

## De 'diagnostic interview schedule' (DIS)

door P. Dingemans, H. van Engeland,  
J.H. Dijkhuis en J. Bleeker

### Inleiding

In de psychiatrie wordt het begrip 'diagnostiek' in tenminste twee betekenissen gebruikt. De eerste, en oudere, is van nosologische aard en omvat een theorie over de zieke of ziekte. Het zogenaamde nosologische denken vond zijn hoogtepunt in Kraepelin, in Nederland in Rümke. De tweede is een taxonomische en houdt een beschrijving of ordening van gegevens van een klasse in. Dit heet classificeren. Vertegenwoordigers van deze betekenis van het begrip diagnostiek zijn de zogeheten neo-Kraepelianen zoals Spitzer. In het vervolg zullen we het woord diagnostiek gebruiken in de betekenis van classificeren en classificatie.

Het eerste belangrijke onderzoek naar interbeoordelaar betrouwbaarheid werd uitgevoerd door Ash (1949). Hij concludeerde dat beoordelaars een te lage ('poor') interbeoordelaar-betrouwbaarheid hadden en dat hun overeenstemming beter was bij hoofdklassen dan specifieke diagnoses. Sinds die tijd zijn er vele betrouwbaarheidsstudies uitgevoerd; niet onbegrijpelijk als men zich realiseert dat de bruikbaarheid van een classificatiesysteem (mede) bepaald wordt door intersubjectieve overeenstemming (Van Engeland, 1980). De feitelijke ontwikkeling van betrouwbare classificatiesystemen, met name in de Verenigde Staten, kwam aan het begin van de jaren zeventig pas goed op gang. Dit

---

De schrijvers zijn verbonden, respectievelijk als wetenschappelijk medewerker aan de Rijks Universiteit Utrecht en de Universiteit van Amsterdam, als hoogleraar in de klinische psychopathologie aan de Rijks Universiteit Utrecht, als hoogleraar in de klinische psychologie aan de Rijks Universiteit Utrecht en als psychiater aan de Psychiatrisch Centrum AZUA in Amsterdam. Deelname aan de DIS-training van de eerste auteur werd mogelijk gemaakt dank zij een stipendium van de Nederlandse Organisatie voor Zuiver Wetenschappelijk Onderzoek. Correspondentieadres: Psychiatrisch Centrum AZUA, Tafelbergweg 25, 1105 BC Amsterdam.

is waarschijnlijk mede beïnvloed door het werk van Kendell e.a. (1968, 1970, 1971), waarin werd gewezen op transcultureel verschillende interpretaties van diagnostische termen door psychiaters en dergelijke.

Na een omvattende bespreking van de betrouwbaarheidsliteratuur kwamen Spitzer en Fleiss (1974) tot de conclusie dat:

1. betrouwbaarheid een noodzakelijke voorwaarde is voor een adequaat classificatiesysteem;
2. dat het onderzoek voor 1975 had aangetoond dat de betrouwbaarheid van classificatiesystemen onvoldoende was, en
3. dat de reden van de onbetrouwbaarheid de vaagheid van DSM-I en DSM-2 was.

Tevens stelden zij in dit artikel voor de mate van betrouwbaarheid niet langer in een maat voor procentuele overeenkomst, doch in een Kappamaat te berekenen. Deze laatste maat begon in de jaren zeventig de eerste in toenemende mate te vervangen, hoewel op de Kappamaat voor betrouwbaarheid (Maxwell, 1977) ook de nodige kritiek is gekomen. In de jaren zeventig is echter in de ontwikkeling van nieuwe classificatiesystemen een (te) grote nadruk komen te liggen op de betrouwbaarheid. Carey en Gottesman (1978) hebben op overtuigende wijze aannemelijk gemaakt dat het ontwikkelen van een zeer betrouwbaar classificatiesysteem ten koste *kan* gaan van de validiteit ervan.

Het gebruik van classificatiesystemen in de psychiatrie loopt nogal uiteen. In Nederland kennen we het CHAM-systeem van Silbermann (1971) en de International Classification of Diseases in zijn negende, al of niet met klinische modificatie, versie (WHO, 1978). Verder mag de Present State Examination (Wing e.a., 1974) zich in een grote populariteit in Engeland en in een wat mindere mate, en dan vooral bij onderzoekers, in de rest van Europa verheugen. Sinds kort probeert de 'Association for Methodology and Documentation in Psychiatry (AMDP; Bobon e.a. Eds., 1983) met het zogeheten AMDP-systeem het classificatiefront in Europa open te breken. De momenteel meest gebruikte is echter de Diagnostic Statistical Manual (APA, 1980), welke thans zijn derde versie kent en binnen niet al te lange tijd wordt opgevolgd door de vierde. Er is geen systeem in de psychiatrie dat zo snel en wijd verbreid ingang heeft gevonden als de DSM-III (Spitzer e.a. Eds., 1983). Onlangs betoogden wij (Dingemans e.a., 1984) met anderen (Slooff e.a., 1982) dat de invoering van de DSM-III in de Nederlandse psychiatrie een redelijk alternatief zou zijn voor de thans gebruikte ICD-9-CM, op grond van het argument dat deze laatste niet voldoet aan een aantal basale eisen die aan een classificatiesysteem gesteld moeten worden. Wij betoogden daarbij tevens dat aan de DSM-III een gestandaardiseerd interview verbonden diende te zijn en noemden toen de Diagnostic Interview Schedule (DIS; Robins e.a., 1981) en de Structured Clinical Interview for DSM-III (SCID); welke laatste in de maak is. Een gestandaardiseerd interview is immers een belangrijke mogelijkheid om een

belangrijk deel van de diagnostische onbetrouwbaarheid op te heffen (Ward e.a., 1962).

Wij noemden met name de DIS omdat dat een gestandaardiseerd interview is dat door leken kan worden afgenomen en als zodanig in epidemiologisch onderzoek gebruikt kan worden en daardoor ook sterk kostenbesparend kan werken. In een tijd waarin de beschikbare onderzoeksgelden verminderen en de vraag naar epidemiologisch onderzoek toeneemt, leek ons dit instrument derhalve een verdere studie waard. Omdat de eerste auteur in de gelegenheid was zich in het gebruik van de DIS te trainen, willen wij hier wat verder ingaan op de achtergronden van het ontstaan, de structuur en de toepassingsmogelijkheden ervan. Daarbij maken we onder andere gebruik van datgene wat er door de ontwerpers zelf over is geschreven (Helzer e.a., 1981; Robins e.a., 1981, 1982) en wat er door de trainers over is gezegd. We zullen dit alles kritisch evalueren met betrekking tot de toepasbaarheid in het Nederlandse taalgebied.

