

## Badanie zgodności rozpoznań psychiatrycznych z kryteriami diagnostycznymi klasyfikacji ICD-10 za pomocą kwestionariusza SCAN

### Study of compatibility of psychiatric diagnoses with ICD-10 diagnostic criteria using the SCAN questionnaire

Tomasz Adamowski, Andrzej Kiejna, Tomasz Hadryś

Katedra i Klinika Psychiatrii AM we Wrocławiu  
Kierownik: prof. dr hab. n. med. A. Kiejna

#### Summary

**Aim.** Authors aimed at testing whether psychiatrists in their diagnostic process obeyed strict ICD-10 diagnostic criteria. Diagnoses made by psychiatrists at discharge were compared with those of SCAN ver.2.1 on admission.

**Methods.** Diagnoses obtained by SCAN I-Shell program were compared with clinical diagnoses obtained by psychiatrists in the psychiatric wards according to ICD-10 criteria on 3 levels: diagnostic group (Fc), diagnostic class (Fcc), and diagnostic category (Fcc.c). Validity assessment was obtained with Cohen's Kappa coefficient, sensitivity, specificity and Yule's Y coefficient.

**Results.** On the diagnostic group level, Cohen's kappa was 0.14-0.65, Yule's Y 0.57-0.71. Sensitivity 0.69-0.95 and specificity 0.41-0.94. In psychotic disorders group F2 kappa was 0.65, Yule's Y 0.71, sensitivity 0.69, specificity 0.94. In affective disorders group F3 kappa was 0.31, Yule's Y 0.57, sensitivity 0.95, specificity 0.41. In neurotic disorders group F4 kappa was low 0.14, Yule's Y 0.62, sensitivity 0.95, specificity 0.50.

**Conclusions.** The study showed a higher level of agreement between SCAN and clinical diagnoses in the group of psychotic disorders with exception of schizoaffective disorders, and lower agreement rates in the group of affective and neurotic disorders where the number of SCAN diagnoses outweighed that of the clinical ones. It could be the result of systematic faults in the coding of diagnoses.

*Słowa klucze:* SCAN, zgodność, diagnoza

*Key words:* SCAN, agreement, diagnosis

#### Wstęp

Diagnostyka w psychiatrii opiera się na wieloosiowym wymiarze zaburzeń psychicznych zawartych w kryteriach klasyfikacyjnych ICD lub DSM. Poza obecnością czynników kryterialnych uwzględnia się też wymiar stygmatyzacyjny zaburzenia

psychicznego, jego znaczenie rentowe i orzecznicze. Psychiatra jako pracownik instytucji poddany jest naciskom administracyjnym – w celu najbardziej efektywnego podejścia do leczenia pacjenta musi jednocześnie godzić interes szpitala z interesem instytucji refundującej koszty. W obliczu tych coraz bardziej skomplikowanych zależności postawienie rzetelnej w sensie naukowym diagnozy ma zasadnicze znaczenie dla wdrożenia odpowiedniego programu leczniczego i rokowania.

W związku z tym powstaje pytanie, jak dalece pomocne w tym procesie może być zastosowanie narzędzia diagnostycznego SCAN, skonstruowanego na podstawie kryteriów współczesnych klasyfikacji psychiatrycznych. Wyniki dotychczas przeprowadzonych badań nie potwierdzają tego jednoznacznie. Diagnoza za pomocą kwestionariusza w porównaniu z diagnozą kliniczną ma różny stopień zgodności, w zależności od konstrukcji narzędzia (strukturalizowane lub nie). Ponadto posługiwanie się narzędziami diagnostycznymi wymaga nabycia umiejętności ich obsługi i jest czasochłonne. Może jednak warto zadać sobie trud wspomaganie diagnostyki za pomocą narzędzi diagnostycznych. W niektórych krajach (np. w USA) już od wielu lat stosowane są systemy, w których finansowanie usług szpitalnych odbywa się na podstawie grup diagnostycznych (ang. diagnosis related group, DRG), do których pacjent zostaje przydzielony już w bardzo wczesnym etapie hospitalizacji.

Jedną z zasadniczych kwestii pozostaje obiektywizacja metod diagnostycznych w psychiatrii. Posiadając rzetelne, strukturalizowane w całości lub w części narzędzia, wraz ze zoperacjonalizowanymi nowoczesnymi kryteriami diagnostycznymi, można porównać diagnostykę z ich wykorzystaniem z tradycyjnymi rozpoznaniami klinicznymi przez psychiatrę. Porównanie jest istotne w celu sprawdzenia, w jakim stopniu lekarz posługuje się kryteriami diagnostycznymi współczesnych klasyfikacji. Dlaczego występują niezgodności, czy mają przyczyny natury klinicznej, czy innej? Stosowanie zoperacjonalizowanych kryteriów pozwala uzyskać wyższe wskaźniki zgodności pomiędzy badaczami. Wykorzystanie akceptowanych międzynarodowych kryteriów badawczych zwiększa rzetelność diagnozy zaburzeń psychicznych w centrach naukowych na całym świecie [1].

Sukces dużych badań zależy od dobrze zaprojektowanych narzędzi oraz poziomu wykszolenia badaczy. Niezgodności we wskaźnikach rozpowszechnienia w badaniach epidemiologicznych oraz słaba zgodność pomiędzy badaczem a klinicystą stanowią problem w badaniach naukowych, w planowaniu programów zdrowia publicznego i w przypadku ubezpieczeń. W jednym z badań sprawdzano zgodność pomiędzy przeszkolonymi epidemiologami a klinicystami w wykorzystaniu SCAN do wykrywania zaburzeń psychotycznych i neurotycznych. Osiągnięto zadowalające wyniki. Zgodność dla wszystkich zaburzeń wynosiła 0,74 (95% CI: 0,57–0,91), a dla specyficznych zaburzeń psychotycznych 0,63 (0,40–0,86), neurotycznych 0,63 (0,43–0,83). Czulość od 0,6 do 0,9, specyficzność 0,8–0,9 [2].

