



Badanie rzetelności skal oceny nasilenia depresji w populacji polskiej: skali melancholii Becha-Rafaelsena, skali depresji Hamiltona oraz skali Montgomery-Åsberg

*Assessment of depression scales reliability in the Polish population:
the Bech-Rafaelsen Melancholia Scale, the Hamilton Depression Scale,
and the Montgomery-Åsberg Scale*

JOANNA MAZUREK, ANDRZEJ KIEJNA, KRZYSZTOF MAŁYSZCZAK,
AGNIESZKA STĘPIEŃ

Z Katedry i Kliniki Psychiatrii AM we Wrocławiu

STRESZCZENIE. *Celem pracy była ocena rzetelności nie stosowanego dotąd w Polsce narzędzia pomiaru depresji – skali melancholii Becha-Rafaelsena w porównaniu z bardziej znanymi skalami: Hamiltona i Montgomery-Åsberg. Badanie przeprowadzono w grupie 34 pacjentów z różnymi kategoriami zaburzeń depresyjnych. Ich stan oceniali dwóch klinicystów, posługujących się wymienionymi trzema narzędziami oraz prostą skalą CGI. Wyniki potwierdziły, że zarówno skala melancholii, jak i pozostałe wykorzystane w badaniu narzędzia są instrumentami rzetelnymi, różniąc się natomiast pod względem wewnętrznej spójności.*

SUMMARY. *The aim of the study was to compare an instrument for depression severity assessment – the Melancholia Scale by Bech-Rafaelsen – which has not been used in Poland so far, with two other more popular depression scales, developed by Hamilton and Montgomery-Åsberg. Subjects in the study were 34 patients with different categories of depressive disorder. Their state was assessed by two clinicians using the three above-mentioned tools and a simple GCI scale. Both the Melancholia Scale and the other two instruments were found to be reliable, but differing in internal consistency.*

Słowa kluczowe: depresja / skala melancholii Becha-Rafaelsena / skala Hamiltona / skala Montgomery-Åsberg / rzetelność

Key words: depression / the Bech-Rafaelsen Melancholia Scale / the Hamilton Depression Scale / the Montgomery-Åsberg Scale / reliability

Zaburzenia depresyjne stanowią problem, z którym nie tylko psychiatrzy, ale i lekarze innych specjalności stykają się coraz częściej. Ich rozpowszechnienie w populacji ogólnej Europy i Ameryki Północnej jest duże – wynosi 4–6% [9, 14], mimo to rozpoznanie depresji, a także określenie jej nasilenia wciąż bywa problemem klinicznym. Oszacowanie głębokości depresji jest jednak istotne w niektórych przypadkach, np. dla

oceny skuteczności terapii. Klasyfikacja ICD-10 definiuje trzy stopnie nasilenia depresji, co może być wystarczające dla niektórych zastosowań. Dokładniejsza ocena stanu pacjenta wymaga stosowania odpowiednich skal psychometrycznych.

Jakość pomiarów wykonywanych za pomocą skal psychometrycznych określa się za pomocą pojęć rzetelności i trafności. Rzetelność określa, w jakim stopniu pomiary

dostarczają dokładnej i powtarzalnej wariancji. Trafność określa, w jakim stopniu pomiar jest użyteczny.

SPOSOBY OKREŚLANIA RZETELNOŚCI NARZĘDZI PSYCHOMETRYCZNYCH

W psychiatrii rzetelność jest określana poprzez wyznaczanie hipotetycznej wartości idealnej za pomocą szeregu pomiarów, w przybliżeniu powtarzalnych. Najczęściej stosowana jest metoda *test-retest* polegająca na wykonaniu pomiaru w dwóch różnych punktach czasu. Pomimo teoretycznej atrakcyjności metoda *test-retest* na wiele sposobów odstaje od hipotetycznego ideału. Z jednej strony drugi pomiar jest często zakłócony przez biologiczne, psychologiczne i społeczne zmiany dotyczące osoby badanej. Takie systematyczne zmiany powodują zwiększenie wariancji błędu, co oznacza niedoszacowanie wartości rzetelności. Z drugiej strony respondent może pamiętać swoje poprzednie odpowiedzi i powtarzać je, próbując być „dobrym”, a wówczas oszacowana wartość rzetelności może być zbyt duża.

