

ICD-9-CM coderen in de psychiatrie

door P. Dingemans, C. Hooijer en J. Dutilh

Inleiding

In Nederland wordt het International Classification of Diseases (ICD-9CM) systeem gebruikt in het Centrale Patiëntenregister Psychiatrische Instituten. Dit classificatiesysteem wordt door de Hoofdinspectie van de Geestelijke Volksgezondheid samen met de Nationale Ziekenhuisraad operationeel gehouden. Doel van het gebruik van de ICD in dit verband is onder andere het verzamelen van epidemiologische gegevens.

In 1980 betoogden Giel en Bogaarts dat '... de vertaling van individuele in groepsverschijnselen... diagnostische classificatie...' van iemands ziekte genoemd kan worden. De voorwaarden daartoe werden door hen benoemd als: het opzoeken van ziekte-eenheden, aetiologie en beloop, welomschreven klassen, criteria voor indeling in klassen samen met uitsluitcriteria en een rangorde van die klassen. Ze stelden dat diagnostische classificatie met de ICD bestaat uit twee fasen: het uitschrijven van de diagnose in duidelijke en algemene termen (classificeren) en het zoeken van de juiste ICD-klasse bij deze classificatie (coderen). In hun onderzoek toonden Giel en Bogaarts (1980) vervolgens aan dat zowel het classificeren als het coderen problematische zaken blijken te zijn. Zij concludeerden: 'Hij (de clinicus) zal op zijn minst moeten vastleggen welk ziektebeeld de aanleiding tot opname was, wat de onderliggende psychopathologische syndromen waren, of er somatische syndromen zijn en of er verband leek te bestaan met één van de psychopathologische syndromen'. En met betrekking tot het coderen: 'Wat de codeur moet doen staat goeddeels in de ICD zelf', en: 'Als de clinicus

Schrijvers zijn respectievelijk verbonden als klinisch psycholoog aan de Rijks Universiteit Utrecht en Amsterdam; psychiater aan de Valeriuskliniek van de Vrije Universiteit van Amsterdam, en psychiater aan de Universiteit van Amsterdam. Correspondentieadres: Psychiatrisch Centrum AZUA, Tafelbergweg 25, 1105 BC Amsterdam.

geen rangorde heeft gegeven in de ziektebeelden die hij in zijn omschrijving noemt, dan zal de codeur een afgesproken hiërarchie van syndromen hanteren'.

Een jaar later (1981) stelde Lit dat het met de ICD onmogelijk is om te diagnosticeren en te coderen en dat deze vervangen zou moeten worden door een systeembenadering. Na de bespreking van een casus waaraan hij een dergelijke systeembenadering toelichtte, kwam hij vervolgens tot de conclusie: 'In tegenstelling tot de gebruikelijke classificaties blijkt het mogelijk om op basis van een systeembenadering een wetenschappelijk verantwoorde ordening van psychiatrische gegevens tot stand te brengen'.

In 1982 merkten Slooff e.a. met betrekking tot de ICD en de DSM-III (Diagnostic Statistical Manual, 3e versie; APA, 1980) in relatie tot het classificeren van niet affectieve functionele psychosen op, dat de eerste (ICD) erg afhankelijk is van de interpretatie van de diagnosticus en dat de betrouwbaarheid matig is. Hun conclusie is: 'De ICD schiet te kort in een aantal voorwaarden die in beginsel aan elk classificatiesysteem moeten worden gesteld, zoals definiëring en weging van kenmerken. De DSM-III is in dit opzicht beter onderbouwd en duidelijker'. Het gematigd optimisme over de ICD heeft in enkele jaren plaats gemaakt, zo lijkt, voor regelrecht pessimisme. Nieuwe verwachtingen gelden de DSM-III. Daarmee is, in kort bestek, ook de Nederlandse literatuur over classificeren en coderen in de psychiatrie van de afgelopen jaren samengevat.

Vraagstelling

Ons onderzoek kwam voort uit een pessimisme over het gebruik van de ICD in de psychiatrische praktijk. De jaaroverzichten van Leidschen-dam met betrekking tot ICD-classificatiecodes, die de deelnemende instanties elk jaar worden toegezonden uit onze kliniek werden voor wat betreft de jaren 1978 en 1981 met elkaar vergeleken. Deze keuze werd bepaald door het gegeven dat het toenmalige Paviljoen III van de Amsterdamse binnenstad verhuisde naar de periferie (Bijlmermeer) en geleid zou kunnen hebben tot opname van verschillende patiëntenpopulaties. In de vergelijking van de uitboekcodes viel op dat in 1978, 108 classificaties van de niet-organische psychosen (196) als schizofreen werden benoemd (55%) en van alle neurotische stoornissen (215) er 167 als depressief werden gecodeerd (78%). In 1981 bleken die percentages respectievelijk: 6/10% en 31%. Een intrigerend verschijnsel dat ons ook deed denken aan goede en slechte schizofrenie- respectievelijk depressiejaren. Navraag leerde dat vanaf 1978 descriptieve uitboek-diagnosen (bv.: paranoïsch-hallucinatoir syndroom overgaand in een melancholisch syndroom naar aanleiding van overlijden van moeder bij een man met schizoïd karakter) door daartoe opgeleide codeurs werden omgezet in ICD-codes. Codeurs, verder aangeduid als medi-