Przeprowadzono badanie, w którym porównano pacjentów z pierwszym epizodem psychozy rozpoznanej zgodnie z klasyfikacją ICD-10 z pacjentami z diagnozą stawianą za pomocą SCAN 2. Całkowita zgodność kappa wynosiła 0,55 (95% CI: 0,44–0,66). W kategoriach diagnostycznych kappa wynosiła 0,47 dla grupy schizofrenii: F20.x, 25.x, 23.2 (czulość 0,49, specyficzność 0,98); 0,46 dla ostrych i przemijających zaburzeń psychotycznych F22.x (0,86; 0,77); 0,63 dla zaburzeń dwubiegunowych: F30.x, F31.x (0,55; 0,99); 0,47 dla nasilonych epizodów depresyjnych z objawami psy-

chotycznymi: F32.3 (0,33; 1). Większe różnice w zgodności związane były z niższym natężeniem objawów negatywnych w skali PANSS, niskim poziomem wykształcenia, wysokim poziomem objawów urojeniowych PANSS (4–7). W grupie schizofrenii były inne rezultaty: nasilenie psychopatologii nie miało wpływu na różnice zgodności, wiek miał pozytywny wpływ, tzn. starszy wiek nasilał niezgodność diagnostyczną, tak samo jak aktywność społeczna poza miejscem zamieszkania [3].

Autorzy w przeprowadzonym badaniu chcieli sprawdzić, w jakim stopniu psychiatra w procesie diagnostyki stosuje precyzyjne kryteria ICD-10. W tym celu porównano rozpoznania postawione przez psychiatrę przy wypisaniu ze szpitala z diagnozami postawionymi przy przyjęciu z wykorzystaniem narzędzia SCAN – wersja 2.1 u tych samych pacjentów [4]. Wybrano do porównania rozpoznania kliniczne wypisowe, ponieważ stawiane przez psychiatrów przy przyjęciu do szpitala mają charakter diagnozy wstępnej, która może ulec ostatecznie zmianie do momentu wypisania. Następuje to w wyniku szczegółowej oceny objawów psychopatologicznych podczas pobytu na oddziale oraz uwzględnienia danych z wywiadu, ewentualnie innych źródeł (rodzina). W przypadku diagnozowania z wykorzystaniem narzędzia SCAN te dodatkowe okoliczności nie wpływały na rozpoznanie. Hipoteza badawcza była następująca: zastosowanie w diagnostyce w części zoperacjonalizowanych kryteriów diagnostycznych narzędzia SCAN powinno uszczegółowić diagnozę postawioną przez lekarza psychiatrę.

### Material

Materiał do badania pochodził z wrocławskiej bazy danych projektu EDEN<sup>1</sup>. Badanie EDEN miało charakter prospektywny i zostało przeprowadzone na podstawie ustalonego wcześniej protokołu, w którym określono kryteria włączania i wyłączenia pacjentów oraz metody i harmonogram [5]. Badanie zgodności rozpoznań psychiatrycznych z kryteriami diagnostycznymi klasyfikacji ICD-10 za pomocą kwestionariusza SCAN nie zostało zaplanowane w ramach tego projektu. Do badania wrocławskiego włączono 257 pacjentów, przyjętych do szpitala psychiatrycznego we Wrocławiu, spełniających kryteria. Analizą objęto 220 pacjentów (n = 111 na oddziale dziennym i n = 109 na oddziale stacjonarnym). Wszyscy zostali zbadani za pomocą kwestionariusza SCAN przy przyjęciu (t1). Przy wypisaniu ze szpitala (t2) każdy pacjent otrzymał diagnozę kliniczną postawioną przez psychiatrę.

### Metoda

#### Narzędzia badawcze

SCAN [4]

Badania za pomocą SCAN-2.1 były przeprowadzane w ciągu trzech dni od przyjęcia pacjenta do szpitala (t1) przez 4 lekarzy, z których dwóch miało specjalizację

<sup>1</sup>(European Day Hospital Evaluation – Europejska Ocena Psychiatrycznych Szpitali Dziennych; [www.edenstudy.com](http://www.edenstudy.com)) Międzynarodowy projekt badawczy 5 PR UE, QLG4-CT-2000-01700 (1999–2003), przeprowadzony w 5 ośrodkach europejskich: Drezno (Niemcy), Londyn (Wielka Brytania), Wrocław (Polska), Michalovce (Słowacja) i Praga (Czechy).

z psychiatrii oraz dwóch, którzy odbywali szkolenie specjalizacyjne – zaliczyli staż z psychiatrii klinicznej i wykazali się wiedzą z psychopatologii. Wszyscy badacze przeszli tygodniowe przeszkolenie w zakresie posługiwania się kwestionariuszem SCAN w Klinice Psychiatrii i Psychoterapii na Wydziale Lekarskim Uniwersytetu Technicznego w Dreźnie w 2000 roku. Szkolenie przeprowadzone było przez eksperta WHO z Centrum Szkoleniowego SCAN profesora Terry Brugę z Kliniki Psychiatrycznej Uniwersytetu w Leicester (UK). Poziom znajomości psychopatologii oceniano w całym okresie badania, kontrolując zgodność uzyskiwanych przez badaczy średnich wartości skali całkowitej i podskal BPRS Ventury, opartej na ICC (ang. intra-class correlation coefficient). Dla skali całkowitej uzyskano wartość ICC = 0,79 (substantial [La]), natomiast dla podskal – od 0,58 do 0,94 [6]. W związku z wymogiem znajomości psychopatologii, kontrolowanej w całym okresie badania, badacze, pomimo różnic kwalifikacyjnych formalnych (zaawansowanie specjalizacji), osiągnęli taki sam poziom kwalifikacji operacyjnych (kontrola BPRS, szkolenie SCAN, zaliczenie z psychopatologii) w zakresie stosowania wiedzy z psychopatologii w posługiwaniu się narzędziem SCAN. Okresem podlegającym ocenie był „the present state (PS)” – „czas obecny”. Brano pod uwagę stan kliniczny pacjenta w ciągu 1 miesiąca przed hospitalizacją. Badanie miało na celu opisanie wszystkich objawów, które się pojawiły w tym czasie. Dane były zbierane komputerowo zgodnie z algorytmem IShell ver. 1.0.1.2, który umożliwiał otrzymanie komputerowej diagnozy według międzynarodowej klasyfikacji ICD-10. W przypadku większej liczby diagnoz uzyskanych przez IShell, zostały ustalone główne diagnozy za pomocą SCAN na podstawie informacji z programu komputerowego. Niezależnie od diagnozy ustalonej za pomocą SCAN każdy pacjent otrzymał rutynowo rozpoznanie kliniczne postawione przez psychiatrię po zakończeniu leczenia na oddziale dziennym lub stacjonarnym.