Inna metoda określania rzetelności polega na jednoczesnym wykonywaniu pomiaru przez dwóch lub więcej badaczy podczas tego samego badania (*inter-rater reliability*). Wartość rzetelności oznaczona w badaniu *inter-rater* jest zazwyczaj większa niż w badaniu *test-retest*. Wynika to z faktu, że warunki pomiaru i sposób przeprowadzenia badania typu *inter-rater* jest jednakowy dla wszystkich pomiarów, co powoduje zawyżenie otrzymanej wartości rzetelności.

Należy wspomnieć także o metodach pozwalających określić rzetelność na podstawie szeregu pojedynczych (niepowtarzalnych) pomiarów. Zakłada się, że pozycje skali psychometrycznej związane są z tą samą mierzoną wielkością i stanowią swoje powtórzenia, czyli są testami równoległymi. Stopień, w jakim pozycje skali są zgodne ze sobą traktowany jest jako wartość rzetelności, nazywanej rzetelnością wewnętrzną

(*internal consistency reliability*). Najbardziej znanymi metodami obliczeniowymi tego rodzaju są: współczynnik α Cronbacha oraz wzory Kudara i Richardsona. Koncepcyjnie podobną metodą jest porównanie części (połówek) tego samego testu (*half-split reliability*). Zaletą metod obliczania rzetelności na podstawie wewnętrznej spójności jest prostota wykonania badania, nie wymagająca powtórzeń pomiarów. Należy jednak wziąć pod uwagę fakt, że obliczony w ten sposób współczynnik rzetelności związany jest ze zgodnością pozycji skali, czyli homogennością, natomiast nie daje informacji na temat powtarzalności pomiarów [6].

RZETELNOŚĆ SKAL OCENY DEPRESJI W INNYCH BADANIACH

Z ponad 30 aktualnie stosowanych w świecie skal, w Polsce wykorzystywana jest najczęściej skala oceny depresji Hamiltona (HAMD). Skala ta, mimo swej popularności, była krytykowana z powodu heterogenności, wykazywano także jej niezbyt dużą trafność [1, 5]. W roku 1970 Montgomery i Åsberg, a w 1980 r. Bech i Rafaelsen zaproponowali inne, wg autorów bardziej trafne i rzetelne narzędzia oceny nasilenia depresji. Zarówno HAMD, jak i dwie pozostałe skale, poddawane były wielokrotnie badaniom rzetelności i trafności. Hamilton, wprowadzając w 1960 r. swą skalę depresji, określał współczynnik jej rzetelności na 0,84 (przy 10-osobowej grupie badanych), a 0,90 przy 70 pacjentach badanych przez dwóch psychiatrów [8]. W badaniu Becha i wsp. z 1979 r. współczynnik rzetelności typu *inter-observer* skali HAMD oraz jej subskali wahał się od 0,69 do 0,97. Średnia interkorelacja pomiędzy 4 badaczami mierzona współczynnikiem zgodności Kendala wynosiła dla całej skali 0,79 ($n=9$, $p<0,01$) [3]. Wysoką rzetelność HAMD wykazało badanie Faravellego i wsp. z 1986 r. – współczynnik rzetelności typu *inter-rater* (Pearsona) wynosił 0,95 [7]. Autorzy skali MADRS potwierdzili wysoką rzetelność włas-

Tablica 1. Rezultaty estymacji rzetelności (*test-retest*) uzyskane przez Maiera [11]

| Skala | Współczynnik korelacji przed leczeniem | Współczynnik korelacji po leczeniu |
|-------------------|--|------------------------------------|
| HAMD – 21 pozycji | 0,70 | 0,69 |
| – 17 pozycji | 0,72 | 0,70 |
| – 5 pozycji | 0,69 | 0,66 |
| MADRS | 0,66 | 0,82 |
| BRMES | 0,79 | 0,88 |

nej skali w badaniu *inter-rater* z 1970 r., uzyskując współczynniki korelacji między 0,89 a 0,93 w zależności od liczby i doświadczenia badających [13]. Rzetelność typu *inter-rater* skali BRMES została określona w badaniu Becha i wsp. na 0,82 [4], a przez Smolkę na 0,8–0,91 [2].