sche codeurs, zijn administratieve medewerkers in dienst van het ziekenhuis, welke getraind zijn in, en het tot hun dagelijkse bezigheid rekenen, om medische diagnoses (inclusief psychiatrische) om te zetten in ICD-9-(CM)-codes. Vanaf 1980 (dus in 1981) werd dit werk overgenomen door assistentpsychiaters op het ochtendrapport bij uitboeken van de patiënt. Omdat de behandelende artsen meer gegevens over de patiënt hebben dan de medische codeurs, zo was de redenering, zouden ze ook beter moeten kunnen diagnostiseren en coderen dan de laatsten. Dit leek ons een plausibele verklaring. Echter, dat er grote verschillen in vaardigheid, training, inzicht en interesse in classificeren en coderen kunnen blijven bestaan, maakte het ochtendrapport ons duidelijk. Ernstige twijfel over de betrouwbaarheid van ICD-codering was daarvan het gevolg. Twijfel die zich ook generaliseerde naar de 'Mededelingen uit het Centrale Patiëntenregister' (o.a. Van Weerden-Dijkstra en Giel, 1974). Besloten werd daarom een explorerend onderzoek op te zetten naar de betrouwbaarheid (interbeoordelaar) van het ICD-coderen*.

Er worden verschillende diagnostische systemen gebruikt in de psychiatrie; ICD- en DSM-III-classificatie zijn van het nominale (klinische syndroom) type.

Voor de beoordeling van tussen-beoordelaarsbetrouwbaarheid zijn verschillende technieken in gebruik: percentageovereenkomst (ratio van overeenstemming en totaal mogelijke overeenstemming), chikwadraattoets (de afwijking van overeenkomst tussen twee beoordelaars die afwijkt van kansovereenkomst), gewogen Kappa (percentage van overeenkomst waarbij de mate van afwijking tussen beoordelaars bepaald gewicht wordt toegekend die voor de correctie van de kansovereenkomst gebruikt wordt) en Kappa (overeenkomst die gecorrigeerd is voor de kans op toevalsovereenkomst). De meest gebruikte maat is de eerste: de overeenkomst tussen twee beoordelaars uitgedrukt in procenten. Op dit gebruik is de laatste jaren nogal kritiek gekomen. De critici zeggen dat een patiënt verschillende diagnoses kan hebben (iemand kan èn schizofreen èn alcoholist èn zwakbegaafd zijn.) Dit zou moeten betekenen dat een beschouwing van hoofd- en nevendiagnosen in de beoordeling van de betrouwbaarheid op zijn minst tot uitdrukking zou moeten komen. Voorts wordt ook betoogd dat de overeenkomst tussen twee beoordelaars, uitgedrukt in een percentage, gecorrigeerd dient te worden voor toevalsovereenkomst (Grove e.a., 1981). Dit heeft er mede toe geleid dat Kappa (Cohen, 1968) als maat voor tussen-beoordelaarsbetrouwbaarheid (met kanscorrectie) als het minst slechte alternatief gekozen wordt.

Recentelijk hebben Mezzich e.a. (1981) drie methoden voor tussen-beoordelaarsbetrouwbaarheidsbepalingen beschreven die meervoudi-

* Een vooronderzoek (Dutilh en Dingemans, 1984) bevestigde het vermoeden dat de betrouwbaarheid van het ICD-coderen bij assistent-psychiaters laag was.

ge diagnoses in aanmerking nemen en corrigeren voor kansovereenkomst tussen twee of meer beoordelaars. Deze benoemden zij als: 'proportional overlap'-, 'intraclass correlation'- en 'rank correlation'-procedure. Hierover zeggen zij: 'The proportional overlap and intraclass correlation procedures are addressed to the situation in which multiple diagnoses for a subject is a list of non-ordered categories. The rank correlation procedure is appropriate when the order of the categories is important' (Mezzich e.a., 1981). Wij gingen er in dit onderzoek van uit, dat de volgorde van de opeenvolgende classificaties per beoordelaar er niet toe deed en dat de verschillende classificaties samen de uiteindelijke betrouwbaarheid tussen de beoordelaars bepaalden. Daarom kozen we er voor om de zogenaamde 'proportionele overlap'-procedure van Mezzich e.a. voor de betrouwbaarheidsbepalingen te gebruiken, welke corrigeert voor kansovereenkomst (uitgedrukt in Kappa) en deze aan te vullen met een maat voor overeenkomst welke niet corrigeert voor kansovereenkomst.