### Enkele achtergrondgegevens van de DIS

De ontwikkeling van de DIS vloeit voort uit het streven naar: 1. betere, meer betrouwbare, rubricerende diagnostiek van psychopathologische toestandsbeelden; 2. eenheid in diagnostische gebruiken; en 3. kostenbesparingen op de gezondheidszorg. De weg naar betere en meer betrouwbare klinisch-psychiatrische diagnostiek in de Verenigde Staten liep via het 'Manhattan Project' (Srole e.a., 1962), de 'St. Louis-500 study' (Feighner e.a., 1972), de 'Research Diagnostic Criteria' (Spitzer e.a., 1978), de 'Schedule of Affective Disorders and Schizophrenia' (Spitzer en Endicott, 1978), de 'Diagnostical Statistical Manual-III' (APA 1980), naar het 'Renard Diagnostic Interview' (Helzer e.a., 1981) en eindigde voorlopig bij de 'Diagnostic Interview Schedule' (Robins e.a., 1981). Deze ontwikkelingsgang is waarschijnlijk in belangrijke mate geïnspireerd door het werk van Kendell (Kendell, 1968, 1970, 1971), waarin bij voortduring wordt gewezen op verschillen in internationale en lokale diagnostische gebruiken. Zij is in ieder geval in sterke mate bepaald door het werk van Spitzer en Robins in hun actieve bijdragen aan de diverse systemen (o.a. de DIS), door Spitzers voorstudies voor een verbeterde versie van de DSM-II (Spitzer en Klein Eds., 1978) in de vorm van de DSM-III en zijn niet aflatende zorg ook dit systeem te perfectioneren door het creëren van een breed forum over het gebruik van de DSM-III (Spitzer e.a., 1983) en door de ontwikkeling van een daarop geënt standaard klinisch interview, de Structured Clinical Interview for DSM-III (SCID; in voorbereiding). De genoemde systemen (RDC, SADS, DIS, DSM-III) hebben een *gemiddelde* betrouwbaarheid, uitgedrukt in Kappa, van .70 (met een verschillende range over de verschillende classificatiecategorieën). De classificatiebetrouwbaarheid is voor de hoofdrubrieken, met name de functionele

psychosen) beter dan voor de subcategorieën en de neurosen. Voorts worden de systemen gekenmerkt door meer specifieke omschrijving en uitbreiding van het aantal diagnostische categorieën (de DSM-I en -II omvatten 34.000, de DSM-III 177.000 woorden; Blashfield, 1984); dooreen loslaten van de nosologische denkwijze (toestand + oorzaak + beloop = ziekte (nosos)) en door een zich concentreren op descriptieve categorieën: een taxonomische benadering.

De totstandkoming van de DIS werd in belangrijke mate geïnspireerd door Renard Diagnostic Interview (RDI; Helzer e.a., 1981). Dit klinische interview voor psychiaters, waarin de Feighner-criteria werden geoperationaliseerd, toonde tevens aan dat leken-interviewers even betrouwbaar classificeerden als psychiaters. De totstandkoming van de DSM-III en het streven van het National Institute for Mental Health om tot eenvormigheid in classificatieprocedures te komen maakten de weg vrij voor de ontwikkeling van de DIS, waarin de classificatiecriteria van verschillende systemen verwerkt zijn en de vragen gesteld worden door een leek. De DIS is, enig in zijn soort, een interview dat, met training, door leken kan worden afgenomen en als zodanig kostenreducerend kan werken, met name in epidemiologisch onderzoek, bij grote groepen mensen. Als zodanig lijkt de DIS voor dit doel ook beter bruikbaar dan de Present State Examination (PSE; Wing e.a., 1974) omdat deze laatste door klinisch ervaren mensen moet worden afgenomen.

### **De structuur van de standaardvragen**

De DIS zelf is een sterk gestructureerd diagnostisch interview, dat is opgebouwd rondom een drietal psychiatrische classificatiesystemen.

Dit zijn de 'Feighner-' (Feighner e.a., 1972), de 'Research Diagnostic Criteria' (Spitzer e.a., 1978) en de 'Diagnostic Statistical Manual' (DSM-III; APA, 1980). Op grond van de resultaten van het interview kunnen de psychopathologische symptomen van de geïnterviewde (in principe) in een van deze drie of alle drie genoemde systemen, al of niet met een computer, gesorteerd worden en in klassen of categorieën worden ondergebracht. Met zijn 263 gestandaardiseerde vragen is de DIS op dit moment het meest uitgebreide psychiatrische interview. De geclaimde tijdsduur voor afname, al naar gelang de aanwezigheid en intensiteit van de symptomatologie, is 45 tot 75 minuten. Onze ervaring met de PSE heeft geleerd dat bij een zorgvuldige afname van een standaardinterview de geclaimde tijd met een factor 2 moet worden vermenigvuldigd.

De vragen (259 met betrekking tot symptomatologie en 4 observatievragen over psychotische verschijnselen) zijn gerangschikt in de volgende categorieën: 1 - 4 demografische gegevens (vooral bedoeld om een indruk te krijgen over de interviewbaarheid van de patiënt; zie ook vragen 240-258); 16-61 somatiseren; 62-67 paniekaanvallen; 68-71

fobieën; 72-99 depressie; 100-117 manie; 118-143 schizofrenie; 144-148 anorexie; 149-170 alcoholmisbruik, -verslaving; 172-180 obsessief compulsief gedrag; 181-195 drugmisbruik/-verslaving; 196-209, S23-234 antisociaal gedrag; 210, S 23 seksualiteit; 214-218 relationeel functioneren; 219-227 werk; 228-234 verzorging van kinderen; 235-239 gokken; 240-258 vragen met betrekking tot hersenorganische syndromen; 260-263 observaties met betrekking tot psychosen.

Het is de bedoeling dat de vragen in de aangegeven volgorde worden afgewerkt. Er is een uitzondering op. In het geval dat er twijfel bestaat over de interviewbaarheid van de patiënt als gevolg van organisch-cerebrale problemen (deze indruk moet in de eerste 30 vragen ontstaan zijn), wordt verder gegaan met de 'mini-mental'. Mocht deze een score van 12 of meer opleveren (vraag 240-258) dan wordt het interview beëindigd. Dit deel van het interview is in feite de 'Mini Mental Status Examination' uit het John Hopkins Ziekenhuis (USA).

De taal van het interview is zo simpel gehouden dat iedereen de vragen moet kunnen begrijpen. Een voorbeeld:

Tabel 1: Voorbeeld van vragen en coderen daarvan in de DIS.

		Deck 01				
27	Heeft u ooit veel problemen met OVERGEVEN (niet tijdens de zwangerschap) gehad?					
	ARTS _____ ZELF _____	1	2	3	4	5 66
28	VRAAG ALLEEN AAN VROUW MET KIND (. 14): Heeft u ooit tijdens de hele zwangerschap moeten OVERGEVEN? Ja ..... (CODEER 5 EN GA NAAR V. 29) Nee ..... (VRAAG A)  A Bent u tijdens een zwangerschap ooit in een ziekenhuis opgenomen omdat u moest overgeven? Ja ..... (CODEER 1) Nee ..... (CODEER 5)					
				1	5	67
29	Heeft u ooit veel last gehad van MISSELIJKHEID – een draaiurig gevoel in uw maag – terwijl u niet echt hoefde over te geven? ALS WAGEN- OF ZEEZIEKTE WORDT GENOEMD, CODEER DAN ALLEEN 4.					
	ARTS _____ ZELF _____	1	2	3	4	5 68
30	Heeft u ooit veel last van BUIKLOOP of diarree gehad?					
	ARTS _____ ZELF _____	1	2	3	4	5 69
		Deck 02				
31	Heeft u ooit veel last van GASVORMING of een opgeblazen gevoel in uw maag of darmen gehad?					
		1	2	3	4	5 08

32 Zijn er bepaalde DINGEN DIE U NIET KUNT ETEN omdat u ontdekt heeft dat ze u ziek maken?

Nee ..... (CODEER 1)

Ja ..... (VRAAG A)

A Kunt u enkele dingen noemen die u niet kunt eten omdat ze u ziek maken?