#### Analizy statystyczne

Diagnozy postawione przez program (algorytm) IShell narzędzia SCAN zostały porównane z rozpoznaniem klinicznymi postawionymi przez psychiatrię na oddziale, zgodnie z kryteriami klasyfikacji ICD-10, na 3 sposoby: w grupach diagnostycznych (Fc), w klasach diagnostycznych (Fcc) oraz na poziomie kategorii diagnostycznych o znaczeniu klinicznym (Fcc.c). Wyróżniono następujące kategorie: (A) zaburzenia spowodowane nadużywaniem substancji (F10.1-F19.9); (B) zaburzenia psychotyczne (F20.0-F29.0, z wyłączeniem F25.0); (C) zaburzenia schizoafektywne (F25.0-F25.9); (D) zaburzenia maniakalne (F30.0-F31.2); (E) zaburzenia depresyjne (F31.3-F31.5; F32.0-F33.9; F34.1); (F) zaburzenia lękowe (F40.0-F42.9, z wyłączeniem F41.2); (G) reakcja na stres i zaburzenia adaptacyjne (F43.0-F43.9); (H) zaburzenia występujące pod postacią somatyczną (F45.0-F45.9); (I) inne zaburzenia nerwicowe (F48.0-F48.9); (J) zaburzenia odżywiania się (F50.0-F50.2); (K) zaburzenia afektywne nieokreślone (F39).

Ocenę trafności narzędzia ustalono z wykorzystaniem współczynnika kappa Cohena, czułości (sensitivity), specyficzności (specificity) oraz współczynnika Yule'a. Uzyskane wyniki dotyczyły dwóch typów analiz. Najpierw ustalono punktację dla każdej z diagnoz z dokładnością do dwóch cyfr (tj. w grupach Fcc), sumując wyniki

rozpoznań szczegółowych. Analizowana zmienna przyjmowała wartość 1, gdy wśród rozpoznań za pomocą SCAN występowała dana diagnoza oraz 0 – w przeciwnym przypadku. Podobnie przeprowadzono analizę każdej z diagnoz z dokładnością do jednej cyfry (tj. w grupach Fc).

W celu porównania diagnozy za pomocą SCAN i diagnozy klinicznej obliczono następujące wskaźniki:

**Całkowita zgodność procentowa (overall percentage agreement – PA)** – stosunek liczby osób, którym program SCAN postawił rozpoznanie kliniczne do osób, którym ani program SCAN, ani psychiatra nie postawili rozpoznania. Wielkość ta zależy od rozpowszechnienia (base rate) i nie jest poprawiona o zgodę przypadkową.

**Czułość (sensitivity)**. Ilość osób z diagnozą kliniczną, u których stwierdzono przynajmniej jedno rozpoznanie za pomocą SCAN z tej samej grupy.

**Specyficzność (specificity)**. Ilość osób, u których nie uzyskano diagnozy za pomocą SCAN w tych wszystkich wypadkach, gdy nie dał jej również psychiatra.

**Kappa ( $\kappa$ )**. Współczynnik  $\kappa$  jest powszechnie stosowanym współczynnikiem zgodności, który uwzględnia korektę zgody obserwowanej o zgodę przez przypadek [7]. Przybiera wartości od -1 (w szczególnych przypadkach zależnych od rozpowszechnienia [7], przez 0 dla braku zgody większej niż zgoda przez przypadek, do +1 dla całkowitej zgodności. Landis i Koch [8] zaproponowali następującą interpretację wielkości kappi: wielkość  $\kappa$  w przedziale pomiędzy  $0,01 < \kappa < 0,20$  jako niewielką (slight), pomiędzy  $0,21 < \kappa < 0,40$  jako zadowalającą (fair), pomiędzy  $0,41 < \kappa < 0,60$  jako umiarkowaną (moderate), pomiędzy  $0,61 < \kappa < 0,80$  jako znaczną (substantial) i  $\kappa > 0,81$  jako prawie całkowitą (almost perfect). Główną wadą współczynnika  $\kappa$  jest wrażliwość na rozpowszechnienie (base rate), tzn. znaczne wahania przy niższych wskaźnikach rozpowszechnienia [9]. Dlatego nie zaleca się stosowania go, gdy obserwowane wskaźniki rozpowszechnienia wynoszą 5% lub są mniejsze [10].

**Wskaźnik Yule'a (Y)**. Wskaźnik Yule'a również określa zgodność pomiędzy badaczami, jest także miarą uwzględniającą zgodność przypadkową (Yule, 1912). Podobnie jak kappi zawiera się w przedziale od -1 do +1. Wielkość Y wykazuje znacznie mniejszą zależność od wskaźnika rozpowszechnienia niż kappi. Wartości mogą być zafałszowane, kiedy rozpowszechnienie wynosi poniżej 1% [9].