Methode

Uit het totale gestandaardiseerde ontslagbrievenbestand van 1981 werden er 20 at random uitgehaald. Deze ontslagbrieven bevatten omschrijvingen betreffende reden van opname, toestandbeeld bij opname (CHAM-systeem; Silberman, 1971), psychiatrische en somatische onderzoeksgegevens, biografische c.q. auto-anamnestiche en hetero-anamnestiche gegevens, aetiologische factoren, medicatie, decursus en beschrijvende ontslagdiagnose en ICD-code(s). Deze 20 ontslagbrieven werden gebruikt in een A- en B-deel van onderzoek. In het A-deel werd alléén de beschrijvende ontslagdiagnose gegeven. In het B-deel werd de ontslagbrief in zijn totaal ondergebracht. Voor beide delen gold dat de naam van de patiënt (en overige persoonsgegevens) werden uitgewist. Voor het B-deel gold tevens dat in de ontslagbrieven de ICD-codes werden uitgewist. Het A- en B-deel betrof dezelfde 20 patiënten.

Dit onderzoekmateriaal werd aangeboden aan twee universiteitsklinieken. Uit de eerste kliniek namen 2 assistent-psychiaters, 2 psychiaters en 1 medisch codeur deel. In de tweede kliniek namen 2 assistent-psychiaters, 1 psychiater en 2 medische codeurs deel. In totaal werden in deel A en deel B afzonderlijk 20 patiënten 10 maal beoordeeld. Deel A werd eerst aangeboden, gevolgd door deel B nadat alle gegevens betreffende het eerste deel uit beide klinieken waren geretourneerd. De tijd tussen aanbidding van deel A en B schatten wij op circa een half jaar.

De gestelde vragen waren voor beide delen hetzelfde:

- A. Welke ICD-9 code(s) zou u op grond van de informatie in de tabel aan deze patiënt geven?
- B. Welke nevendiagnose(n), in de vorm van een ICD-9 code, zou u (eventueel) aan deze patiënt geven?

- C. Welke differentiaaldiagnose(n), in de vorm van een ICD-9 code zou u aan deze patiënt geven?
- D. Wilt u uw keuze en/of volgorde van de ICD-9 codes kort toelichten?

Bij deze vraagstellingen konden de coderingen van somatische aandoeningen achterwege blijven. Uitgegaan werd van een ICD-9-CM codering die voor beide klinieken gebruikelijk is.

Resultaten

In de uiteindelijke statistische verwerking van de gegevens werd voor de bepaling van de interbeoordelaarbetrouwbaarheid van het coderen uitgegaan van de hoofd- en nevendiagnosen samen (vragen A en B). De ICD-9-CM codes werden gehercodeerd in 28 klassen. Geen rekening werd gehouden binnen een klasse met de verschillende sub-klasse aanduidingen. Bijvoorbeeld: als iemand als lijdend aan een schizofrene psychose van het paranoïde subtype werd gecodeerd door één beoordeelaar en door de andere beoordeelaar als lijdend aan schizofrene psychose van het gedesorganiseerde type, dan werden beide diagnosen 'teruggecodeerd' naar 'schizofrenie' (ICD 295.). Kort samengevat: in de herco-

Tabel 1: Gemiddelde betrouwbaarheid van ICD-9-CM-codering van twee psychiatrische universiteitsklinieken

groepen	onderzoek	
	A-deel	B-deel
assistent- psychiaters (n = 4)	a 37 ¹	41
	b 15	15
	c 26	31
	d > 50 ²	> 50
psychiaters- supervisors (n = 3)	a 50	49
	b 20	18
	c 38	38
	d > 40	> 40
medische codeurs (n = 3)	a 51	56
	b 18	19
	c 40	46
	d > 30	> 20

Legenda: a percentuele overeenkomst

b kans op overeenkomst

c overeenkomst gecorrigeerd voor kans (Kappa)

d p-waarde van Kappa

1 waarden zijn vermenigvuldigd met 100 en hebben betrekking op 20 gecodeerde patiëntdiagnosen.

2 df = 19, tweezijdig getoetst.

dering werd geen rekening gehouden met het 4e en/of 5e cijfer uit de code; in de analyse werd uitgegaan van de overeenstemming binnen één klasse. Dit heeft tot resultaat dat de feitelijke diagnostische coderings-overeenkomst lager zal zijn dan de gerapporteerde overeenkomst. De analyses van de beoordelingsgegevens werden met een voor dat doel geconstrueerd computerprogramma uitgevoerd, uitgaande van de door Mezzich e.a. (1981) opgestelde uitgangspunten.

In tabel 1 zijn de beoordelaars uit de beide klinieken in categorieën ondergebracht en zijn de overeenkomsten uitgesplitst naar het A- en B-deel van het onderzoek. Uit deze gegevens blijken drie dingen:

1. De interbeoordelaarbetrouwbaarheid, uitgedrukt in Kappa-waarden, neemt iets toe als we de groepen met elkaar vergelijken. De medische codeurs onderling zijn betrouwbaarder dan de psychiaters, welke op hun beurt betrouwbaarder zijn dan de assistent-psychiaters; dit geldt voor zowel het A- als het B-deel.