INTERVIEWER: ALS HET ER MINDER DAN 3 ZIJN, CODEER DAN 1; ALS HET ER DRIE OF MEER ZIJN BEGIN DAN MET NAVRAGEN.

1 4 5 09

ARTS \_\_\_\_\_ ZELF \_\_\_\_\_

De vragen worden begeleid door instructies, welke in hoofdletters zijn geschreven en vanzelfsprekend niet aan de patiënt of de geïnterviewde worden voorgelezen. Voorbeeld:

59 IK GA U EEN AANTAL PROBLEMEN NOEMEN WAAR U HET OVER HAD. IK ZOU GRAAG WILLEN WETEN WANNEER U VOOR HET EERST LAST HAD VAN ÉÉN VAN DEZE. BIJVOORBEELD: U HEEFT LAST GEHAD VAN (NOEM HIER ALLE GECODEERDE VRAGEN VAN V. 16-58). WAT WAS DE VROEGSTE LEEFTIJD DAT U VAN ÉÉN VAN DEZE LAST HAD? (CODEER LEEFTIJD).

Specifieke opdrachten aan de interviewer zijn over het algemeen opgenomen in een gesloten blok en zijn ook in hoofdletters geschreven:

INTERVIEWER: ALS DE GEÏNTERVIEWDE HET NIET WEEET EN HIJ IS JONGER DAN 30, CODEER DAN 01. ALS DE GEÏNTERVIEWDE HET NIET WEEET EN HIJ IS OUDER DAN 30, STEL DAN VRAAG A.

De scoremogelijkheden van de vraag en het antwoord van de geïnterviewde maken dat op een specifieke wijze met het 'navragen' verder wordt gegaan; dit wordt in eerste instantie door de codes bij de standaardvraag aangegeven en vervolgens verder uitgewerkt door middel van het stroomdiagram voor de navragen of vragen om toelichting, welke ook gestandaardiseerd zijn.

Het type vraag met de bijbehorende code en de daaraan gekoppelde vragen om toelichting vormen het hart van het interview, zijn (naar onze ervaring) moeilijk onder de knie te krijgen en vormen daarom het grootste obstakel voor de uitvoering van het interview. Zij zullen er daarom hier wat verder op ingaan.

Voorbeelden van het type vragen:

a. Heeft u het ooit met een arts over uw problemen met roken gehad?

Nee ..... 1

Ja ..... 5

b. Heeft u ooit *problemen met lopen gehad*? 1 2 3 4 5

c. Bent u ooit *verlamd* geweest? Dat betekent dat u tenminste een paar minuten helemaal niet in staat was een

- deel van uw lichaam te bewegen. 1 3 4 5  
 d. Bent u ooit uw *stem* 30 minuten of langer *kwijt* geweest,  
 zodat u alleen nog kon fluisteren? 1 4 5  
 e. Heeft u ooit *pijn* op uw *borst* gehad? 1 2 4 5

Voor de codes geldt dat een 1 altijd een 'nee'- en een 2, 3, 4 of 5-code een 'ja'-antwoord inhoudt. Een 2-code betekent dat het symptoom waar naar gevraagd werd niet ernstig genoeg was om verder nog in aanmerking te komen (bereikte geen drempelwaarde). Een 3-code betekent dat het symptoom de drempelwaarde bereikte, serieus aanwezig was of is, doch altijd veroorzaakt wordt of werd door gebruik van alcohol, drugs of medicijnen. Een 4-code betekent dat het symptoom altijd serieus aanwezig was of is en altijd verklaard kon of kan worden door zowel letsel of lichamelijke ziekte als wel het gevolg was van alcohol-, drug- of medicijngebruik. Een 5-code betekent dat het altijd ernstig genoeg was en niet altijd door zowel letsel of lichamelijke ziekte als wel drug-, medicijnen- of alcoholgebruik verklaard kon worden. Uit de voorbeeldvragen wordt duidelijk dat niet op alle vragen de 5-codes van toepassing zijn. Zo zal voor bepaalde symptomen zelden een lichamelijke oorzaak aanwezig zijn. In dat geval wordt de code-4 weggelaten. Soms zijn symptomen zo ernstig dat de 2-code zonder meer kan worden weggelaten (voorbeeld c en d). Elke vraag heeft echter op zijn minst een 1- en een 5-code (voorbeeld a). In een dergelijk geval behoeven geen vragen om toelichting gesteld te worden, hetgeen ook geldt voor alle codes die onder elkaar (verticaal) in het interview voorkomen:

11 Hoe lang is het geleden dat u een dergelijke aanval had?

CODEER DE	In de laatste twee weken .....	1
MEEST RECENTE	In de laatste maand .....	2
PERIODE	In de laatste 6 maanden .....	3
	In het laatste jaar .....	4
	Langer dan 1 jaar geleden .....	5

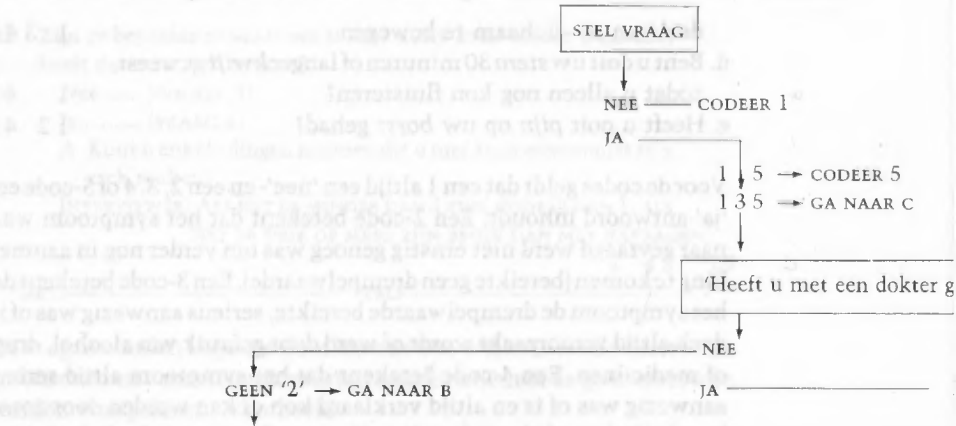
Om nu te bepalen hoe een 'ja'-antwoord verder dient te worden gecoördineerd is een stroomdiagram ontwikkeld. Dit wordt bij de standaardvragen door een horizontale weergave van de codes aangeduid, voorbeeld b t/m e).

Het stroomdiagram ziet er als volgt uit: zie tabel 2 op de volgende bladzij.

Het stroomdiagram dient als volgt gebruikt te worden. Eerst wordt de vraag aan de geïnterviewde gesteld, bijvoorbeeld:

- 30 Heeft u wel eens veel last gehad van buikloop of diarree?  
 ARTS ————— ZELF ————— 1 2 3 4 5

Tabel 2: Stroomdiagram voor het stellen van vragen in de Diagnostic Interview Schedule.



**A**

'Heeft u het met een andere hulpverlener over (S<sub>x</sub>) gehad?'  
 'Heeft u voor (S<sub>x</sub>) meer dan eens medicijnen ingenomen?'  
 'Leverde (S<sub>x</sub>) veel problemen in uw leven of uw dagelijkse bezigheden op?'

