

Richtlijnen voor klinische diagnostiek van DSM-IV-persoonlijkheidsstoornissen

R. VERHEUL, W. VAN DEN BRINK, PH. SPINHOVEN, R. HARINGSMA

ACHTERGROND Klinische diagnostiek van persoonlijkheidsstoornissen is onvoldoende betrouwbaar en valide. Er bestaat in de klinische praktijk geen consensus over de meest aangewezen diagnostische procedure.

DOEL In dit artikel worden richtlijnen geformuleerd voor klinische diagnostiek van DSM-IV-persoonlijkheidsstoornissen.

METHODE Als uitgangspunt voor deze overzichtsstudie van diagnostische procedures voor de klinische diagnostiek van persoonlijkheidsstoornissen zijn drie kwaliteitseisen gekozen: betrouwbaarheid, validiteit en uitvoerbaarheid. De bestaande meetmethoden worden systematisch geëvalueerd aan de hand van deze criteria, alvorens te komen tot richtlijnen waarbij getracht wordt een optimale balans te bereiken in de mate waarin aan deze criteria is voldaan.

RESULTATEN De betrouwbaarheid van de meting is gebaat bij structurering van de dataverzameling en standaardisatie van de scoring. Hiertoe vormt een semi-gestructureerd interview het beste middel. Als waarborgen voor voldoende validiteit worden aangemerkt: de beschikbaarheid van anamnestiche gegevens, multipale bronnen van informatie en heldere operationele definities van de algemene en specifieke diagnostische criteria. In het kader van de uitvoerbaarheid wordt bepleit om, voorafgaande aan het afnemen van het semi-gestructureerde interview, onderzoekshypothesen te genereren op basis van een combinatie van een screeningsvragenlijst, klinische indrukken, en eventueel andere informatiebronnen.

CONCLUSIES Betrouwbare en valide diagnostiek van persoonlijkheidsstoornissen is een veel-eisende taak die zowel voldoende kennis van DSM-IV als uitgebreide en intensieve training in het afnemen van een semi-gestructureerd interviewschema vereist.

[TIJDSCHRIFT VOOR PSYCHIATRIE 42 (2000) 6, 409-422]

TREFWOORDEN diagnostiek, persoonlijkheidsstoornissen, richtlijnen

De aandacht voor persoonlijkheidspathologie bij patiënten in de GGZ en verslavingszorg is de laatste jaren snel toegenomen. Deze aandacht komt onder andere tot uitdrukking in het grote aantal wetenschappelijke publicaties op dit terrein en de ontwikkeling van een breed scala van (semi-)gestructureerde meetmethoden om de betrouwbaarheid en validiteit van diagnostiek te bevorderen. In de klinische praktijk heeft deze

ontwikkeling echter vooralsnog slechts in beperkte mate haar beslag gekregen. Uit een recente studie bleek dat er wat betreft de diagnostiek van persoonlijkheidsstoornissen een grote kloof bestaat tussen wetenschappelijke onderzoeksmethoden en de klinische praktijk (Westen 1997). Westen liet zien dat, hoewel klinici As I-stoornissen vaak diagnosticeren aan de hand van vragen die direct zijn afgeleid van de DSM-

IV-criteria, zij deze werkwijze over het algemeen minder bruikbaar achten voor de meting van persoonlijkheidsstoornissen. Favoriet bleken daarentegen het luisteren naar hoe patiënten zichzelf en interacties met anderen beschrijven alsook het observeren van het gedrag van de patiënt in de klinische setting. Deze bevinding bleek onafhankelijk van de theoretische oriëntatie van de clinicus en de specifieke persoonlijkheidsstoornis. Het behoeft geen verder betoog dat de betrouwbaarheid en daarmee de validiteit van de klinische diagnostiek van persoonlijkheidsstoornissen met deze situatie niet is gediend. Met het thans voorliggende artikel wordt beoogd om, aan de hand van een drietal kwaliteitseisen en een evaluatie van bestaande methoden, te komen tot richtlijnen voor een betrouwbare, valide en tegelijkertijd klinisch uitvoerbare en haalbare diagnostiek van DSM-IV-persoonlijkheidsstoornissen.

KWALITEITSEISEN VOOR KLINISCHE DIAGNOSTIEK

In het algemeen kunnen aan klinische diagnostiek drie kwaliteitseisen worden gesteld: betrouwbaarheid, validiteit en haalbaarheid. Onder betrouwbaarheid wordt verstaan de mate van overeenstemming tussen de testuitslagen die verkregen worden wanneer een test herhaaldelijk wordt toegepast bij dezelfde persoon. Voldoende betrouwbaarheid is een noodzakelijke, maar op zichzelf onvoldoende voorwaarde voor de validiteit van een diagnostische test ofwel de mate waarin een diagnostische observatie de werkelijke klinische toestand benadert. Een perfecte betrouwbaarheid kan immers hand in hand gaan met systematische afwijkingen tussen de waargenomen en de werkelijke klinische toestand. Ten slotte valt of staat de acceptatie van iedere diagnostische methode met haar haalbaarheid. In de gezondheidszorg worden bij voorkeur diagnostische tests gebruikt die snel, eenvoudig, goedkoop en weinig belastend zijn.