2. Binnen elke codeurs-groep draagt een toevoeging van informatie over de patiënt (A-deel weinig en in het B-deel veel informatie over dezelfde patiënt) niet significant bij aan een verbetering of verhoging van de interbeoordelaarbetrouwbaarheid.

3. De percentuele overeenkomst tussen de beoordelaars is altijd groter dan de overeenkomst gecorrigeerd voor kans en geen enkele gevonden Kappa-waarde bereikt statistische significantie.

In tabel 2 zijn de interbeoordelaargegevens voor alle codeurs per kliniek en per onderzoeksonderdeel samengebracht. Uit deze gegevens blijkt het volgende:

1. In kliniek A daalt de interbeoordelaarbetrouwbaarheid iets, in kliniek B neemt deze iets toe als naast de beschrijvende ontslagdiagnose (onderzoek A) de gehele gestandaardiseerde ontslagbrief (onderzoek B) wordt gegeven.

Tabel 2: Gemiddelde betrouwbaarheid van ICD-9-CM-codering van 20 patiënten uitgesplitst naar de twee universiteitsklinieken

	kliniek A		kliniek B	
	onderzoek A	onderzoek B	onderzoek A	onderzoek B
totaal aantal	a 50 ¹	44	33	36
codeurs per	b 17	19	14	13
kliniek	c 39	31	23	26
(n = 5)	d >30 ²	> 50	> 50	> 40

Legenda: a percentuele overeenkomst

b kans op overeenkomst

c overeenkomst gecorrigeerd voor kans (Kappa)

d p-waarde van Kappa

1 waarden zijn vermenigvuldigd met 100.

2 df = 19, tweezijdig getoetst.

2. De gemiddelde betrouwbaarheid voor kliniek A is iets groter dan die voor B, doch in de uiteindelijke beoordeling ervan kan men veilig stellen dat ze elkaar niet veel ontlopen.

Uit de bovenstaande gegevens kan worden afgeleid dat alle interbeoordelaarbetrouwbaarheden, uitgedrukt in Kappa, geen statistische significantie bereiken, c.q. te laag zijn. Dit geldt voor allebei de klinieken. Het toevoegen van meer informatie, in de zin zoals Giel en Bogaarts (1980) dat omschreven en in dit onderzoek geoperationaliseerd in termen van een volledige gestandaardiseerde ontslagbrief, leidt noch bij de assistent-psychiaters noch bij de psychiaters noch bij de medische codeurs als groep tot een wezenlijke verbetering van ICD-9-CM coderen. Het oorspronkelijke motief om aan assistent-psychiaters het ICD-coderen over te laten, in plaats van aan de medische codeurs, omdat zij op grond van hun kennis van de patiënt beter in staat zouden (moeten) zijn tot coderen, wordt door deze onderzoeksresultaten niet ondersteund. Integendeel, de assistent-psychiaters blijken in dit onderzoek de minst betrouwbare codeurs, de medische codeurs de minst onbetrouwbare. Te stellen dat de medische codeurs het altijd beter doen dan de psychiaters en assistent-psychiaters wordt, door verdere analyse van de afzonderlijke groepen per kliniek, echter ook niet bevestigd.

Navraag bij de medische codeurs in kliniek B leverde echter de bevinding op dat zij gebruik maken van de 'explicatie diagnose-codering' (PIGG, 1978), welke in de loop van de jaren is uitgegroeid naar een lijst van 63 ICD-9-CM codes voor nog meer syndromen. De aldus afgesproken codeerregels hebben in deze kliniek er waarschijnlijk toe bijgedragen dat de medische codeurs betrouwbaarder codeerden dan de andere groepen codeurs (Fokkema, 1980).

De dagelijkse praktijk van het ICD-9-coderen bestaat overwegend uit het omzetten van een geschreven uitboekdiagnose (bijvoorbeeld: paranoïd syndroom overgaand in een depressief syndroom n.a.v. overrijden van haar hond bij een vrouw zonder duidelijke karakterstoornissen) in de daarop betrekking hebbende ICD-codes (aangegeven op de zogenaamde P-kaart van Leidschendam). De betrouwbaarheidsgegevens van onderzoeksdeel A zijn dan ook de beste afspiegeling van de codeurspraktijk. Onbetrouwbaar naar het lijkt.