Wysoką rzetelność skal HAMD, MADRS i BRMES potwierdził Korner i wsp., uzyskując kolejno współczynniki rzetelności *intra-class*: 0,83, 0,86, 0,81 dla tych skal [10].

Maier i Philipp przeprowadzili natomiast badanie rzetelności wewnętrznej spójności wszystkich wspomnianych tu skal, uzyskując wyniki świadczące, że choć wszystkie narzędzia mają zadowalającą rzetelność, to ich jednorodność jest niejednakowa: najmniejsza w przypadku HAMD (wsp. Cronbacha 0,77) średnia dla BRMES (0,81), największa w przypadku MADRS (0,86) [12]. Poniżej (tab. 1) przedstawiamy natomiast wyniki badania Maiera, określające rzetelność typu *test-retest* wszystkich trzech wymienionych skal (1990). Badanie objęło 40 pacjentów ocenianych przez 2 psychiatrów na dzień przed rozpoczęciem leczenia i w 21 dniu terapii [11].

CEL BADANIA

Przeprowadzone badanie miało na celu określenie dla polskiej populacji pacjentów rzetelności dotąd w Polsce nie wykorzystywanej skali melancholii Becha i Rafaelsena (BRMES) oraz porównanie jej z rzetelnością skal już uznanych – skali Hamiltona (HAMD) oraz Montgomery-Åsberg (MADRS).

BADANE OSOBY

Badaniem objęto grupę 34 pacjentów (tab. 2) hospitalizowanych w Klinice Psychiatrii AM oraz Specjalistycznym Zespole Psychiatrycznej Opieki Zdrowotnej we Wrocławiu w związku z zaburzeniami depresyjnymi oraz pacjentów leczonych z powodu depresji w poradni zdrowia psychicznego (PZP). Dobór pacjentów przebiegał dwuetapowo i polegał na: (1) wstępnej diagnozie zespołu depresyjnego dokonanej przez lekarza przyjmującego pacjenta w izbie przyjęć lub leczącego w PZP i (2) potwierdzeniu wstępnej diagnozy zespołu depresyjnego przez lekarza prowadzącego badanie.

Tablica 2. Podstawowe cechy badanej grupy

| Zmienne | Ogółem |
|------------------|--------|
| Liczba pacjentów | 34 |
| Średni wiek | 54,9 |
| Płeć: kobiety | 29 |
| mężczyźni | 5 |

Doboru dokonywano na podstawie kryteriów zaburzeń afektywnych wg klasyfikacji ICD-10. Włączano pacjentów z rozpoznaniami: zaburzenie afektywne dwubiegunowe – epizod depresyjny (F31), epizod depresyjny (F32), zaburzenia depresyjne nawracające (F33), dystymia (F34), zaburzenia depresyjne i lękowe mieszane (F41.2). Z badania wykluczono pacjentów, u których występowały: organiczne zaburzenia psychiczne, uzależnienie od substancji psychoaktywnych oraz schizofrenia i zaburzenia typu schizofrenii.

METODA

Każdy pacjent poddany został jednorazowemu badaniu psychiatrycznemu w trzeciej lub czwartej dobie od dnia przyjęcia do szpitala lub w dniu zgłoszenia się do PZP. Badanie przeprowadzane było zawsze w godzinach przedpołudniowych.

Chorzy byli poddani typowemu badaniu psychiatrycznemu (niestandardyzowany wywiad), trwającemu 30–60 minut. Obejmowało ono, oprócz oceny stanu psychicznego, także wywiad dotyczący obciążenia rodzinnego badanego chorobą psychiczną, wiek pacjenta w momencie pierwszego zachorowania oraz obecność czynnika reaktywnego. Bezpośrednio po badaniu dokonywano oceny ciężkości depresji pacjenta za pomocą skali BRMES, HAMD oraz MADRS. U każdego badanego dokonywana była również ocena globalnej ciężkości stanu psychicznego przy użyciu siedmiostopniowej skali *Clinical Global Impression* (CGI) o następujących przedziałach:

-
-
- 1 – niewielkie odchylenia od normy
 - 2 – pograniczne zaburzeń psychicznych
 - 3 – niewielkie nasilenie zaburzeń psychicznych
 - 4 – umiarkowane nasilenie zaburzeń psychicznych
 - 5 – wyraźne zaburzenia psychiczne
 - 6 – ciężkie zaburzenia psychiczne
 - 7 – należał do chorych w najcięższym stanie
-
-

Badanie przeprowadziło 2 lekarzy: psychiatra i lekarz ogólny specjalizujący się w psychiatrii, którzy oceniali stan psychiczny każdego pacjenta oraz głębokość jego zaburzeń depresyjnych za pomocą wymienionych 3 skal oraz skali CGI, niezależnie od siebie podczas tego samego badania.

Rzetelność typu *inter-rater* oceniono metodą *intra-class reliability* wg modelu ICR(2,1) oraz za pomocą współczynnika korelacji Spearmana. Dla oceny rzetelności spójności wewnętrznej skal wykorzystano współczynnik α Cronbacha. Zbadano także, czy różnice pomiędzy wynikami uzyskanymi przez osoby oceniające są istotne statystycznie. Obliczenia dla całych skal, jak i wszystkich pozycji skal wykonano za pomocą testu U Mana-Whitneya.

WYNIKI

Tab. 3 przedstawia podstawowe wyniki oceny nasilenia depresji uzyskane przez dwie osoby badające. Różnice średnich wyników skalowania uzyskanych przez obu klinycystów są małe i nieistotne statystycznie.

Wartości współczynników rzetelności przedstawiono w tabelicy 4. Dla wartości ICR(2,1) przedstawiono także granice przedziałów ufności dla prawdopodobieństwa znalezienia się współczynnika poza wyznaczonym przedziałem wynoszącym 0,05. Współczynnik Cronbacha wyliczono dla każdej osoby badającej oddzielnie.

Tablica 3. Wyniki oceny nasilenia depresji uzyskane przez dwóch badaczy

| Skala | Średnia | Minimum | Maksimum | Odchylenie standardowe |
|-------|---------|---------|----------|------------------------|
| BRMES | 18,59 | 7 | 39 | 7,37 |
| | 17,47 | 6 | 40 | 8,52 |
| HAMD | 20,29 | 7 | 37 | 7,37 |
| | 20,38 | 4 | 43 | 9,44 |
| MADRS | 26,79 | 7 | 58 | 11,45 |
| | 26,62 | 4 | 60 | 12,53 |
| CGI | 3,79 | 1 | 7 | 1,27 |
| | 3,79 | 1 | 7 | 1,43 |

Tablica 4. Wartości rzetelności skal (współczynnik α Cronbacha wyliczono dla każdej z dwóch osób badających oddzielnie)

| Skala | Wartość współczynnika rzetelności ICR(2,1) | Współczynnik korelacji Spearmana | Współczynnik α Cronbacha |
|-------|--|----------------------------------|---------------------------------|
| BRMES | 0,84 < 0,91 < 0,95 | 0,90 | 0,83 0,88 |
| HAMD | 0,82 < 0,90 < 0,94 | 0,87 | 0,72 0,83 |
| MADRS | 0,92 < 0,96 < 0,97 | 0,95 | 0,90 0,94 |
| CGI | 0,83 < 0,90 < 0,94 | 0,88 | – |

Uzyskane wartości rzetelności typu *inter-rater* są wysokie dla wszystkich oceniających skal i świadczą o dużej zgodności uzyskanych wyników. Największą rzetelność osiągnęła skala MADRS, zarówno mierzoną współczynnikiem ICR, jak i współczynnikiem korelacji Spearmana. Współczynniki spójności wewnętrznej α Cronbacha należy ocenić także jako wysokie dla skal BRMES i MADRS, natomiast jako średnie dla skali HAMD.

Dla dokładniejszej analizy wewnętrznej struktury zastosowanych skal zbadano rów-

nież na ile zmieni się spójność danej skali po usunięciu poszczególnych jej pozycji. Obliczono przy tym korelacje między każdą pozycją a ogólnym wynikiem (tab. 5, 6, 7).