Wykorzystując tabele 2 x 2 (tab.1) ustalono, że diagnoza kliniczna będzie standardem. Powyższe wskaźnik obliczono według następujących wzorów [11]:

Tabela 1

Tabela 2 x 2		Diagnoza za pomocą SCAN	
		Obecna (+)	Negatywna (-)
Diagnoza kliniczna	Obecna (+)	a	b
	Brak (-)	c	d

$$PA = (a+d)/N \times 100\%$$

$$\text{Czułość} = a/(a+b)$$

$$\text{Specyficzność} = d/(d+c)$$

$$\text{Kappa } (\kappa) = (p_o - p_e) / (1 - p_e)$$

$p_o$  jest całkowitą zgodnością procentową i wyraża się wzorem  $(a+d)/N$

$p_c$  jest wielkością odpowiadającą zgodności przypadkowej i dla kappy Cohena liczone jest następująco:  $PC = ((a+b)/N)((a+c)/N) + ((c+d)/N)((b+d)/N)$

$$\text{Wskaźnik Yule'a (Y)} = \frac{\sqrt{ad} - \sqrt{bc}}{\sqrt{ad} + \sqrt{bc}}$$

$$\text{Rozpowszechnienie (base rate)} = (a+b)/N$$

## Wyniki

### Charakterystyka populacji badanej

W grupie 220 pacjentów badanych za pomocą kwestionariusza SCAN przeważały kobiety – stanowiły 66,8%. Średni wiek pacjentów wynosił 42 lata. W związku małżeńskim pozostawała mniej niż połowa pacjentów – 41,8%, a samotnie mieszkało 12,7%. Ponad połowa badanej grupy otrzymywała świadczenia społeczne, w tym rentę lub emeryturę 58,6% oraz zasiłek socjalny 55,5%. Pracowało 17,3% uczestników badania. W grupie badanej przeważali pacjenci wielokrotnie hospitalizowani na oddziałach stacjonarnych (1-3 hospitalizacji – 38,1%) i mający wcześniej objawy powyżej trzech epizodów zaburzenia (43,7%). Średni czas hospitalizacji wyniósł 101 dni, w tym na oddziale dziennym był dłuższy (143 dni) niż na oddziale stacjonarnym (109 dni). W grupie badanej najczęściej (u 41,8%) rozpoznawano zaburzenia psychiatryczne, w drugiej kolejności (u 37,3%) zaburzenia afektywne. Najmniej rozpoznano dotyczyło uzależnień – 0,5%.

Badania dotyczące porównania diagnozy za pomocą SCAN i diagnozy klinicznej (tab. 2, 3, 4)

Tabela 2

### Porównanie współczynników zgodności w poszczególnych grupach rozpoznania F0-F5 klasyfikacji ICD-10 w badaniu EDEN we Wrocławiu (Fc)

Grupa ICD	+		PA%	Czułość (sensitivity)	Specyficzność (specificity)	Kappa κ	Wskaźnik Yule'a Y	Base rate % <sup>3</sup>
	a	b						
	-							
	c	d						
F0	1	4	95	0,20	0,97	0,14 <sup>1</sup>	0,49	2
	6	20						
F1	1	0	81	1	0,80	0,04 <sup>1</sup>	NaN <sup>2</sup>	0,5
	43	17						
F2	60	27	84	0,69	0,94	0,65	0,71	39,5
	8	12						
F3	79	4	61	0,95	0,41	0,31	0,57	37,7
	81	56						
F4	19	1	54	0,95	0,50	0,14	0,62	9
	101	99						
F5	0	0	69	NaN	0,69	0 <sup>1</sup>	NaN <sup>2</sup>	
	69	15						

<sup>1</sup> Współczynnik kappa może być zafałszowany z powodu wskaźnika rozpowszechnienia (base rate) poniżej 5% (Grove i wsp. [10])

<sup>2</sup> Wskaźnik Yule'a może być zafałszowany z powodu wskaźnika rozpowszechnienia (base rate) poniżej 1% (Spitznagel, Helzer [9])

<sup>3</sup> Base rate: rozpowszechnienie

Na poziomie grup diagnostycznych zgodność oceniona za pomocą współczynnika kappa osiągnęła wartości od 0,14 do 0,65, a ustalana na podstawie współczynnika Yule'a (Y) od 0,57 do 0,71. Czulość mieściła się w przedziale od 0,69 do 0,95, a specyficzność od 0,41 do 0,94. W grupie zaburzeń psychotycznych F2 kappa wyniosła 0,65, Y – 0,71, przy czulości 0,69 i specyficzności 0,94. W grupie F3 kappa wyniosła 0,31, Y – 0,57, przy czulości 0,95 i specyficzności 0,41. Dla grupy F4 wskaźnik kappa był niski – 0,14, Y wyniósł 0,62, przy czulości 0,95 i specyficzności 0,50.

Tabela 3

**Porównanie współczynników zgodności w poszczególnych klasach rozpoznania (Fcc) klasyfikacji ICD-10 w badaniu EDEN we Wrocławiu**

Grupa ICD	+		PA%	Czulość (sensitivity)	Specyficzność (specificity)	Kappa κ	Wskaźnik Yule'a Y	Base rate % <sup>3</sup>
	a	b						
	-							
	c	d						
F06	1	4	96	0,20	0,98	0,16 <sup>1</sup>	0,53	2,3
	5	210						
F11	1	0	100	1	1	1 <sup>1</sup>	NaN <sup>2</sup>	0,5
	0	219						
F20	36	21	87	0,63	0,96	0,64	0,72	25,9
	7	156						
F25	5	18	88	0,22	0,95	0,21	0,41	10,5
	9	188						
F30	2	0	94	1	0,94	0,22 <sup>1</sup>	NaN <sup>2</sup>	0,9
	13	205						
F31	2	11	93	0,15	0,98	0,17	0,46	5,9
	5	202						
F32	12	11	70	0,52	0,72	0,13	0,25	10,5
	56	141						
F33	11	32	80	0,26	0,93	0,23	0,37	19,5
	12	165						
F40	2	0	81	1	0,81	0,07 <sup>1</sup>	NaN <sup>2</sup>	0,9
	41	177						
F41	1	1	64	0,50	0,64	0,01 <sup>1</sup>	0,15 <sup>2</sup>	0,9
	78	140						
F42	4	0	98	1	0,98	0,66 <sup>1</sup>	NaN	1,8
	4	212						
F43	1	9	96	0,10	0,10	0,15 <sup>1</sup>	0,66	4,5
	1	209						
F45	2	0	68	1	0,68	0,04 <sup>1</sup>	NaN	0,9
	70	148						