De ICD-9-CM-coderingsonenigheid nader onderzocht

Met uitzondering van de medische codeurs, gaven de assistent-psychiaters en psychiaters in kliniek A in het B relatief meer hoofd diagnoses dan in het A-onderzoek. Hiervan was echter geen sprake in kliniek B. In kliniek B werden echter, meer dan in kliniek A, meer differentiaal-diagnostische codes gegeven. Dit gegeven bleef bestaan als meer informatie voor coderen werd gegeven (volledige ontslagbrief in plaats van geschreven uitboekdiagnose). Dit betekende dat de diagnostische on-

Tabel 3: Diagnostische overeenkomst tussen 10 beoordelaars over 20 patiënten uit het B-onderzoek

	ICD-9-CM uitboek		Kliniek A		Kliniek B		Diagnose ^d
	code(s) A	code(s) B	procentuele overlap ^b	spreiding ^c	procentuele overlap	spreiding	
1	296	296	7	291, 295 ²	18	292, 295 ²	par. hall. syndroom n.a.v. staken med., chron. psych., eenzaamheid orale + sens. trekken.
2	298, 301	298, 301	43	295, 298 ³	46	293, 298 ⁴	amanticeel syndr. in puerp., karakt. neur., passief afh. trekken.
3	297	297	46	295 ⁴ , 297	30	295 ² , 297 ³	par. hall. syndr. n.a.v. gezinsprobl. bij chron. psychose.
4	297	297	35	295 ² , 297 ³	20	295 ¹ , 296, 297 ³	rec. par. syndr. n.a.v. invol. + gezinsprobl., chron. psychose.
5	296, 317	296, 317	15	294 ³ , 296 ²	30	294, 295, 296 ³	maniform syndr., zwakbegaafdheid, epilepsie.
6	300	300	40	300 ² , 307 ² , 303	24	300, 301, 307 ²	depr. syndr., anorexie n.a.v. losmakingsprobl., karakter neurose.
7	300	300	80	300 ⁵	70	300 ⁸	angst syndr. n.a.v. relatieprobl., ernstige hyst. karakter neurose.
8	301, 309	301, 309	73	300 ⁴ , 296	47	300 ³ , 296 ²	geag. depr. syndr. n.a.v. werk- + relatieprobl., hyst. neurose, psychopath. trekken.
9	297	297	22	297 ² , 295, 297	15	295, 297 ²	rec. par. syndr. n.a.v. t.s. zoom, karakter neurose.
10	300	300	55	298, 300 ⁴	35	300 ³ , 311, 313	depress. syndr. i.k.v. tobisch syndr. n.a.v. adolescentieprobl., hyst. karakt. neur.
11	296	296	36	296 ³ , 298 ²	70	296 ² , 300, 307 ²	agitatie syndr. n.a.v. relatieprobl. met cyclische psychose.
12	301, 309, 317	301, 309, 317	35	296, 300 ² , 309 ²	24	295 ² , 300, 301, 309	gedragsstoorn. i.k.v. geag. depr. syndr. n.a.v. overbelasting, zwakbegaafd. ontw. psychop.
13	290	290	25	290 ² , 295, 300, 308	43	290, 311 ² , 298 ²	stup. tst. i.k.v. depressie syndr., zelfverwaal. n.a.v. opname echgenoot, globale dementie.
14	300	300	43	298, 300 ² , 311	36	300 ³ , 311 ²	depr. syndr. n.a.v. eenzaamheid, karakter neurose, psychasthene trekken.
15	290, 296	290, 296	30	296 ⁴ , 298	36	296 ³ , 298 ²	rec. melan. syndr. n.a.v. lich. klachten, karakter neurose, psychasth. trek., lichte dementie.
16	307, 301, 304	307, 301, 304	58	305, 307 ⁴	67	307 ⁵	anorexia nervosa na verslaving Dulcolax, ernstige karakter neurose.
17	297	297	30	295 ² , 297 ² , 301	40	297 ⁴ , 298	par. hall. syndr. n.a.v. invol. probl., premorbid. paranoïde karakterstr.
18	295, 301	295, 301	46	295 ⁴ , 297	21	295, 297 ² , 298 ²	par. hall. syndr. n.a.v. losmakingsprobl., karakterneurotische ontwikk.
19	296, 301	296, 301	100	300 ⁸	28	300 ⁹ , E 958 ²	t.s. i.k.v. depress. syndr. n.a.v. gezins-, relatie-, seksprobl., ernst. hyst. karakt. neur.
20	298, 301	298, 301	49	293, 295, 298 ³	22	293, 294 ² , 298 ²	amanticeel syndr. n.a.v. geboorte te kind, hyst. karakt. neurose, psychasthene trekken.

N.B.: a Uitboekcode van de patiënt uit de kliniek

b Proportionele overeenkomst van 5 beoordelaars uit de desbetreffende kliniek.

c Eerste hoofddiagnose (code) van elke beoordelaar; grootste aantal en tues zijn gecursiveerd.

d Beschrijvende uitboekdiagnose in de ontslagbrief.

zekerheid (meer informatie betekent meer onzekerheid) van het A-naar het B-deel niet optrad. De verschillen werden, zo leek het, meer door lokale gebruiken bepaald. Zo werden in de B-kliniek, ongeacht het onderzoeksdeel, in totaal meer classificatiecodes (voor hoofd-, neven- en differentiaaldiagnosen) toegekend dan in kliniek A. Ook hebben we elke codeur uit de afzonderlijke groepen met zichzelf vergeleken (onderzoek A versus B). We namen de Kappa als maat voor consistentie. De gemiddelde Kappa voor alle codeurs samen was: $K = .10$ (met een range van $< .0 - 0.20$). Elke codeur bleek weinig consistent te zijn.