ALLEMAAL NEE → CODEER 2  
 EERSTE JA

---

**B**

GEEN '4' → GA NAAR BLOK C

'Was (S<sub>x</sub>) ooit het gevolg van een lichamelijke ziekte of letsel?'

JA → GA NAAR ▲ IN BLOK D  
 NEE

---

**C**

GEEN '3' → CODEER 5

'Was (S<sub>x</sub>) ooit het gevolg van medicatie-, drugs- of alcoholgebruik?'

NEE → CODEER 5  
 JA

'Was (S<sub>x</sub>) altijd het gevolg van medicatie-, drugs- of alcoholgebruik?'

NEE → CODEER 5  
 JA → ('Wat veroorzaakte (S<sub>x</sub>)?'  
 NOTEER ALC/MED. EN CODEER 3

**D**

'To...

- Zen...
- Stre...
- Bez...
- Dep...
- Psy...

NOTE...

CODE...

---

'Als...

dru...

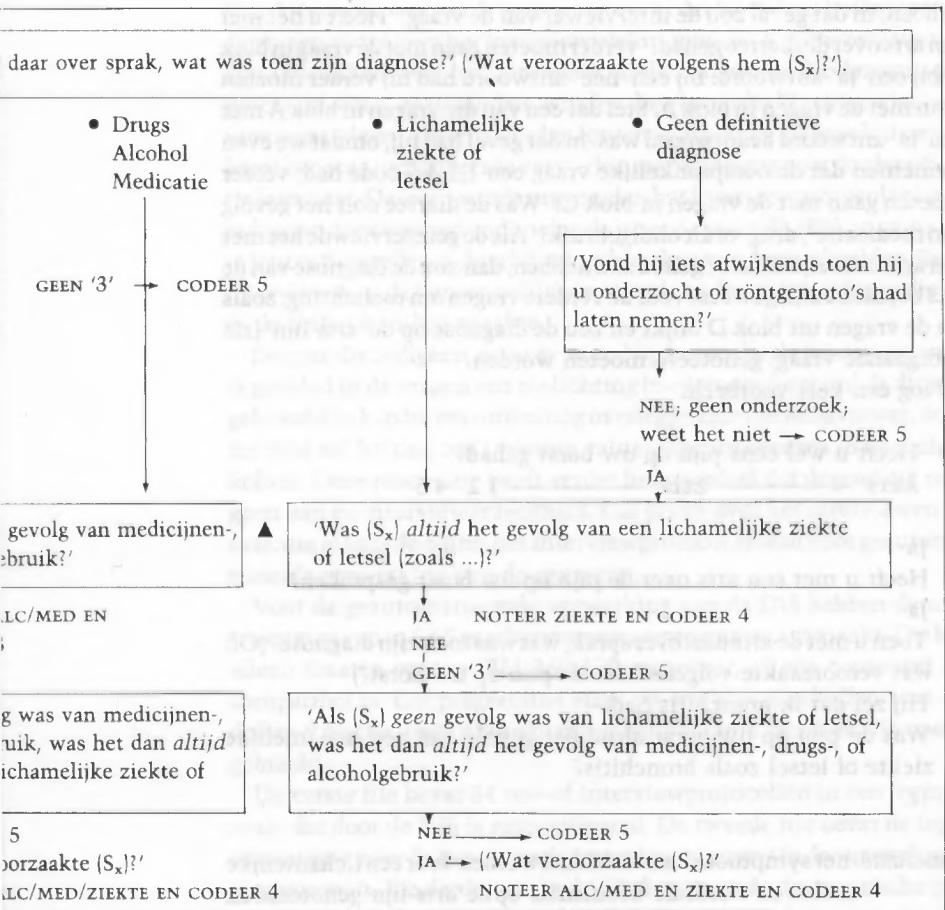
het...

lets...

Andere hulpverleners houden in: sociale werkers, verpleegkundigen, geestelijken, psychologen, tandartsen, chiropracters en voetspecialisten.



Deze flowchart is bedoeld om de diagnose van de ziekte van de lever te stellen. Het begint met de vraag: 'Wat veroorzaakte volgens hem (S<sub>x</sub>)?' De gebruiker wordt gevraagd te kiezen uit drie opties: Drugs Alcohol Medicatie, Lichamelijke ziekte of letsel, of Geen definitieve diagnose. De flowchart leidt de gebruiker door een reeks vervolgvragen, zoals 'Vond hij iets afwijkends toen hij u onderzocht of röntgenfoto's had laten nemen?' en 'Was (S<sub>x</sub>) altijd het gevolg van een lichamelijke ziekte of letsel (zoals ...)?' De gebruiker wordt uiteindelijk geadviseerd welke codeer te gebruiken (CODEER 4 of CODEER 5) en welke informatie te noteren (ALC/MED EN ZIEKTE).



Het nee-antwoord op deze vraag wordt direct gescoord met een 1 (niet aanwezig). Stel, de geïnterviewde geeft een ja-antwoord. In dat geval (zie stroomdiagram) moeten we hem de volgende vraag stellen ('Heeft u met een arts over (S<sub>x</sub>) gesproken?'), omdat de vraag geen (1 5) of een (1 3 5)formaat heeft maar een (1 2 3 4 5)formaat. Als het antwoord op deze vraag om toelichting met 'ja' beantwoord wordt, dan moet verder gegaan worden met de volgende vraag om toelichting: 'Toen u met de arts daarover sprak, wat was toen zijn diagnose?' (blok D). Was het antwoord op de vraag of hij er met een arts over gesproken had negatief, dan had de interviewer verder moeten gaan met de vragen uit blok A, totdat een van de drie vragen met een 'ja' beantwoord was, om dan vervolgens weer verder te gaan met de vragen in blok B en dergelijke. Stel dat de vraag over diarree of buikloop een 1,2,3,5-code gehad zou hebben. In dat geval zou de interviewer van de vraag: 'Heeft u het met een arts over de diarree gehad?' verder moeten gaan met de vraag in blok D bij een 'ja'-antwoord. Bij een 'nee'-antwoord had hij verder moeten gaan met de vragen in blok A. Stel dat een van die vragen in blok A met een 'ja'-antwoord beantwoord was. In dat geval had hij, omdat we even aannemen dat de oorspronkelijke vraag een 1,2,3,5-code had, verder moeten gaan met de vragen in blok C: 'Was de diarree ooit het gevolg van medicatie-, drug- of alcoholgebruik?' Als de geïnterviewde het met een arts over zijn diarree gehad zou hebben, dan zou de diagnose van de arts bepalend zijn geweest voor de verdere vragen om toelichting, zoals uit de vragen uit blok D blijkt en zou de diagnose op de 'arts-lijn' (zie voorgaande vraag) genoteerd moeten worden.

Nog een kort voorbeeld.

20 Heeft u wel eens pijn op uw borst gehad?

ARTS \_\_\_\_\_ ZELF \_\_\_\_\_ 1 2 4 5

A Ja

V Heeft u met een arts over de pijn op uw borst gesproken?

A Ja

V Toen u met de arts daarover sprak, wat was toen zijn diagnose? (Of: wat veroorzaakte volgens hem de pijn op uw borst?)

A Hij zei dat ik bronchitis had.