Osoby badające uzyskały w przypadku wszystkich skal zbliżone średnie wartości nasilenia depresji. Ocena różnic wyników skalowania pomiędzy badającymi nie wykazała istotnych statystycznie rozbieżności. Aby dokładniej ocenić zgodność skalowania analizowano także różnice dla każdej pozycji skal (z wyjątkiem CGI, która składa się z jednej tylko pozycji). W przypadku skali

Tablica 5. Wartości współczynników α Cronbacha skali BRMES po usunięciu poszczególnych pozycji skali oraz współczynników korelacji odpowiednich pozycji skali z wynikiem ogólnym

| Streszczenie dla skali: średnia = 19,3800, odchylenie standardowe = 7,26522, n = 100 | | |
|--|------------------------|-------------------------------|
| Współczynnik α Cronbacha: 0,812696, standaryzowany współczynnik α : 0,813061 | | |
| Przeciętna korelacja między pozycjami skali: 0,297255 | | |
| Pozycja skali Becha-Rafaelsena | Korelacja pozycja–suma | α po usunięciu pozycji |
| B1 aktywność motoryczna | 0,68 | 0,78 |
| B2 aktywność werbalna | 0,62 | 0,78 |
| B3 spowolnienie psychiczne | 0,69 | 0,78 |
| B4 lęk psychiczny | 0,21 | 0,82 |
| B5 impulsy suicydalne | 0,24 | 0,83 |
| B6 obniżony nastrój | 0,63 | 0,78 |
| B7 niska samoocena i poczucie winy | 0,46 | 0,80 |
| B8 zahamowanie uczuciowe | 0,64 | 0,78 |
| B9 zaburzenia snu | 0,28 | 0,82 |
| B10 zmęczenie i bóle | 0,29 | 0,81 |
| B11 praca i zainteresowania | 0,55 | 0,79 |

Tablica 6. Wartości współczynników α Cronbacha skali HAMD po usunięciu poszczególnych pozycji skali oraz współczynników korelacji odpowiednich pozycji skali z wynikiem ogólnym

| Streszczenie dla skali: średnia = 22,1414, odchylenie standardowe = 7,29553, n = 99 | | |
|--|------------------------|-------------------------------|
| Współczynnik α Cronbacha: 0,732445, standaryzowany współczynnik α : 0,757975 | | |
| Przeciętna korelacja między pozycjami skali: 0,161794 | | |
| Pozycja skali Hamiltona | Korelacja pozycja–suma | α po usunięciu pozycji |
| H1 nastrój depresyjny | 0,63 | 0,68 |
| H2 poczucie winy | 0,35 | 0,72 |
| H3 zniechęcenie do życia, tendencje samobójcze | 0,11 | 0,76 |
| H4 zaburzenia zasypiania | 0,25 | 0,73 |
| H5 sen płytki, przerywany | 0,34 | 0,72 |
| H6 wczesne budzenie się | 0,24 | 0,73 |
| H7 aktywność złożona, praca | 0,40 | 0,71 |
| H8 spowolnienie, zahamowanie | 0,48 | 0,70 |
| H9 niepokój, podniecenie ruchowe | 0,09 | 0,74 |
| H10 lęk – objawy psychiczne | 0,32 | 0,72 |
| H11 lęk – objawy somatyczne | 0,24 | 0,73 |
| H12 brak apetytu, zaparcia | 0,52 | 0,71 |
| H13 objawy somatyczne ogólne | 0,37 | 0,72 |
| H14 utrata libido, zaburzenia miesiączkowania | 0,30 | 0,72 |
| H15 hipochondria | 0,34 | 0,72 |
| H16 ubytek masy ciała | 0,32 | 0,72 |
| H17 krytycyzm | 0,45 | 0,71 |