<sup>1</sup> Współczynnik kappa może być zafałszowany z powodu wskaźnika rozpowszechnienia (base rate) poniżej 5% (Grove i wsp. [10])

<sup>2</sup> Wskaźnik Yule'a może być zafałszowany z powodu wskaźnika rozpowszechnienia (base rate) poniżej 1% (Spitznagel, Helzer [9])

<sup>3</sup> Base rate: rozpowszechnienie

Na poziomie klas diagnostycznych ICD-10 porównanie diagnozy uzyskanej za pomocą SCAN z kliniczną wykazało zgodność na poziomie kappa od 0,13 do 0,64, Yule'a od 0,25 do 0,72, czułość od 0,15 do 0,63 oraz specyficzność od 0,72 do 0,98 dla wyników istotnych.

Tabela 4

**Zgodność rozpoznań za pomocą SCAN z klinicznymi w kategoriach diagnostycznych ICD-10 o znaczeniu klinicznym (Fcc.c)**

Kat. <sup>2</sup>	+	-	PA%	Czułość (sensitivity)	Specyficzność (specificity)	Kappa $\kappa$	Base rate % <sup>3</sup>	Liczba diagnoz za pomocą SCAN	Liczba diagnoz klinicznych
	+ a	b							
<sup>A</sup> Zaburzenia spowodowane nadużywaniem substancji	1	0	98	1	0,98	0,32 <sup>1</sup>	1	5	1
	4	215							
<sup>B</sup> Zaburzenia psychotyczne	52	14	88	0,79	0,92	0,71	30	65	66
	13	141							
<sup>C</sup> Zaburzenia schizofreniczne	5	14	90	0,26	0,96	0,25	9	14	19
	9	192							
<sup>D</sup> Zaburzenia maniackalne	4	4	93	0,50	0,94	0,30 <sup>1</sup>	4	16	8
	12	200							
<sup>E</sup> Zaburzenia depresyjne	48	20	70	0,71	0,69	0,36	31	95	68
	47	105							
<sup>F</sup> Zaburzenia lękowe	8	2	63	0,80	0,62	0,09	5	87	10
	79	131							
<sup>G</sup> Reakcja na stres i zaburzenia adaptacyjne	1	9	95	0,10	0,99	0,15	5	2	10
	1	209							
<sup>H</sup> Zaburzenia występujące pod postacią somatyczną	2	0	68	1	0,68	0,04 <sup>1</sup>	0,9	72	2
	70	148							
<sup>I</sup> Inne zaburzenia nerwicowe			95	NaN	0,95	0 <sup>1</sup>	0	10	0
	10	210							
<sup>J</sup> Zaburzenia odżywiania się			100	NaN	1	NaN <sup>1</sup>	0	0	0
		220							
<sup>K</sup> Zaburzenia afektywne nieokreślone			67	NaN	0,67	0 <sup>1</sup>	0	73	0
	73	147							

<sup>1</sup> Współczynnik kappa może być zafałszowany z powodu wskaźnika rozpowszechnienia (base rate) poniżej 5% (Grove i wsp. [10])

<sup>2</sup> Kategorie kliniczne zgodne z ICD-10: F10.1-19.9 – A, F20.0-29.0 – B, F25.0-25.9 – C, F30.0-31.2 – D, F31.3-31.5, 32.0-33.9, 34.1 – E, F40.0-42.9 bez 41.2 – F, F43.0-43.9 – G, F45.0-45.9 – H, F48.0-48.9 – I, F50.0-50.2 – J, F39 – K

<sup>3</sup> Base rate: rozpowszechnienie

Szczegółowa analiza zgodności diagnostycznej na poziomie rozpoznań w kategoriach diagnostycznych o znaczeniu klinicznym potwierdza znaczną zgodność wśród zaburzeń psychotycznych, z wyjątkiem zaburzeń schizofrenicznych. Za pomocą



narzędzia SCAN uzyskano 65 diagnoz z tej kategorii, czyli prawie tyle samo (66), ile postawił psychiatra.