V Was de pijn op uw borst altijd het gevolg van een lichamelijke ziekte of letsel zoals bronchitis?

A Ja.

*Conclusie:* het symptoom kan verklaard worden door een lichamelijke ziekte en wordt als 'bronchitis' op de 'arts-lijn' genoteerd en gecodeerd met een '4'.

Dit deel van het interview, nogmaals, leverde in de training (1 week full-time) problemen op voor iedereen. Een goed kunnen hanteren van

het stroomdiagram is echter van vitaal belang voor het goed coderen van de antwoorden van de geïnterviewde, vooral omdat de geautomatiseerde verwerking van de DIS daar met een ijzeren logica op aansluit en elke inconsistentie in het codeerresultaat weigert te verwerken. Een uitgebreide handleiding voor het hanteren van specifieke antwoorden en hoe daarmee verder te gaan, is voor de interviewer met de DIS ontwikkeld.

In de training van het gebruik van de DIS zijn per symptoomsectie ook nog weer specifieke situaties of antwoorden verwerkt. Deze zijn ook beschikbaar via een trainershandleiding.

### **Redigeren, coderen en verwerken**

Van de interviewer wordt tevens verwacht dat hij tijdens het interview de vragen codeert en het interview afsluit met een korte weergave van zijn indrukken over het interview en de patiënt. In dit dynamische proces moeten antwoorden worden her-beoordeeld, uitgewist, opnieuw gecodeerd. Daarin worden fouten gemaakt en als gevolg daarvan loopt het geautomatiseerde verwerkingsprogramma voor de classificatie later vast. Daarom wordt aangeraden het interviewprotocol te laten redigeren door iemand anders dan de interviewer zelf. Hij kan gemaakte fouten herstellen en hierbij gebruik maken van de voorbeelden die de interviewer op de 'arts en zelflijn' noteerde en de aantekeningen die hij in de linker kantlijn maakte.

Degene die redigeert moet er vooral op letten dat de juiste sequentie is gevolgd in de vragen om toelichting (zie stroomdiagram). Is dit niet gebeurd dan kan hij eventueel nog overleggen met de interviewer, doch meestal zal hij dan een 'missing value' moeten noteren in de rechter kolom. Deze procedure heeft verder het voordeel dat degene die redigeert aan de interviewer feedback kan geven over het interviewen. In essentie maakt de editor het interviewprotocol gereed voor geautomatiseerde verwerking van de gegevens.

Voor de geautomatiseerde verwerking van de DIS hebben de constructeurs een in SAS geschreven sorteerprogramma gemaakt. Dit kan alleen draaien op een IBM 360-370 computer, of een computer die compatibel is. Dit programma staat op een aan te schaffen tape (50 dollar; 9 tracks, 1600 BPI en niet gelabeld) waarop vier files zijn ondergebracht.

De eerste file bevat 34 test-of interviewprotocollen in een formaat zoals dat door de DIS is gespecificeerd. De tweede file bevat de input statements voor de input van de interviewgegevens in de verwerkingsprogramma's. De derde en vierde file bevatten de diagnostische programma's van de DSM-III en de RDC. Deze twee diagnostische programma's moeten apart gedraaid worden. Bij de tape wordt verder de benodigde documentatie geleverd.

Momenteel (1984) worden verschillende pogingen ondernomen in

de VS om de afname en de verwerking (beide of afzonderlijk) van de DIS te automatiseren. Het probleem daarbij is vooral dat zowel de afname als wel de verwerking van de DIS behoorlijk meer geheugenruimte van een computer vraagt dan de meeste micro's standaard hebben. Nodig is op zijn minst 365 kB (meestal aanwezig 64-128 kB). Het zal echter niet lang meer duren voordat de geautomatiseerde verwerkingsprogramma's, al of niet op commerciële basis, op de markt zullen komen. Verder bestaat er de mogelijkheid de DIS met de hand te verwerken. Dit is een procedure die zeer tijdrovend is (een halve dag), aanleiding geeft tot het maken van nogal wat fouten en daarom wordt afgeraden.

Met behulp van de twee laatstgenoemde programma's kunnen de volgende classificaties in het DSM-III-systeem worden gemaakt: seniele en preseniele dementieën, schizofrenieën, depressies in engere zin, bipolaire stoornissen, enkele fobieën, obsessief-compulsieve stoornis, somatiseringsstoornissen, antisociale persoonlijkheidsstoornissen, psychoseksuele stoornissen, misbruik of verslaving aan alcohol of diverse drugs – anorexie en pathologisch gokken (Robins e.a., 1981, p. 382). Het programma laat (nog) geen DSM-III-classificatie toe van stoornissen bij kinderen. Hiervoor is een interview naar analogie van de DIS ontwikkeld dat bij de ouders wordt afgenomen.

Voor de RDC-classificaties is het aantal mogelijkheden beperkter: schizofrenieën, depressies, verschillende fobieën, obsessief-compulsieve stoornissen, Briquets syndroom, antisociale persoonlijkheid en alcoholisme.

### De betrouwbaarheid en de validiteit van de DIS

De beschikbare betrouwbaarheidscijfers van de DIS hebben betrekking op de tweede versie (Robins e.a., 1981). In dit onderzoek werden 216 mensen geïnterviewd: eenmaal door een lekeninterviewer en eenmaal door de psychiater. Beide malen werd de DIS gebruikt. In aanvulling daarop werd de psychiater gevraagd om in een zogenaamde 'vrije periode' na het interview aan de patiënt alle vragen te stellen die hij maar wilde om zijn klinische impressies die hij bij de afname van de DIS had opgedaan verder te onderzoeken. Na het vrije deel werd hem gevraagd om zijn klinische indrukken opnieuw bij te stellen. De psychiater werd, met andere woorden gevraagd navraag te plegen over die antwoorden op de DIS-vragen, waar hij onzeker over was gebleven, of die tegen zijn klinisch oordeel ingingen.

Het patiëntenbestand bestond uit 118 psychiatrische patiënten, 39 ambulante psychiatrische patiënten, 24 verzorgers van de universiteit zonder bekende psychopathologie, 10 mensen van de 'Gamblers (Anonymous, en 26 ex-patiënten. De volgorde van interviews door de lekeninterviewer versus de psychiater werd gerandomiseerd.

Bij de beëindiging van het onderzoek bleken 97 lekeninterviewers het eerste interview te hebben afgenomen en in 119 gevallen bleken de

psychiaters de eerste interviewers te zijn geweest. De interviews werden op de gebruikelijke wijze geredigeerd (zie bovenstaande) en vervolgens verwerkt. Vier maten werden gebruikt voor verdere studie:

1. een maat voor afwijking in het lekeninterview;
2.  $K(\kappa)$  als een concordantiemaat tussen de leek en de psychiater;
3. een sensitiviteitsmaat (percentage van correct geïdentificeerde gevallen);
4. een specificiteitsmaat (percentage van correct geïdentificeerde niet-gevallen);
5. een proportionele maat voor het gegeven dat de lekeninterviewer een casus als positief beoordeelt, terwijl het geen 'geval' is (false positives).