Tablica 7. Wartości współczynników α Cronbacha skali MADRS po usunięciu poszczególnych pozycji skali i współczynników korelacji odpowiednich pozycji skali z wynikiem ogólnym

| Streszczenie dla skali: średnia = 27,9700, odchylenie standardowe = 9,91220, n = 100 | | |
|--|------------------------|-------------------------------|
| Współczynnik α Cronbacha: 0,842191, standaryzowany współczynnik α : 0,853255 | | |
| Przeciętna korelacja między pozycjami skali: 0,382145 | | |
| Pozycja skali Montgomery-Åsberg | Korelacja pozycja–suma | α po usunięciu pozycji |
| M1 smutek obserwowany | 0,79 | 0,81 |
| M2 skargi na smutek | 0,75 | 0,81 |
| M3 poczucie napięcia wewnętrznego | 0,40 | 0,84 |
| M4 skrócenie snu | 0,23 | 0,86 |
| M5 zmniejszenie apetytu | 0,58 | 0,82 |
| M6 trudności w koncentracji uwagi | 0,66 | 0,82 |
| M7 znużenie, męczliwość | 0,57 | 0,82 |
| M8 niemożność przeżywania uczuć | 0,58 | 0,82 |
| M9 pesymistyczne myśli | 0,61 | 0,82 |
| M10 myśli samobójcze | 0,38 | 0,84 |

Hamiltona największe różnice pomiędzy osobami badającymi wystąpiły w trzech pozycjach (spłylenie snu, objawy somatyczne

lęku, spadek masy ciała), a w przypadku skali melancholii – w jednej pozycji (aktywność motoryczna). Prawdopodobieństwo,

że różnica ma charakter losowy nie przekraczało 0,05. W skali Montgomery-Åsberg różnice między badaczami były nieistotne statystycznie.

OMÓWIENIE

W przeprowadzonym badaniu uzyskano wysoką rzetelność typu *inter-rater* wszystkich zastosowanych skal. Otrzymane wartości współczynników rzetelności od 0,87 do 0,96 (metodą ICR oraz korelacji Spearmana) potwierdzają dużą zgodność wyników badania.

Najwyższą rzetelność wykazała skala MADRS, a stosunkowo niską – rzetelność skali Hamiltona. Jest to zastanawiające, jako że jest to skala najdłuższa (17 pozycji), a teorie psychometryczne wskazują, iż większa liczba pozycji testu wpływa na wzrost jego rzetelności [11].

Stosunkowo dużą rzetelność uzyskała także skala CGI, która jest skalą jednopozycyjną, więc wg wspomnianych już założeń nie należało oczekiwać jej wysokiej rzetelności. Oznacza to, że dla niektórych zastosowań jest to wystarczające narzędzie pomiaru głębokości depresji.

Potwierdzenie wysokiej rzetelności wszystkich badanych skal, które stosowane były w oparciu o niestrukturalizowane badanie psychiatryczne, dowodzi również, że wykorzystanie strukturalizowanych wywiadów, proponowanych przez autorów, nie jest konieczne. Korzystanie z tych narzędzi poprzedzone powinno być jedynie treningiem przygotowującym do badania skalami.

Badanie jednorodności skal metodą Cronbacha wykazało, że największą wewnętrzną spójnością cechuje się MADRS. Z wyjątkiem pozycji oceniającej zaburzenia snu, dla wszystkich innych uzyskano wysoką korelację z wynikiem ogólnym, a usunięcie tych pozycji zmniejszałoby spójność skali (tab. 6). Struktura skali BRMES jest mniej jednorodna. Cztery jej pozycje (lęk, zaburzenia snu, impulsy samobójcze, zmęczenie i bóle) nie są wystarczająco skorelowane z wynikiem ogólnym, natomiast po ich usunięciu spójność

skali wzrosłaby (tab. 4). Najmniejszą jednorodność ma w przeprowadzonym badaniu skala HAMD. Świadczy o tym wyraźnie niższy niż w poprzednich przypadkach współczynnik oraz mała zgodność poszczególnych pozycji skali z wynikiem ogólnym (tab. 5). Być może rezygnacja z niektórych pozycji skali, co czyniono już w innych badaniach tego typu, podniosłaby wartość psychometryczną tego narzędzia.

Badanie oceniało także różnice między pozycjami skal u obu badaczy. W większości okazały się one nieistotne statystycznie. Najwięcej rozbieżności wystąpiło przy wykorzystaniu skali Hamiltona – dotyczyły one pomiaru spłycenia snu, objawów somatycznych lęku oraz spadku masy ciała. W skali BRMES badacze różnili się w ocenie aktywności motorycznej. Rozbieżności mogą być spowodowane nieprecyzyjnym zdefiniowaniem objawów. Przyczyną wystąpienia różnic między badaczami może być także ich niejednakowe doświadczenie kliniczne.