Om een inzicht te krijgen in de coderingsovereenkomst van de 10 deelnemende codeurs (uitgesplitst over 2 klinieken) hebben we tabel 3 geconstrueerd. Het gaat hier om de gegevens betreffende het B-deel. Van de hoofddiagnose(n) hebben we (uit ruimte-overwegingen) alleen de eerste code per codeur opgenomen. De getallen boven de codes geven de frequentie van voorkomen aan; de cursieve codes zijn de meest gebruikte codes (de 'ties' werden allebei cursief gezet).

Een globale inspectie van de ICD-uitboekcodes duidt op een redelijk evenwichtige verdeling van patiënten welke met psychotische (13) en neurotische (7) problemen (als één van de hoofddiagnosen) werden opgenomen.

Aan casus 1 geeft de behandelend arts (uitboekdiagnose) een 296.4 code (manisch-depressief, cyclisch). In het licht van de ontslagbrief een misser. De schizofrenie code (295)* komt in de beide klinieken vaker voor en refereert waarschijnlijk aan de 'chronische psychose' in de ontslagbrief. Echter de 2 medische codeurs uit kliniek B coderen 'een paranoïsch-hallucinatoir toestandsbeeld bij opname' als verwijzend naar chronisch paranoïde psychose (297.1), niet gespecificeerde psychose (298, al dan niet reactief) of een psychose die door alcohol (291) of een drug (292) is geïnduceerd. Deze casus laat duidelijk zien dat de medische codeurs vooral af gaan op de reden van opname (het acute beeld), terwijl de psychiaters het acute beeld in een groter beeld plaatsen: zij zien het acute in relatie tot de chronische psychose. In ieder geval draagt dit verschijnsel bij tot het, in beide klinieken, vaak voorkomende verschijnsel dat de codes 295, 297 en 298 vaak samen voorkomen. De casussen 3, 4, 9, 17, 18 laten nog eens duidelijk zien hoe moeilijk het diagnostiseren c.q. coderen van niet-affectieve functionele psychosen met de ICD is.

Casus 19 benadrukt het verschil tussen codeurs en psychiaters: deze patiënte werd opgenomen omdat ze een t.s. had gedaan. De beide medische codeurs uit kliniek B boeken haar dan ook met een t.s.-code (E 958) uit, terwijl de psychiaters het t.s. plaatsen in het kader van een depressief syndroom en haar als lijdend aan een depressie uitboeken (300). Haar behandelend arts vond zelfs dat de depressieve vitale kenmerken had (296).

* Getallen tussen haakjes verwijzen naar de ICD-9-CM-code.

Casus 2 en 20 zijn beide zwangerschapspsychosen. Hiervoor wordt in het merendeel van de gevallen een niet gespecificeerde reactieve psychose (298), een organisch psychotische aandoening (chronisch!), en in één geval een passagère psychose gegeven (zoals elders (Giel en Bogaarts 1980) gepropageerd werd te doen).

Casus 5 was een zeer zwakbegaafde persoon met epilepsie, die een maniform beeld ontwikkelde (296). In de praktijk komt ook wel eens een paranoïde toestand (297) of een opwindingsstoestand (298) daarbij voor. Een dergelijk beeld wordt in beide gevallen ook nog wel eens een 'debiliteitspsychose' genoemd. De ICD heeft hiervoor geen code. Ons leek de ICD-uitboekcode van deze patiënt de meest adequate (296 + 317). Bij de codeurs in beide klinieken zien we dat een dergelijke casus verschillend gezien werd: een aantal codeurs uit kliniek A codeerden in eerste instantie de epilepsie (194); de codeurs uit de tweede kliniek benoemden op de eerste plaats het maniforme beeld (296). Een codeur benoemde de toestand van de patiënt als een psychotische opwindingsstoestand (298). Een codeur ging zelfs zover dit te coderen als schizofrenie (295).

Casus 7 en 10 laten zien dat de codeurs het goed met elkaar eens zijn over het coderen van neurotische depressies, doch het soms oneens zijn (casus 8) met de behandelend arts (uitboekcode) die de persoonlijkheidsstoornis (301, 309) voorop stelt.

Casus 15 stelt het probleem van de hiërarchische ordening van syndromen in het coderen aan de orde. De uitboekcodes lijken een correcte volgorde te hebben: eerst de organische, vervolgens de psychotische en dan de neurotische syndromen. In de onderhavige casus blijkt dat geen enkele codeur uit beide klinieken deze hiërarchie aanhield en het melancholisch syndroom van deze patiënt afwisselend als een niet gespecificeerde affectieve psychose (296) of als een overige niet organische psychose (298) codeerde.

Interessant was het gegeven, tenslotte, dat slechts 1 patiënt in kliniek B uitgeboekt werd als lijdend aan een schizofrene psychose (ICD 295), terwijl alle codeurs in beide klinieken vaker een patiënt als schizofreen codeerden.