Hieronder worden eerst de meest voor de

hand liggende strategieën besproken die de betrouwbaarheid, validiteit en haalbaarheid van de diagnostiek van persoonlijkheidsstoornissen bevorderen.

Waarborgen voor de betrouwbaarheid

Volgens de DSM-IV dienen persoonlijkheidsstoornissen te worden gediagnosticeerd aan de hand van algemene en specifieke diagnostische criteria. De specifieke diagnostische criteria, die de basis vormen voor de meeste (gestructureerde) meetinstrumenten, hebben inmiddels in brede kring ingang gevonden. De algemene diagnostische criteria (zie voor een bespreking onder 'Stroomdiagram 2, 8. Is er voldaan aan de algemene diagnostische criteria?'), die van toepassing zijn op elk van de 79 specifieke diagnostische criteria, worden daarentegen in de alledaagse praktijk vaak veronachtzaamd. De beoordeling van de specifieke criteria en de verificatie van de algemene criteria zijn veelomvattende taken die gebaat zijn bij een optimale structurering van de dataverzameling en standaardisatie van de scoring.

Structurering van de dataverzameling Bij gebrek aan structuur loopt de clinicus het gevaar om deze beoordelingsprocedure halfslachtig uit te voeren alsook om de nadruk te veel op de meest opvallende kenmerken te leggen. Volledige structurering wordt bij diagnostiek van persoonlijkheidsstoornissen ook niet ideaal geacht, met name omdat de beoordeling nogal wat reflectie en interpretatie vergt. Om deze redenen hebben de meeste interviewinstrumenten voor persoonlijkheidsstoornissen een semi-gestructureerd karakter.

Standaardisatie van de scoring Bij de diagnostiek van persoonlijkheidsstoornissen hebben we te maken met pathologische persoonlijkheidseigenschappen die zich met name kwantitatief (en veel minder duidelijk kwalitatief) van normale eigenschappen onderscheiden (Widiger & Costa 1994). De beslissing of een bepaalde eigen-

schap al dan niet binnen het maladaptieve spectrum valt, is per saldo arbitrair. Dit gegeven vormt een potentiële bron van onbetrouwbaarheid, waartegen wij ons enigermate kunnen wapenen door standaardisatie (of explicitering) van de scoringsregels (bij voorbeeld door middel van ankerpunten).

Waarborgen voor de validiteit

Er bestaat geen algemeen aanvaard extern beoordelingscriterium (of gouden standaard) voor persoonlijkheidsstoornissen. De validiteit van het bestaande instrumentarium kan daarom eigenlijk alleen op indirecte wijze worden bestudeerd. Een van de meest gebruikelijke strategieën in dit opzicht is het vergelijken van de meetuitkomsten van een instrument met die van een ander instrument dat geacht wordt hetzelfde of een nauw verwant concept te meten (convergente validiteit). Een andere manier om validiteit in te schatten is nagaan of het instrument het totale domein van het te meten fenomeen in al zijn aspecten dekt (inhoudsvaliditeit). In dit kader kunnen als waarborgen van de validiteit worden aangemerkt: de beschikbaarheid van (1) anamnestiche gegevens, (2) meerdere bronnen van informatie, en heldere operationele definities van de (3) specifieke en (4) algemene diagnostische criteria.

Anamnestiche gegevens Vragen over de persoonlijkheid dienen betrekking te hebben op relatief symptoomvrije perioden uit het alledaagse leven van de geïnterviewde. Anamnestiche gegevens stellen de interviewer in staat om de perioden tussen aanvang en remissie van eventuele psychiatrische symptomatologie als ook de perioden van hospitalisatie of detentie zoveel mogelijk uit te sluiten.

Meerdere bronnen van informatie Soms is de patiënt zelf niet de meest betrouwbare en valide bron van informatie. Dit geldt vooral wanneer de patiënt kampt met ernstige psychiatrische

symptomen, een sterke neiging heeft tot ontkennen of het geven van sociaal-wenselijke antwoorden, of anderszins in beperkte mate in staat of bereid is zichzelf adequaat te beoordelen. Het gebruik van meerdere bronnen van informatie kan hier uitkomst bieden.