Het vrije deel van het interview voor de psychiaters was bedoeld om de klinische indrukken opgedaan tijdens de afname van de DIS te vergelijken met de klinische indrukken die werden verkregen over symptoomgebieden die mogelijk niet in het interview aan bod waren gekomen of onduidelijk waren gebleven. Hiermede zouden, zo redeneerden de makers van de DIS, aanwijzingen verkregen kunnen worden over welke delen van het interview nog niet goed ontwikkeld zouden zijn. Post hoc analyses brachten echter aan het licht dat psychiaters zelden twijfel hadden over de diagnose die de DIS gesteld had (0-7% per classificatiecategorie).

Uit de afwijkingsmaat kwam naar voren, dat lekeninterviewers alcoholverslaving, somatisering en paniekaanvallen significant onderdiagnostiseerden in de DSM-III en in de RDC-criteria, en drugmisbruik in de DSM-III en drugverslaving in de Feighnercriteria overdiagnostiseerden. Somatisering bleek het sterkst te worden ondergediagnostiseerd.

De gemiddelde waarden voor betrouwbaarheid ( $K$ ), specificiteit en sensitiviteit voor de DSM-III criteria waren respectievelijk: .69, 75 en 94 procent. De resultaten voor de Feighner criteria waren vergelijkbaar en die voor de RDC minder vergelijkbaar.

Dit laatste werd door de auteurs vooral toegeschreven aan het gegeven dat met de RDC classificaties waren gemaakt in die diagnostische groepen waarvoor een lage concordantie bestond. Als classificaties werden vergeleken die zowel met de RDC alswel met een ander systeem gemaakt waren, waren de uitkomsten wederom vergelijkbaar.

Met de validiteit hebben de auteurs begrijpelijkerwijze moeite, want zoals in veel psychiatrische onderzoeken hebben zij het psychiateroordeel als criterium genomen, waarmee ze het lekenoordeel vergeleken hebben. De uitkomst daarvan bespraken ze in termen van validiteit. Dat lijkt ons onjuist, evenals de benaming van 'procedural validity', zoals die door Heltzer e.a. (1981) wordt gebruikt: ook het artsenoordeel hoeft niet absoluut juist te zijn (Dingemans e.a., 1984). Het probleem van validiteit wordt niet opgelost door een te sterke aandacht voor betrouwbaarheid van het classificatiesysteem per se, maar misschien

wel door de klinische diagnose of classificatie te vergelijken met het natuurlijke verloop van de ziekte.

Een tweede kritiekpunt is dat de auteurs van de DIS spreken over betrouwbaarheid van leken- en psychiateroordeel, daar waar ze voor beide groepen een verschillende beoordelingsprocedure hanteerden. In dat geval lijkt het ons strikt genomen onjuist om nog te kunnen spreken van interbeoordelaarbetrouwbaarheid. We komen daarop terug.

### Vergelijking van de DIS met andere systemen

De DIS is voortgekomen uit de RDI, welke laatste werd ontwikkeld vanuit de vraag of niet-artsen net zo goed als psychiaters met een gestructureerd interview zouden kunnen rubriceren. De opzet van de DIS en de RDI lijken als gevolg daarvan sterk op elkaar (vraagstellingen en codeersysteem van antwoorden). De RDI laat echter alleen een classificatie met behulp van de Feighner criteria toe, doch omvat, dit in tegenstelling tot de DIS, vragen naar de (psychiatrische) ziektegeschiedenis binnen het gezin. De concordantiecijfers voor de RDI zijn gemiddeld wat lager ( $K = .60$ ) dan die van de DIS.

De DIS laat zich moeilijk vergelijken met de Schedule of Affective Disorders and Schizophrenia (SADS). Deze laatste richt zich vooral op een onderverdeling van de affectieve en psychotische stoornissen met behulp van de Research Diagnostic Criteria.

Behalve met de RDI en de SADS, laat de DIS zich nog het beste vergelijken met de Present State Examination (PSE, Wing e.a., 1974). De PSE kan alleen door getrainde en ervaren klinici worden afgenomen. De afname is gebaseerd op de symptomatologie die zich manifesteerde in de maand voorafgaande aan het interview. Wing e.a. (1974, p. 13 e.v.), probeerden in de ontwikkeling van de PSE tijdperioden van onder andere 1 week tot 3 maanden uit en kwamen tot de bevinding dat de een-maandperiode voor de patiënt het gemakkelijkst te hanteren was. De DIS heeft de claim om zowel de huidige toestand van de patiënt te kunnen omvatten als wel de symptomatologie die ooit in zijn leven is opgetreden te kunnen classificeren ('current and lifetime diagnosis'). Op grond van onderzoek naar de zogenaamde life events, waarbij bleek dat naarmate een groter tijdsbereik werd gehanteerd met betrekking tot de meegemaakte gebeurtenissen, geheugen-lacunes een steeds grotere rol gingen spelen (Uhlenhuth e.a., 1977) en op grond van recent PSE-onderzoek (Sloof e.a., 1983 a en b: '... dat acuut psychotische patiënten op twee opeenvolgende tijdstippen een verschillend verslag geven van hun symptomatologie in de maand voorafgaande aan de interviews') lijkt deze claim overtrokken.

Er is tot dusver één vergelijkend onderzoek tussen de DIS en een aangepaste versie van de PSE, de Psychiatric Assessment Interview (PAI, Carpenter e.a., 1976) gedaan (Pulver en Carpenter, 1983). In dit

onderzoek werden 43 van de 68 personen die in St. Louis aan het International Pilot Study for Schizophrenia deelnamen (1968-1969) met de DIS geïnterviewd nadat zij in eerdere instantie (indexjaar, 2, 5 en 11 jaar na follow-up) met een aangepaste versie van de PSE (de PAI) geïnterviewd waren. De PSE/PAI-items die betrekking hadden op psychotische verschijnselen, zoals hallucinaties (visueel, auditief, tactiel/somatisch en geur) en wanen (boodschappen, spionage, vergif/achtervolging) werden met elkaar vergeleken. De DIS faalde in gemiddeld 38% van de gevallen om op levensbasis hallucinaties en in 14% van de gevallen wanen op te sporen in vergelijking met de uitkomsten van de voorafgaande PSE/PAI-afnames.

Als de DIS werd vergeleken met de index diagnose op grond van een eerdere (psychose) opname, dan bleek het instrument in 33% van de gevallen te falen. De auteurs concludeerden dat het gevaar groot is dat de DIS de psychotische verschijnselen op levensbasis onderscoort (false negatives). Als mogelijke oorzaken noemen de auteurs: gebrek aan inzicht bij de patiënt, bizarre ervaringen die moeilijk bespreekbaar zijn, vooral als ze relatief nieuw zijn en aan een vreemde verteld moeten worden, defensiemechanismen en stigmatiseringsgevaar.

Helzer e.a. (1984) zetten in antwoord op dit onderzoek vraagtekens bij de gevolgde methodiek (slechts 43 van de oorspronkelijke 68 deelnemers deden in dit onderzoek mee en 13 van de 43 werden per telefoon met de DIS geïnterviewd. Zij laten de mogelijke oorzaken voor false negatives die Pulver en Carpenter noemden buiten beschouwing, suggereren een overdiagnostisering van psychotische verschijnselen met de PSE/PAI en het IPSS-project en stellen dat het onderzoek van Pulver en Carpenter geen antwoord geeft op de vraag of leken beter of slechter zijn in het vaststellen van klinische symptomen. In antwoord hierop zeggen Pulver en Carpenter geen reden te zien hun mening ten aanzien van de onderdiagnostisering van psychotische symptomen (false negatives) te moeten herzien.