We zullen het bij deze korte en zeker niet uitputtende bespreking laten, hoewel we ervan overtuigd zijn dat elke gecodeerde casus (op grond van, nogmaals, volledige ontslagbrieven) open is voor discussie. Duidelijk moge zijn geworden dat veel codeurs, op grond van dezelfde, gestandaardiseerde, grote hoeveelheid uitboekinformatie tot allerlei coderingen komen die onderling nogal van elkaar afwijken.

Samenvatting en discussie

In dit onderzoek werd de betrouwbaarheid van het coderen van structurele uitboekdiagnosen (deel A) en gestandaardiseerde ontslagbrieven (deel B) in twee klinieken onderzocht met behulp van de ICD-9-CM. Over het algemeen kan men stellen dat de interbeoordelaarbetrouw-

baarheid slecht tot zeer slecht was en dat de hoeveelheid gegeven informatie (weinig versus veel informatie over de patiënt) dit gegeven niet veranderde. Hierop uitgezonderd bleken twee medische codeurs uit een kliniek, welke met behulp van de volledige ontslagbrieven wel betrouwbaarder codeerden. (Navraag hierover leerde dat zij in de loop van de jaren een ICD-aanvullend systeem hebben ontwikkeld wat door de andere lokale kliniek niet gehanteerd wordt.) Hierover dient wel te worden opgemerkt dat de medische codeurs in de praktijk eerder met een beschrijvende uitboekdiagnose dan met een volledige ontslagbrief geconfronteerd worden om tot ICD-9-codes te geraken. De resultaten uit dit onderzoek laten zien dat de opvatting dat assistent-psychiaters en psychiaters beter de ICD-codering voor hun rekening kunnen nemen (en dus beter coderen) omdat zij meer informatie over de patiënt hebben, niet opgaat. Toevoeging van informatie leidde juist bij assistent-psychiaters en psychiaters tot een verslechterde betrouwbaarheid, bij de medische codeurs (in twee van de drie gevallen) tot een verbeterde betrouwbaarheid. We noemen een instrument valide als het datgene doet of meet waarvoor het gemaakt is en betrouwbaar als het dat onder alle omstandigheden doet. Validiteit kan derhalve nooit groter zijn dan betrouwbaarheid. Als we daarvan uitgaan, dan moeten we op grond van onze gegevens concluderen dat het besluit om assistent-psychiaters en psychiaters te laten coderen negatieve gevolgen heeft gehad voor zowel de betrouwbaarheid als de validiteit van de ICD-codes. Als uitvloeisel van deze bevindingen kan daaraan de vraag gekoppeld worden in hoeverre waarde gehecht kan worden aan de ICD-codering en dus de epidemiologische gegevens daarover in Nederland, gegeven dat de resultaten uit dit onderzoek een afspiegeling zijn van het ICD-coderen in de diverse Nederlandse klinieken. In dit verband lijkt het ons bijvoorbeeld ook nuttig onderzoek op te zetten naar de betrouwbaarheid van ICD-coderen in andere branches van het ziekenzorgwezen: gegevens daarover ontbreken. Zo merkte de Geneeskundig Hoofdinspecteur voor de Geestelijke Volksgezondheid in zijn brief aan de instellingen die deelnemen aan het Centrale Patiëntenregister Psychiatrische Instituten recentelijk (17 mei 1983) het volgende op: 'Gaandeweg is echter gebleken dat bij een aantal instellingen het coderen van diagnoses is toevertrouwd aan personen die weinig of geen opleiding hebben genoten om dit op een adequate wijze te doen'. Het antwoord op deze vraag lijkt derhalve niet erg hoopvol.

Een summiere analyse van de eerste hoofddiagnose, -classificatie, of code bracht de nodige afwijkingen in coderingsgebruiken aan het licht. Dit kan niet alleen aan onervarenheid van de codeurs (bijvoorbeeld assistent-psychiaters) worden toegeschreven. Zo werd bijvoorbeeld duidelijk dat het coderen van niet affectieve (functionele) psychosen met behulp van de ICD-9 een moeilijke taak is, welke er niet beter op wordt als alle gegevens over de patiënt in de codering betrokken kunnen worden. Alle codeurs werd toegestaan om verschillende hoofd- en

