nikt inny.	<b>0,61</b>	-0,27	-0,38	0,20
HP4	Moje zdrowie jest doskonałe.	<b>0,57</b>	-0,23	-0,49	0,27
HP5	Ostatnio czułem/am się źle.	<b>0,66</b>		-0,14	-0,50
PP	Czy w ciągu ostatnich 4 tygodni odczuwał/a Pan/i ból?	<b>-0,58</b>	0,13	0,15	0,48
RF1	Czy stan Pana/i zdrowia uniemożliwia podjęcie pracy zawodowej lub wykonywanie prac domowych (niemożność podjęcia i wykonania pracy)?	<b>0,61</b>	-0,20	-0,21	0,19
RF1	Czy z powodu stanu zdrowia był/a Pan/i niezdolny/a do wykonania pewnego rodzaju lub ilości pracy lub zajęć domowych (ograniczenia w wykonywaniu pewnych zajęć)?	<b>0,69</b>	-0,26		0,14
SF	Jak często w ciągu ostatniego miesiąca, ze względu na zdrowie ograniczał/a Pan/i swoje życie towarzyskie (np. wizyty u przyjaciół, bliskich, sąsiadów)?	<b>0,42</b>	0,18	0,14	-0,33



Otrzymane wartości potwierdzają akceptowalną jakość testu jako narzędzia służącego ocenie zdrowia osób starszych. Jednak analiza głównych składowych pokazuje, że pewne zmienne wchodzące w skład testu wykazują zbyt niskie korelacje z mierzonymi czynnikami, aby mogły nadal stanowić z pozostałymi zmiennymi skalę o wysokiej wartości – zatracona została częściowo trafność teoretyczna testu. Wynika to prawdopodobnie z faktu, iż wraz z wiekiem i związanymi z nim zmianami i procesami zachodzącymi w życiu człowieka, zmienia się rozumienie pewnych pojęć, spojrzenie na siebie i otaczający świat. Także różnice kulturowe pomiędzy Polską a USA, gdzie test został skonstruowany mogą być powodem zaobserwowanych niedoskonałości testu. Aby precyzyjniej zlokalizować przyczynę obniżonej rzetelności testu należy zbadać jego trafność i rzetelność w ogólnej populacji mieszkańców Polski, niezawężonej do określonej grupy wiekowej.

#### WNIOSKI

1. Kwestionariusz Oceny Ogólnego Stanu Zdrowia SF-20 zastosowany w populacji osób w podeszłym wieku wykazuje się akceptowalną trafnością i rzetelnością.
2. Rzetelność testu w polskiej populacji osób starych jest niższa niż określona przez jego autorów w populacji, na której była badana jego trafność i rzetelność przez twórców testu.
3. W populacji osób starszych ujawniają się wpływy czynników kulturowych oraz związanych z wiekiem, co wymaga dostosowania testu do warunków charakteryzujących tę populację.
4. Zaleca się dalsze badania skal należących do testu, które charakteryzowały się niższą trafnością w niniejszych badaniach, aby otrzymać narzędzie charakteryzujące się równie wysokimi parametrami jak oryginalna wersja testu.

*P Brzyski, T Knurowski, B Tobiasz-Adamczyk*

#### VALIDITY AND RELIABILITY OF SHORT FORM HEALTH SURVEY SF-20 IN POPULATION OF ELDERLY PEOPLE

#### SUMMARY

Aim of this study was assessing validity and reliability of Short Form General Health Survey (SF-20) and its usefulness in evaluation of health status of elderly people in Polish conditions. Study group consisted of 529 randomly chosen elderly (household residents) living in Krakow. Theoretical validity of the test was evaluated using exploratory factor analysis (principal components method). Reliability was calculated by Cronbach's alpha index and by checking if questions forming each sub scale fulfill Kline's criterion.

Analysis showed some faults in theoretical validity of scales concerning mental and physical health, which decrease their reliability. Nevertheless, scales constituting this Survey fulfill Kline's and Nunnally's criterias, which means that they have high internal consistency.

Results confirm satisfactory value of SF-20 test as an instrument for evaluation of health of elderly people in Polish conditions.

#### PIŚMIENNICTWO

1. Bowling A., Measuring disease. Open University Press, Buckingham, Philadelphia 1995.
2. Bowling A., Measuring health A review of quality of life measurement scales. Open University Press, Buckingham, Philadelphia 1999.

3. World Health Organization, The First Ten Years. The Health Organization. Geneva, World Health Organization, 1958.
4. World Health Organization, Targets for Health for All by the Year 2000. Copenhagen, World Health Organization, Regional Office for Europe, 1985.
5. Brzeziński Z., Metodologia badań psychologicznych. Wydawnictwo Naukowe PWN, Warszawa 2002.
6. McDowell I., Newell C., Measuring health. Oxford University Press, New York, Oxford, 1996.
7. Stewart AL., Hays RD, Ware JE jr, The MOS Short-Form General Health Survey: reliability and validity in a patient population. *Med. Care* 1988;26:724–35.
8. Kline P. A handbook of test construction. London, Methuen, 1986.
9. Nunnally J., Psychometric Theory. New York, McGraw-Hill, 1978.
10. Bowling A., Research methods in health. Investigating health and health services. Open University Press, Buckingham, Philadelphia 2002.

**Adres autorów:**

Piotr Brzyski

Zakład Socjologii Medycyny, Katedra Epidemiologii i Medycyny Zapobiegawczej,

Collegium Medicum, Uniwersytet Jagielloński

ul. Kopernika 7, 31-034 Kraków

Tel.: (12) 422-31-82 w. 13

Fax: (12) 422-31-82 w. 16

e-mail: mylisy@cyf-kr.edu.pl