### Omówienie wyników

W badaniu wrocławskim na poziomie grup diagnostycznych wyniki porównania zgodności pomiędzy diagnozą uzyskaną za pomocą SCAN a diagnozą kliniczną były zbliżone do wartości uzyskanych w badaniach innych autorów, szczególnie dla diagnoz z grupy F2. Dla tej grupy diagnostycznej klasyfikacji ICD-10 osiągnięto względnie wysokie wskaźniki zgodności, czułości i specyficzności, co świadczy o zaletach narzędzia SCAN w diagnostyce zaburzeń z tej grupy. W przypadku zaburzeń z grup F3 i F4, niejednorodnych pod względem profilu zaburzeń, osiągnięto niższą zgodność niż w porównywanych badaniach. Dla współczynnika Yule'a otrzymano wynik zgodności przybliżony do wyników innych badań. Należy jednak zwrócić uwagę, że w większości przypadków, przy wysokiej czułości, diagnoza za pomocą SCAN była podobna do rozpoznania psychiatry. Natomiast niższe parametry specyficzności wskazują, że wśród rozpoznań postawionych za pomocą SCAN wystąpiły rozpoznania fałszywie pozytywne. Na poziomie klas diagnostycznych ICD-10 zwraca uwagę dość wysoka zgodność, jeśli chodzi o czułość i specyficzność, dla zaburzeń psychotycznych o typie schizofrenii, w porównaniu z innymi badaniami. W przypadku zaburzeń schizoaфекtywnych przy wysokiej specyficzności uzyskano niewielki poziom czułości i słabą zgodność, która niewiele się poprawiła, gdy analizowano ją współczynnikiem Yule'a Y. Wynika z tego, iż SCAN nie był wystarczająco czuły w porównaniu z badaniem psychiatry w klasie zaburzeń schizoaфекtywnych, co potwierdzają też inne badania. Bardzo podobna sytuacja występuje w klasie zaburzeń aфекtywnych dwubiegunowych. W klasach zaburzeń depresyjnych i depresyjnych nawracających uzyskano niewielką zgodność, która jednak wzrosła do zadowalającej przy analizie przeprowadzonej z zastosowaniem współczynnika Y. Niższe wskaźniki czułości oraz dość dobre wyniki specyficzności wskazują, że za pomocą SCAN nie uzyskiwano zbyt wielu diagnoz fałszywie pozytywnych. Otrzymane wyniki znajdują potwierdzenie w innych badaniach, gdzie uzyskano wyższą zgodność, ale także niskie wskaźniki czułości przy wysokiej specyficzności. Szczegółowa analiza zgodności diagnostycznej na poziomie rozpoznań w kategoriach diagnostycznych o znaczeniu klinicznym potwierdza znaczną, porównywalną z literaturą zgodność wśród zaburzeń psychotycznych z wyjątkiem zaburzeń schizoaфекtywnych. W kategorii diagnostycznej „zaburzenie schizoaфекtywne” uzyskano dane pokrywające się z wcześniejszymi wynikami z poziomu klas diagnostycznych. Słaba czułość spowodowała, że program SCAN postawił 14 rozpoznań w porównaniu z 19, które postawił psychiatra, nie wykazywał jednakże rezultatów fałszywie dodatnich. W przypadku zaburzeń z kategorii depresji, przy zadowalającej zgodności, nieznacznie niższej niż podaje piśmiennictwo, uzyskano dość dobrą czułość, ale niższą specyficzność w porównaniu z danymi literatury (95 diagnoz uzyskanych za pomocą SCAN w porównaniu z 68 klinicznymi). Podobnie w kategorii zaburzeń lękowych, przy niskiej zgodności, wysokiej czułości (program SCAN w większości przypadków stawiał rozpoznanie z tej samej kategorii, co psychiatra), osiągnięto niższą specyficzność, co oznaczało, że psychiatra postawił ponad 8 razy mniej diagnoz z tej kategorii.

W kategorii „reakcja na stres i zaburzenia adaptacyjne” diagnozy za pomocą SCAN nie różniły się znacznie od klinicznych (były to tylko 2 rozpoznania w porównaniu z 10 postawionymi przez psychiatrę). Z kolei w przypadku zaburzeń somatyzacyjnych program SCAN postawił bardzo dużo diagnoz w porównaniu z psychiatrą (72 vs. 2); podobnie w kategorii zaburzeń afektywnych nieokreślonych (73 vs. 0).

Na podstawie powyższej analizy zgodności diagnostycznej narzędzia SCAN i diagnozy stawianej przez psychiatrę na poziomie grup, klas i kategorii diagnostycznych klasyfikacji ICD-10 można stwierdzić, że uzyskane wyniki badania zgodności zbliżone były na poszczególnych poziomach diagnostycznych do tendencji wcześniej opisywanych w literaturze, chociaż wystąpiły pewne odmienności charakterystyczne dla ośrodka wrocławskiego. Przede wszystkim należy podkreślić, że uzyskano dobrą zgodność dla zaburzeń psychotycznych ze słabszą jedynie czułością dla zaburzeń schizoafektywnych. Być może wynika ona z bardziej złożonej struktury zaburzenia schizoafektywnego, które obejmuje objawy psychozy, depresji i manii ze specyficznymi kryteriami czasowymi. Szczególne zwrócenie uwagi na kryteria czasowe może przyczynić się do poprawy czułości narzędzia SCAN w wykrywaniu tego zaburzenia. Ponadto uwzględniając dobrą specyficzność narzędzia, jaką uzyskano w zakresie dwóch z trzech składowych (manii i psychozy), a słabą w zakresie zaburzeń depresyjnych, można przypuszczać, że program SCAN zamiast zaburzenia schizoafektywnego o typie depresyjnym postawił diagnozę zaburzenia depresyjnego. Należy w tym miejscu podkreślić, że w literaturze naukowej nowoczesne kryteria zaburzenia schizoafektywnego pozostają nadal kontrowersyjne. Dowodem na to może być fakt, iż w badaniu dotyczącym wpływu różnic kulturowych na diagnostykę psychiatryczną pomiędzy psychiatrami norweskimi a rosyjskimi tylko 20% psychiatrów z obu krajów trafnie postawiło to rozpoznanie. Norwegowie częściej rozpoznawali zaburzenia afektywne, natomiast Rosjanie schizofrenię [12].

Bardzo charakterystyczna była też znaczna dysproporcja w liczbie rozpoznanych zaburzeń depresyjnych i lękowych na korzyść programu SCAN, którego diagnozy w większości przypadków były zgodne z klinicznymi. Tak wysoki wskaźnik diagnoz fałszywie dodatnich dla kategorii diagnostycznych jest inny niż podaje to piśmiennictwo. Warto zauważyć, że tam, gdzie w toku prac nad międzynarodową klasyfikacją kryteria diagnostyczne nie ulegały dużym zmianom (zaburzenia psychotyczne), osiągnięto wyższą zgodność, natomiast w przypadku innych zaburzeń (lękowych i depresyjnych), których kryteria ulegały pozycjonowaniu w czasie, wyniki zgodności były niższe. Może to wskazywać na słabszą znajomość psychiatrów nowoczesnych kryteriów diagnostycznych części zaburzeń. W okresie prowadzenia badania wprowadzano nowe zasady rozliczeń z ówczesną kasą chorych w ramach wdrażania reformy finansowania służby zdrowia. Mogły zatem występować błędy systemowe w kodowaniu rozpoznań w celu uzasadnienia przyjęcia do szpitala psychiatrycznego lub dla uzyskania wyższego finansowania osobodnia leczenia, tym bardziej że zaburzenia lękowe były relatywnie nisko opłacane. Takie błędy systemowe w kodowaniu rozpoznań są opisywane w literaturze. Uważa się, że są one naturalną konsekwencją złych rozwiązań refundacyjnych za usługi szpitalne [13, 14]. Stąd nadrozpoznawalność zaburzeń depresyjnych można wyjaśnić przesunięciem do tej kategorii zaburzenia schizoafektywnego.