Uzyskane wyniki dotyczące rzetelności wykorzystanych skal można porównać z cytowanymi w pierwszej części pracy rezultatami innych badań w tej dziedzinie. Bardzo podobne są także wyniki badania spójności wewnętrznej tychże skal, które niemal pokrywają się z otrzymanymi w badaniu Maiera i Philippa [12].

WNIOSKI

1. Skala melancholii BRMES okazała się rzetelnym narzędziem pomiaru głębokości depresji.
2. Wyniki potwierdzają także zadowalającą rzetelność pozostałych narzędzi: skali Hamiltona i Montgomery-Åsberg oraz prostszej, jednopozycyjnej skali CGI.
3. W ocenie wewnętrznej spójności metodą Cronbacha najbardziej jednorodna okazała się skala Montgomery-Åsberg. Spójność skali melancholii była mniejsza, zaś skala Hamiltona zaprezentowała się w przeprowadzonym badaniu jako narzędzie niehomogenne.

PIŚMIENNICTWO

1. Bech P.: Rating scales for affective disorders, their validity and consistency. *Acta Psychiatr. Scand.* 1981, suppl. 295, 1–101.
2. Bech P.: The Bech, Hamilton and Zung scales for mood disorders: screening and listening. Springer-Verlag, Berlin, Heidelberg 1996.
3. Bech P., Bolwig T.G., Kramp P., Rafaelsen J.: The Bech-Rafaelsen Mania Scale and the Hamilton Depression Scale. Evaluation of homogeneity and inter-observer reliability. *Acta Psychiatr. Scand.* 1979, 59, 420–430.
4. Bech P., Gjerris A., Andersen J., Bojholm S., Kramp P., Bolwig T.G., Kastrup M., Clemmesen L., Rafaelsen O.J.: The melancholia scale and the Newcastle scales. Item combinations and inter-observer reliability. *Br. J. Psychiatry* 1983, 143, 58–63.
5. Bech P., Rafaelsen O.J.: The use of rating scales exemplified by a comparison of the Hamilton and The Bech-Rafaelsen Melancholia Scale. *Acta Psychiatr. Scand.* 1980, 62, suppl. 285, 128–132.
6. Brzeziński J.: *Metodologia badań psychologicznych.* Wyd. Nauk. PWN, Warszawa 1996.
7. Faravelli C., Albanesi G., Poli E.: Assessment of depression: a comparison of rating scales. *J. Affect. Disord.* 1986, 11, 245–253.
8. Hamilton M.: A rating scale for depression. *J. Neurol. Neurosurg. Psychiatry* 1960, 23, 56–62.
9. Kessler R.C.: Epidemiology of Psychiatric Comorbidity. *Textbook in Psychiatric Epidemiology.* Wiley-Liss Inc. 1995, 179–197.
10. Korner A., Nielsen B.M., Eschen F., Moller-Madsen S., Stender A., Christensen E.M., Aggernaes H., Kastrup M., Larsen J.K.: Quantifying depressive symptomatology: inter-rater reliability and inter-item correlations. *J. Affect. Disord.* 1990, 20, 143–149.
11. Maier W.: The Hamilton depression scale and its alternatives: a comparison of their reliability and validity. *Psychopharmacology* 1990, 9, 65–71.
12. Maier W., Philipp M.: Comparative analysis of observer depression scales. *Acta Psychiatr. Scand.* 1985, 72, 230–245.
13. Montgomery S., Åsberg M.: A new depression scale designed to be sensitive to change. *Br. J. Psychiatry* 1979, 134, 382–389.
14. Wittchen H.U., Essau C.: Comorbidity and Mixed Anxiety-Depressive Disorders: Is There Epidemiologic Evidence? *J. Clin. Psychiatry* 1993, 54, 1, suppl., 9–15.

*Adres: Dr Joanna Mazurek, Katedra i Klinika Psychiatrii AM,
ul. Kraszewskiego 25, 50-229 Wrocław*