Karno e.a. (1984) staven het vermoeden van Pulver en Carpenter met betrekking tot het false negatives probleem met onderzoek van Fish (1982): 'We found that in later retrospective histories these same events were often forgotten, blurred, or completely distorted to fit a more comfortable self concept'. Tevens noemde hij de eerder genoemde onderzoeksresultaten van Uhlenhuth e.a. (1977) en concludeert: '... recall declines over time, with a significant decline detectable after 6-9 months, even for dramatic and serious experiences'. Karno wijst tenslotte ook op de verschillende bewoording van de vragen over psychotische verschijnselen in de PSE en de DIS. De PSE zou de clinicus een groter 'net' geven om de psychotische symptomen te vangen, hetgeen aanleiding zou kunnen zijn voor 'false positives' in de PSE, welke door Pulver en Carpenter systematisch als 'false negatives' van de DIS worden geïnterpreteerd.

Met dit laatste punt zijn Culver en Carpenter het niet eens omdat

## PSYCHOFARMACA

volgens hen de PSE een voldoende ingeburgerd en algemeen aanvaard klinisch interviewsysteem is om deze als vergelijkingscriterium voor de DIS te rechtvaardigen. Zij blijven dan ook vasthouden aan het 'false negatives' probleem en spreken de hoop uit dat dit in de toekomst verder wordt uitgezocht.

Hiermede is, in kort bestek, de bestaande onderzoeksliteratuur met betrekking tot de DIS besproken. Het lijkt in de verwachting te liggen (Robins, persoonlijke mededeling, 1984) dat de DIS en de PSE in het International Pilot Study of Schizophrenia (IPSS) onderzoek verder met elkaar vergeleken gaan worden.

### Samenvatting en discussie

Op grond van de DIS-training en de besproken onderzoeksliteratuur komen wij tot de volgende puntsgewijze overwegingen:

1. in tegenstelling tot wat de auteurs beweren kan de DIS niet volledig (niet met de hand en ook niet met de machine) in het DSM-III-systeem gescoord worden: de DSM-III-categorieën werden opgenomen voor zover ze met de Feighner en de Research Diagnostic Criteria overlaptten en voor zover er eenduidige criteria voor bepaalde andere categorieën aanwezig waren (met name drugmisbruik en -verslaving). De DIS kan zowel in de Feighner- als RDC-systemen gescoord worden, doch alleen de DSM-III en de RDC zijn met de machine scoorbaar. De specificiteit van de depressieclassificatie van de RDC (9 typen depressies) is in de geautomatiseerde versie weggefallen.

2. hoewel de suggestie gewekt wordt dat de DIS met weinig training door leken te gebruiken is, lijkt ons dat niet het geval. De interviewers moeten op zijn minst enige interviewervaring hebben, liefst met psychiatrische patiënten, om een gespreksband op te kunnen bouwen. Daarbij komt nog, dat de interviewer het interviewsysteem grondig moet kennen omdat anders de geautomatiseerde verwerking problemen gaat opleveren. Beide vaardigheden zijn natuurlijk niet in een week onder de knie te krijgen en vereisen naast de intensieve DIS-training van een week een begeleiding en supervisie van tenminste 10 interviews.

3. het onderdiagnostiseren van somatiseringsverschijnselen door lekeninterviewers, en dit zijn symptomen waarbij het belangrijk is om af te schatten of ze een medische dan wel psychiatrische oorzaak hebben, vereisen bijna automatisch dat de interviewer niet alleen enige klinische ervaring heeft, doch ook dat het interviewprotocol geredigeerd wordt door iemand met klinische ervaring. Dit gegeven, gevoegd bij het feit dat het redigeren nogal wat tijd vraagt, maakt dat het ons voorkomt dat het voordeel van een lekeninterview pas echt tot zijn recht komt in de grotere epidemiologische onderzoeksprojecten.

4. in het verlengde van het genoemde onder punt 2 kan de vraag worden gesteld of het methodisch en ethisch gezien een juist uitgangs-



punt is, dat een patiënt bereid en in staat is zich door een hem onbekend persoon, die bovendien geen psychiater is, een psychiatrische anamnese te laten afnemen.

5. een ander punt betreft het grote aantal vragen van de DIS (365) dat in de vierde versie nog zal worden uitgebreid met vragen over de secties 'Algemene Angststoornis', 'Posttraumatische Stressstoornis' en 'Boulimie'. Als de geïnterviewde gevraagd wordt of hij een bepaald symptoom gehad heeft of nog heeft, kan hij met een 'ja' of 'nee' antwoorden. Hij kan er van verzekerd zijn dat als hij met een 'ja' antwoordt, hem een groot aantal andere vragen gesteld zal worden, terwijl dit niet het geval is als hij met een 'nee' antwoordt. Als de geïnterviewde dit principe begint door te krijgen – en tijdens de training bleek dat dat al snel gebeurt – dan is de kans niet ondenkbeeldig dat zich een bepaalde antwoordtendens begint af te tekenen, die de betrouwbaarheid van de verkregen gegevens niet ten goede komt. Een aanwijzing hiervoor wordt in de onderzoeksliteratuur door Heltzer zelf gegeven (Heltzer e.a., 1981), waar hij de 'procedural validity' van de RDI onderzocht. Hij liet een en dezelfde patiënt driemaal interviewen en vond dat het gemiddelde aantal diagnoses per patiënt (de scoring van de symptomatologie in de RDI wordt gedaan ongeacht de hiërarchische ordening van klassen) afnam en in het eerste interview 2,7, in het tweede 2,3 en in het derde 2,0 was. Het lijkt dat verdere uitbreiding van het interviewformaat en betrouwbaarheidsonderzoek hier rekening mee dient te houden.

6. als laatste punt geldt het probleem van de 'life time diagnosis' of de 'false negatives'. Op grond van het gebruik van verschillende tijdsperiodes waarop het interview betrekking had, kwamen Wing e.a. (1974, p. 13) tot de conclusie dat de maand voorafgaande aan de psychiatrische opname het beste uitgangspunt biedt voor een betrouwbaar klinisch interviewen van de patiënt. Hij baseerde deze mening onder andere op de vergeetachtigheid van de patiënt. Dit lijkt overigens ook, zoals reeds opgemerkt (Sloof e.a., 198 ), voor de Nederlandse PSE te gelden. Soortgelijke commentaren werden gegeven door Uhlenhuth (1977) met betrekking tot life event onderzoek, Fish (1982), Pulver en Carpenter (1983) en Karno (1984) met betrekking tot longitudinaal onderzoek van psychiatrische patiënten. De argumenten die zij gaven zijn overtuigend genoeg om aarzeling te hebben met de claim van de DIS (en ook de SADS) om een diagnose op levensbasis te kunnen geven.

Aan voornoemde kritische punten zal ongetwijfeld bij de volgende versies van de DIS en het onderzoek daarmee, aandacht geschonken worden. Het lijkt ons in ieder geval een gestandaardiseerd interview dat voor verder gebruik in extra-muraal- en met name in epidemiologisch onderzoek (bijvoorbeeld voor screeningsdoeleinden) in aanmerking dient te komen: het is het beste wat er momenteel op dat terrein is.

Als laatste kunnen wij nog opmerken dat we bezig zijn de DIS en de

instructiehandleiding te vertalen en te bewerken voor Nederlands gebruik. Inlichtingen hierover kunnen bij de eerste auteur worden opgevraagd.