## Wnioski

Wyniki wrocławskiego badania zgodności diagnozy psychiatrycznej z diagnozą za pomocą SCAN opartą na ścisłych kryteriach klasyfikacji ICD-10 są zbliżone do wcześniejszych opisywanych w literaturze. Wykazano wyższy poziom zgodności w zaburzeniach psychotycznych, z wyjątkiem zaburzeń schizoafektywnych oraz niższe wyniki zgodności w zaburzeniach nerwicowych i afektywnych, w których program SCAN ujawnił więcej rozpoznań niż psychiatra. Wyniki wskazują na słabszą znajomość psychiatrów nowoczesnych kryteriów diagnostycznych części zaburzeń. Częściowo może to mieć związek z błędami systemowymi w kodowaniu rozpoznań.

### Оценка согласности психиатрических диагнозов с диагностическими критериями классификации ICD-10 при помощи анкеты

#### Содержание

**Задание.** Авторы работы, в проведенном исследовании, хотели проверить в какой мере психиатр в своей диагностике применяет критерии ICD-10. Для этого, проведены сравнения диагноза, поставленного психиатром при выписке больного с диагнозами, поставленными при приеме с использованием анкеты SCAN ver 2.1 у тех же самых пациентов.

**Метод.** Поставленные диагнозы программой (алгоритм) IShell анкеты SCAN были сравнены с клиническими распознаваниями психиатра согласно с критериями ICD-10 тремя способами: в диагностических группах (Fc), в диагностических классах (Fcc) и на уровне диагностических категорий с клиническим значением (Fcc, c). Оценка достоверности анкеты установлена при использовании коэффициента Каппа Cohen, чувствительности (sensitivity), специфичности (specificity) и коэффициента Yule'a.

**Результаты.** На уровне диагностических групп получена согласность, измеренная коэффициентом Каппа с 0,14 до 0,65, а Yule'a Y с 0,57 до 0,71. Целостность умещалась в границах 0,69–0,95, а специфичность в границах 0,41–0,94. В группе психотических нарушений F 2 Каппа равнялась 0,65 Yule'a Y 0,71, чувствительность 0,69 и специфичности 0,69. В группе F 3 Каппа равнялась 0,31, Yule'a Y 0,57, чувствительность 0,95 и специфичность 0,41. Для группы F 4 показатель Каппа был низкий – 0,14, Yule'a 0,62, чувствительность 0,95 и специфичность 0,50.

**Выводы.** В ходе исследований обнаружен высший уровень согласности при психотических нарушениях, за исключением шизоаффективных нарушений, а также низкие результаты согласности при невротических состояниях и аффективных нарушениях, в которых то анкета SCAN проявила больше диагнозов, чем психиатр. Это обстоятельство может указывать на слабое знание современных диагностических критерий части нарушений психиатрами, а также это может быть связано с системными ошибками при кодировании диагнозов.

### Studie an der Übereinstimmung der psychiatrischen Diagnosestellung mit diagnostischen Klassifizierungskriterien ICD-10 mit Hilfe des SCAN - Fragebogens

#### Zusammenfassung

**Ziel.** Die Autoren wollten in der durchgeführten Studie prüfen, wie oft der Psychiater bei seiner Diagnosestellung die ICD-10 Kriterien anwendet. Für diesen Zweck verglich man die Diagnosen, die vom Psychiater bei der Entlassung gestellt wurden, mit den Diagnosen, die bei der Aufnahme mit Hilfe des Analysewerkzeuges SCAN, Ver. 2.1 bei denselben Patienten gestellt wurden.

**Methode.** Die durch das Programm (Algorithmus) IShell des Werkzeuges SCAN gestellten Diagnosen wurden mit den klinischen Diagnosen des Psychiaters nach den Klassifizierungskriterien ICD-10 auf drei Weisen verglichen: in den diagnostischen Gruppen (Fc), in den diagnostischen

Klassen (Fcc) und in den diagnostischen Kriterien mit klinischer Bedeutung (Fcc.c). Die Beurteilung der Richtigkeit des Mittels wurde mit Hilfe des Kappa - Faktors von Cohen, Sensibilität (sensitivity), Eigentümlichkeit (specifity) und des Yule - Faktors bestimmt.

**Ergebnisse.** Auf dem Level der diagnostischen Gruppen wurde die Übereinstimmung erreicht, die mit dem Kappa - Faktor von 0,14 bis 0,65, Yule - Faktor von 0,57 bis 0,71 gemessen wurde. Die Sensibilität lag zwischen 0,69 bis 0,95 und die Eigentümlichkeit zwischen 0,41 bis 0,94. In der Gruppe der psychotischen Störungen F2 betrug Kappa 0,65, Yule Y 0,71 und die Sensibilität 0,95 und die Eigentümlichkeit 0,41. Für die Gruppe F4 war der Kappa - Index niedrig 0,14, Yule Y 0,62 und die Sensibilität 0,96, die Eigentümlichkeit 0,50.

**Schlussfolgerungen.** Es wurde ein höherer Level der Übereinstimmung in den psychotischen Störungen gezeigt; die Ausnahme waren schizoaktive Störungen. Es erwies sich, dass die Ergebnisse der Übereinstimmung bei den Nervenstörungen und affektiven Störungen niedriger waren, in denen SCAN mehr Diagnosen als der Psychiater gestellt hatte. Es kann darauf zeigen, dass die modernen diagnostischen Kriterien eines Teils der Störungen von den Psychiatern weniger bekannt sind und es kann auch teilweise einen Zusammenhang mit den Systemfehlern beim Kodieren von Diagnosen haben.