Specifieke diagnostische criteria De specifieke diagnostische criteria worden in eerste instantie beoordeeld aan de hand van vragen over de (gedragsmatige) manifestatie van het betreffende criterium. Het is daarom van belang om te beschikken over beschrijvingen van de specifieke diagnostische criteria die concreter en explicieter zijn dan de vaak beknopte en welhaast cryptische definities in de DSM-IV.

Algemene diagnostische criteria Ten slotte dient te worden nagegaan in hoeverre het geconstateerde gedragspatroon voldoet aan de algemene diagnostische criteria. Dit verificatieproces kan alleen op valide wijze geschieden door gericht te vragen naar frequentie, intensiteit, beloop, gevolgen en mogelijke oorzaken van het gedrag.

Waarborgen voor de haalbaarheid

De belangrijkste strategie om de haalbaarheid van de meting van persoonlijkheidspathologie te vergroten is het minimaliseren van de interviewtijd, bijvoorbeeld door die stoornissen buiten beschouwing te laten die waarschijnlijk niet aanwezig zijn.

EVALUATIE VAN DE BESCHIKBARE MEETMETHODEN

Grofweg kunnen worden onderscheiden: zelfbeoordelingvragenlijsten, semi-gestructureerde interviews, het traditionele klinische interview en protocollen waarbij meerdere bronnen van informatie worden geïntegreerd. Wat betreft betrouwbaarheid, validiteit en haalbaarheid zijn er duidelijke verschillen (tabel 1).

TABEL 1 Classificatie van meetmethoden ten behoeve van As II

Waarborg	Betrouwbaarheid			Validiteit			Haalbaarheid
	Methode	Mate van structurering	Mate van standaardisatie	Anamnestiche gegevens	Meerdere bronnen	Verificatie algemene criteria	
Zelfbeoordelingsvragenlijsten							
PDQ-4+	hoog	hoog	nee	nee	beperkt	beperkt	uitstekend
VKP	hoog	hoog	nee	nee	nee	beperkt	uitstekend
ADP-IVh	hoog	hoog	nee	nee	beperkt	beperkt	uitstekend
Semi-gestructureerde interviews							
SIDP-IV	middel	middel	advies	ja + advies	beperkt	ja	redelijk
SCID-II	middel	middel	nee	ja	nee	ja	goed
IPDE	middel	hoog	beknopt	ja	beperkt	ja	redelijk
Ongestructureerde meetmethoden							
Klinisch interview	laag	laag	ja	mogelijk	nee	beperkt	goed
Dossier-evaluatie	laag	laag	mogelijk	nee	nee	mogelijk	uitstekend
Protocollen voor bronnenintegratie							
LEAD-procedure	middel	consensus	ja	ja	ja	mogelijk	slecht
Richtlijnen in dit artikel	middel	hoog	ja	ja	ja	ja	redelijk

Zelfbeoordelingsvragenlijsten Voor het Nederlands taalgebied zijn drie vragenlijsten beschikbaar voor de diagnostiek van persoonlijkheidsstoornissen: de Personality Diagnostic Questionnaire DSM-IV-versie (PDQ-4+; Akkerhuis e.a. 1996), Assessment van DSM Persoonlijkheidsstoornissen DSM-IV-versie (ADP-IV; Schotte & De Doncker 1994), en de Vragenlijst voor Kenmerken van de Persoonlijkheid DSM-III-R-versie (VKP; Duijsens e.a. 1993; DSM-IV-versie in voorbereiding). De vragenlijsten komen in grote lijnen overeen, in die zin dat de meeste

specifieke diagnostische criteria worden beoordeeld aan de hand van een enkel ja/nee-item. Er zijn onderlinge verschillen in de wijze waarop wordt omgegaan met de algemene diagnostische criteria. In de PDQ-4+ vormt een score van 2 of hoger op de 'Impairment and Distress' (ID)-schaal, naast de vereiste scores op de items voor de specifieke criteria, een voorwaarde voor het stellen van een diagnose. In de ADP-IV wordt bij iedere bevestiging van een item afzonderlijk gevraagd naar de ervaren lijdensdruk en/of functionele beperkingen als gevolg van de betreffende

persoonlijkheidseigenschap. In beide vragenlijsten wordt aldus getracht criterium C van de algemene diagnostische criteria te meten.

De betrouwbaarheid van zelfbeoordelingsvragenlijsten is, door de hoge mate van gestructureerdheid en standaardisatie, in principe gewaarborgd. De haalbaarheid is uitstekend te noemen. De validiteit is daarentegen niet of nauwelijks gewaarborgd. De persoonlijkheid wordt gemeten zonder dat rekening wordt gehouden met het beloop van eventuele As I-stoornissen en de behandelgeschiedenis. Voorts wordt niet of nauwelijks geverifieerd in hoeverre voldaan is aan de algemene diagnostische criteria. Voor de verificatie van de specifieke diagnostische criteria geldt dat er bij het gebruik van vragenlijsten vaak ten onrechte van wordt uitgegaan dat patiënten over voldoende introspectief vermogen beschikken. Vanwege de validiteitsproblemen zijn vragenlijsten ongeschikt voor het stellen van klinische diagnoses.