## Literatuur

- American Psychiatric Association, Committee on Nomenclature and Statistics. *Diagnostic and Statistical Manual, ed. 3* (1980), American Psychiatric Association, Washington DC.
- Ash, P. (1949), The Reliability of Psychiatric Diagnosis. *Journal of Abnormal and Social Psychology*, 44, 272-276.
- Blashfield, Roger K. (1984), *The Classification of Psychopathology*. Plenum Press, New York.
- Bobon, D., U. Baumann, J. Angst, H. Helmchen (Eds.) (1983), *Modern Problems of Pharmacopsychiatry*, Vol. 20. The AMDP-System in pharmacopsychiatry. S. Karger, Basel.
- Carpenter, W.T., M.H. Sacks Jr., J.S. Strauss, J.J. Bartko en J.N. Rayner (1976), Evaluating signs and symptoms: Comparison of structured interview and clinical approaches. *British Journal of Psychiatry*, 128, 397-403.
- Carpenter, W.T. Jr. en Ann E. Pulver (1984), Lifetime Psychotic Symptoms assessed with the DIS. *Schizophrenia Bulletin*, 10, 5-7.
- Carey, G. en J. Gottesman (1978), Reliability and validity in binary ratings. *Archives of General Psychiatry*, 35, 1454-1459.
- Dingemans, P., C. Hooijer en J. Dutilh (1984), ICD-9-CM coderen in de psychiatrie. *Tijdschrift voor Psychiatrie*, 6, 432-444.
- Dutilh, J. en P.M.A.J. Dingemans (1984), Een betrouwbaarheidsonderzoek van ICD-9-CM-codering. *Tijdschrift voor Psychiatrie*, 26, 212-221.
- Engeland, H. van (1980), Over ontwikkelingspsychosen. *Academisch Proefschrift*, Rijks Universiteit Utrecht. Krips Repro, Meppel.
- Feighner, J.P., E. Robins, S.B. Guze, R. Woodruff, G. Winokur en R. Munoz (1972), Diagnostic Criteria for use in Psychiatric Research. *Archives of General Psychiatry*, 26, 57-63.
- Greist, John, James W. Jefferson and Robert L. Spitzer (Eds.) (1982), *Treatment of Mental Disorders*. Oxford University Press, New York, Oxford.
- Helzer, John E. en Lee N. Robins (1984), Lifetime Psychotic Symptoms assessed with the DIS. *Schizophrenia Bulletin*, 10, 4-5.
- Helzer, John E., Lee N. Robins, Jack L. Croughan and Amos Welner (1981), Renard Diagnostic Interview. *Archives General Psychiatry*, 38, 393-399.
- Karno, Marvin (1984), Lifetime Psychotic Symptoms Assessed with the DIS. *Schizophrenia Bulletin*, 10, 2, 154-157.
- Kendell, R.E., B. Everett en J.E. Cooper (1968), The Reliability of the 'Present State Examination'. *Social Psychiatry*, 3, 123-129.
- Kendell, R.E. (1968), *The Classification of Depressive Illness*. Oxford University Press, London.
- Kendell, R.E. en A.J. Gourlay (1970), The clinical distinction between psychotic and neurotic depression. *British Journal of Psychiatry*, 117, 257-260.
- Kendell, R.E., J.E. Cooper, A.J. Gourlay, J.R.M. Copeland, L. Sharpe en B.J. Gurland (1971), Diagnostic Criteria of American and British Psychiatrists. *Archives of General Psychiatry*, 25, 123-130.
- Maxwell, A.E. (1977), Coefficients of agreement between observers and their

- interpretation. *British Journal of Psychiatry*, 130, 79-83.
- Pen, A., W.A. Arrindell, W. Lawant en J. Pols (1977), De betrouwbaarheid van de 'Present State Examination'. *Tijdschrift voor Psychiatrie*, 10, 375-388.
- Pulver, A.E. en W.T. Carpenter Jr. (1983), Lifetime Psychotic Symptoms assessed with the DIS. *Schizophrenia Bulletin*, 9, 377-382.
- Robins, Lee N., John E. Helzer, Jack Croughan en Kathryn Ratcliff (1981), National Institute of Mental Health Diagnostic Interview Schedule. *Archives General Psychiatry*, 38, 381-390.
- Robins, Lee N., John E. Helzer, Jack Croughan, B.W. Janet en Robert L. Spitzer (1979), *NIMH Diagnostic Interview Schedule: version III*. National Institute of Mental Health, Bethesda, Maryland.
- Silbermann, R.M. (1971), *CHAM, A Classification of Psychiatric States*, 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, 6/7, 454-469.
- Slooff, C.J., W.R.E.H. Mulder-Hajonides van der Meulen en R.H. van den Hoofdkakker (1983a), De Nederlandse vertaling van de Present State Examination, 9e editie. 1. Betrouwbaarheidsaspecten. *Tijdschrift voor Psychiatrie*, 25, 3, 151-163.
- Slooff, C.J., W.R.E.H. Mulder-Hajonides van der Meulen en R.H. van den Hoofdkakker (1983b), De Nederlandse vertaling van de Present State Examination, 9e editie. 2. De relatie tussen Catego-klassen en categorieën van de 'International Classification of Diseases'. *Tijdschrift voor Psychiatrie*, 25, 4, 280-289.
- Spitzer, R.L. en J.L. Fleiss (1974), A re-analysis of the reliability of psychiatric diagnosis. *British Journal of Psychiatry*, 25, 341-347.
- Spitzer, R.L., J. Endicott en E. Robins (1978), Research Diagnostic Criteria: rational and reliability. *Archives General Psychiatry*, 35, 773-782.
- Spitzer, R.L. en Donald F. Klein (Eds.) (1978). *Critical Issues in Psychiatric Diagnosis*. Raven Press, New York.
- Spitzer, R.L. en J. Endicott (1979), Schedule of Affective Disorders and Schizophrenia. Biometrics Research Division, New York State Department of Mental Hygiene, New York.
- Spitzer, Robert L., Janet W.B. Williams en Andrew E. Skodol (Eds.) (1983), *International Perspectives on DSM-III*. American Psychiatric Press, Washington DC.
- Srole, L., T.S. Langer en S.T. Michael (1962), Mental Health in Metropolis: The Midtown Manhattan Study, McGraw-Hill Book Co. vol. 1, New York.
- Uhlenhuth, E.H., M.D. Balter, R.S. Lipman en S.J. Haberman (1977), Remembering Life Events. In: Strauss, J.S., H.M. Babigian en M. Roff (Eds), *The Origins and Course of Psychopathology: Methods of Longitudinal Research*, p. 117-132. Plenum Press Inc., New York.
- Ward, C.H., A.T. Beck, M. Mendelson, J.E. Mock en J.K. Eerbangh (1962), The Psychiatric Nomenclature. *Archives of General Psychiatry*, 7, 198-205.
- Wing, J.K., J.E. Cooper en N. Sartorius (1974), *The Measurement and Classification of Psychiatric Symptoms*, Cambridge University Press, New York.
- World Health Organization. Mental Disorders: *Glossary and guide to their classification in accordance with the Ninth Revision of the International Classification of Diseases*, 1978, World Health Organization, Geneva (Uitg.: Nijhof, Den Haag).