### L'analyse de la compatibilité du diagnostic psychiatrique et des critères diagnostiques de la classification d'ICD-10 avec l'usage du questionnaire SCAN

#### Résumé

**Objectif.** Les auteurs veulent voir comment le psychiatre se sert des critères ICD-10 dans sa pratique diagnostique. Ils comparent donc le diagnostic du début de l'hospitalisation et après à l'aide du questionnaire SCAN version 2.1.

**Méthode.** Les diagnostics faits par le programme Ishell sont comparés avec les diagnostics des psychiatres basés sur les critères de la classification ICD-10. Cette comparaison est faite trois fois : groupes diagnostiques (Fc), classes diagnostiques (Fcc), catégories diagnostiques (Fcc.c). L'analyse de la comptabilité est basée sur le coefficient kappa de Cohen, de sensibilité, de spécificité et sur le coefficient d'Yule.

**Résultats.** Au niveau des groupes diagnostiques on note le coefficient kappa de 0,14 jusqu'à 0,65 et le coefficient d'Yule de 0,57 jusqu'à 0,71. La sensibilité – de 0,69 jusqu'à 0,95, la spécificité – de 0,41 jusqu'à 0,94. Dans le groupe des troubles psychiques F2 kappa atteint 0,65, d'Yule – 0,57, la sensibilité – 0,95, la spécificité – 0,41. Dans le groupe F4 le coefficient kappa – 0,14, d'Yule – Y=0,62, la sensibilité – 0,95, la spécificité – 0,50.

**Conclusions.** On démontre le niveau plus élevé de la comptabilité dans les troubles psychotiques sauf les troubles schizoaffectifs, et le niveau moins élevé dans les troubles nerveux et affectifs où le SCAN diagnostique plus des troubles que le psychiatre. Cela peut résulter de la plus faible connaissance des critères diagnostiques nouveaux chez certains psychiatres ou des erreurs du système dans les codes diagnostiques.

#### Piśmiennictwo

1. Sartorius N, Ustun TB, Korten A, Cooper JE, van Drimmelen J. *Progress toward achieving a common language in psychiatry, II: Results from the international field trials of the ICD-10 diagnostic criteria for research for mental and behavioural disorders*. Am. J. Psychiatry 1995; 152 (10): 1427–1437.
2. Brugha TS, Nienhuis F, Bagchi D, Smith J, Meltzer H. *The survey form of SCAN: the feasibility of using experienced lay survey interviewers to administer a semi-structured systematic clinical assessment of psychotic and non-psychotic disorders*. Psychol. Med. 1999; 29 (3): 703–711.
3. Isohanni M, Makikyro T, Moring J, Rasanen P, Hakko H, Partanen U, Koiranen M, Jones P. *A comparison of clinical and research DSM-III-R diagnoses of schizophrenia in a Finnish na-*

- tional birth cohort. Clinical and research diagnoses of schizophrenia.* Soc. Psychiatry Psychiatr. Epidemiol. 1997; 32 (5): 303–308.
4. Adamowski T, Kiejna A, Hadryś T. *System SCAN – częściowo strukturalizowany wywiad oparty na kryteriach diagnostycznych.* Psychiatr. Pol. 2006; 4: 751–760.
  5. Kallert TW, Glockner M, Priebe S, Briscoe J, Rymaszewska J, Adamowski T, Nawka P, Reguliova H, Raboch J, Howardova A, Schützwohl M. *A comparison of psychiatric day hospitals in five European countries: implications of their diversity for day hospital research.* Soc. Psychiatry Psychiatr. Epidemiol. 2004; 39 (10): 777–788.
  6. Schützwohl M, Jarosz-Nowak J, Briscoe J, Szajowski K, Kallert T. *Eden Study Group. Inter-rater reliability of the Brief Psychiatric Rating Scale and the Groningen Social Disabilities Schedule in a European multi-site randomized controlled trial on the effectiveness of acute psychiatric day hospitals.* Int. J. Methods Psychiatr. Res. 2003; 12 (4): 197–207.
  7. Cohen J. *A coefficient of Agreement for Nominal Scales.* Edu. Psychol. Measur. 1960; 20: 37–46.
  8. Landis JR, Koch GG. *The measurement of observer agreement for categorical data.* Biometr. 1977; 33 (1): 159–174.
  9. Spitznagel EL, Helzer JE. *A proposed solution to the base rate problem in the kappa statistic.* Arch. Gen. Psychiatry 1985; 42 (7): 725–728.
  10. Grove WM, Andreasen NC, McDonald-Scott P, Keller MB, Shapiro RW. *Reliability studies of psychiatric diagnosis. Theory and practice.* Arch. Gen. Psychiatry 1981; 38 (4): 408–413.
  11. Glaros AG, Kline RB. *Understanding the accuracy of tests with cutting scores: the sensitivity, specificity, and predictive value model.* J. Clin. Psychol. 1988; 44 (6): 1013–1023.
  12. Rezvyy G, Rießvold T, Parniakov A, Olstad R. *A comparative study of diagnostic practice in psychiatry in Northern Norway and Northwest Russia.* Soc. Psychiatry Psychiatr. Epidemiol. 2005; 40 (4): 316–323.
  13. English JT, Sharfstein SS, Scherl DJ, Astrachan B, Muszyński IL. *Diagnosis-related groups and general hospital psychiatry.* Am. J. Psychiatry 1986; 143 (2): 131–139.
  14. English JT, McCarrick RG. *DRGs: an overview of the issues.* Gen. Hosp. Psychiatry 1986; 8 (5): 359–364.

Otrzymano: 7.06.2005

Zrecenzowano: 2.11.2005

Przyjęto do druku: 1.03.2006

Adres: Tomasz Adamowski

Katedra i Klinika Psychiatrii AM

50-367 Wrocław, ul. Pasteura 10