nevendiaagnosen te gebruiken. In de analyse van de gegevens werd niet gelet op de (hiërarchische) volgorde van de codes. Gegeven al deze 'benefits of the doubt' bleef de interbeoordelaarbetrouwbaarheid laag, te laag. Als we, aan het eind gekomen van dit onderzoek naar ICD-9-CM-codering, de voorwaarden die Giel en Bogaarts (1980) noemden voor classificatie en codering overzien dan komen we tot de conclusie dat de geringe betrouwbaarheden die we in dit onderzoek vonden onder andere zijn toe te schrijven aan de beperkte en verwarrende mogelijkheden van opzoeken van ziekte-eenheden en de gebrekkige omschrijving van klassen en indelingscriteria van de ICD-9-CM. We zijn het dan ook eens met de conclusie van Slooff e.a. (1982), waarin zij stellen dat de ICD tekort schiet in een aantal voorwaarden die aan een classificatiesysteem gesteld moeten worden. De ICD-9-CM heeft als classificatiesysteem onvoldoende kwaliteiten en faalt derhalve ook als coderings-systeem in de psychiatrie. De DSM-III heeft als classificatiesysteem betere kwaliteiten (o.a. Slooff e.a., 1982). Het ligt derhalve voor de hand om de DSM-III als classificatiesysteem te gaan gebruiken en voor het coderen daaraan eventueel de ICD te koppelen, voor zover de diagnostische codes niet parallel lopen. In die gevallen waar geen overeenkomst bestaat tussen ICD-9-CM en DSM-III-codes, kan tot een afsprakensysteem of consensus gekomen worden op een landelijk en misschien internationaal niveau. De meer gespecificeerde diagnostische criteria van de DSM-III geven echter nog geen garantie voor een identieke classificatie van dezelfde patiënten door verschillende beoordelaars. Om de steeds complexer wordende classificatiecriteria goed te kunnen doorlopen, zijn toevoegingen in de vorm van symptoomdefinities en een gestandaardiseerd interview aan de DSM-III noodzakelijk. Het classificatieproces kan bovendien zinvol ondersteund worden door inschakeling van een computer en is als zodanig wenselijk.

Op korte termijn kan de DSM-III classificatie enigszins verbeterd worden door deelname aan de trainingen die door de DSM-III werkgroep in Nederland georganiseerd worden. Een aanvulling op die training zou de invoer van een gestructureerd interview kunnen zijn. Dit is in beperkte mate reeds mogelijk met het Diagnostic Interview Schedule (DIS; Robins e.a., 1981). Hierin is de DSM-III als systeem ten dele verwerkt. Een interview dat alle classificatiecriteria van de DSM-III omvat (Standard Clinical Interview for DSM-III; SCID) is in de maak, zo vernamen we op het laatste wereldcongres voor de psychiatrie in Wenen (juli 1983).

Literatuur

- American Psychiatric Association (1980), *Quick reference to the Diagnostic Criteria from DSM-III*. APA, Washington DC.
- Cohen, J. (1968), Weighted Kappa: Nominal scale agreement with provision for scaled disagreement of partial credit. *Psychological Bulletin*, 70, 213-220.

- Dulith, J. en P. M. A. J. Dingemans (1984), Een betrouwbaarheidsonderzoek van ICD-9-CM-codering. *Tijdschrift voor Psychiatrie*, 26, 212-221.
- Fokkema, A. (1980), *Codes voor psychiatrie-registratie Leidschendam*. Interne publikatie. Academisch Medisch Centrum, Amsterdam.
- Giel, R. en F. Bogaarts (1980), Diagnostische classificatie en codering in de psychiatrie. *Tijdschrift voor Psychiatrie*, 22, 5, 327-335.
- Giel, R., A. Busnello, C. E. Climent, A. S. E. D. Elhakim, H. H. Ibrahim, L. Ladrado-Ignacio en N. N. Wig (1981), The classification of Psychiatric Disorder. *Acta Psychiatrica Scandinavica*, 63, 61-74.
- Grove, W. M., N. C. Andreasen, P. McDonald-Scott, M. B. Keller en R. W. Shapiro (1981), Reliability Studies of Psychiatric Diagnoses. *Archives of General Psychiatry*, 38, pp. 408-413.
- Lit, A. C. (1981), Over het classificeren in de psychiatrie. *Tijdschrift voor Psychiatrie*, 23, 1, 5-20.
- Mezzich, J. E., H. C. Kraemer, D. R. L. Worthington en G. A. Goffman (1981), Assessment of Agreement among Several Raters formulating Multiple Diagnoses. *Journal of Psychiatric Research*, 16, 29-39.
- Robins, L. N., J. E. Helzer, J. Croughan en K. S. Ratcliff (1981), National Institute of Mental Health Diagnostic Interview Schedule. *Archives of General Psychiatry*, 381-389.
- Silbermann, R. M. (1971), *CHAM; een systeem voor de classificatie van psychiatische toestanden*. Excerpta Medica, Amsterdam.
- Slooff, C. J., R. Giel, A. de Jong en D. Wiersma (1982), De classificatie van de functionele niet-affectieve psychosen. *Tijdschrift voor Psychiatrie*, 24, 7/8, 454-469.
- Weerden-Dijkstra, J. R. van, R. Giel (1974), Mededelingen uit het Centrale Patiëntenregister voor Psychiatische Ziekenhuizen. *Tijdschrift voor Psychiatrie*, 16, 453-461.
- World Health Organization (1978), *Mental Disorders: Glossary and guide to their classification in accordance with the Ninth Revision of the International Classification of Diseases*. World Health Organization, Geneva. (Uitgave: Nijhof, Den Haag.)