Semi-gestructureerde interviews Voor het Nederlands taalgebied zijn drie semi-gestructureerde interviewschema's beschikbaar voor het brede spectrum van As II. Dit zijn de SIDP-IV (Gestructureerd interview voor de DSM-IV persoonlijkheidsstoornissen; De Jong e.a. 1996), SCID-II (Gestructureerd klinisch interview voor DSM-IV persoonlijkheidsstoornissen; Weertman e.a. 1996), en IPDE (International Personality Disorder Examination; Diekstra e.a. 1996). Deze interviewschema's bestaan uit voorgestructureerde vragen die zoveel mogelijk letterlijk aan de patiënt moeten worden gesteld. Doorgaans neemt de interviewer op basis van het antwoord op de openingsvraag de beslissing om al dan niet door te vragen. De belangrijkste verschillen betreffen onder andere de mate waarin de criteria worden beschreven en expliciete gedragsindicatoren worden benoemd (alleen IPDE), en de mate waarin voorzien is in criteriumspecifieke ankerpunten voor scoring (alleen SCID-II en IPDE). Verder hanteren de interviews verschillende tijdsperioden als referentie-

kader: zo hanteren de SIDP-IV en SCID-II de afgelopen vijf jaar als referentiekader, terwijl volgens de IPDE geen positieve score kan worden gegeven wanneer het gedrag of de trek gedurende ten minste vijf jaar afwezig is geweest.

De betrouwbaarheid van semi-gestructureerde instrumenten is over het algemeen redelijk goed (Zimmerman 1994). Voor de validiteit geldt dat deze bij de semi-gestructureerde interviewschema's over het algemeen beter is gegarandeerd dan bij zelfbeoordelingsvragenlijsten, zij het dat er onderling nogal wat verschillen zijn. Van de beschikbare instrumenten biedt de IPDE de meeste standaardisatie (in de vorm van een uitvoerig glossarium met criteriumbeschrijvingen en criteriumspecifieke scoringsrichtlijnen). De benodigde interviewtraining en relatief lange afnameduur van deze instrumenten vormen echter belangrijke drempels die een brede implementatie vooralsnog in de weg hebben gestaan.

Het traditionele klinische interview Een voordeel van het traditionele klinische interview is dat anamnestiche gegevens, mits verzameld, volop bij het oordeel kunnen worden betrokken. Voorts is de clinicus in staat om een inschatting te maken van de waarde van de zelfbeoordeling van de patiënt als een betrouwbare bron van informatie. Afhankelijk van zijn inschatting van het toestandbeeld en het introspectieve vermogen van de patiënt maakt hij gebruik van dossier- en observatiemateriaal als primaire of secundaire bron. Zo bezien vormt het klinische interview een mooi voorbeeld van het gebruik van meerdere bronnen van informatie, zij het zonder een welomschreven algoritme voor integratie.

Aan het klinische interview kleven echter verschillende bezwaren. Verschillende onderzoeken laten zien dat clinici geneigd zijn af te wijken van de As II-criteria volgens de DSM-classificatie (Blashfield & Herkov 1996). Ook blijken zij vaak maar één diagnose op As II te stellen, waar met (semi-)gestructureerde instrumenten doorgaans meerdere diagnoses naast elkaar worden gevon-

den (Westen 1997). Het gevaar van het (deels) loslaten van de diagnostische criteria is ook dat kenmerken van de diagnosticus een rol gaan spelen bij de diagnose. Bij de borderline persoonlijkheidsstoornis bijvoorbeeld tenderen vrouwelijke klinici naar overrapportage en mannelijke klinici naar onderrapportage, en zijn minder ervaren klinici geneigd de diagnose te stellen bij patiënten die aan minder dan het benodigde aantal criteria voldoen, terwijl meer ervaren klinici vaker geneigd zijn om zelfs bij patiënten die ruim aan de criteria voldoen, de diagnose niet te stellen (Morey & Ochoa 1989). Mede als gevolg van het loslaten van de diagnostische criteria is de betrouwbaarheid van het traditionele klinische interview onvoldoende. De interbeoordelaarsbetrouwbaarheid van onderzoeken vóór de DSM-III met ongestructureerde methoden is met een gemiddelde kappa van 0,32 uitgesproken laag (Spitzer & Fleiss 1974). Anders dan men zou verwachten, heeft de komst van de tweede as van de DSM-classificatie deze lage waarden voor de interbeoordelaarsbetrouwbaarheid niet of nauwelijks verbeterd (Zimmerman 1994).

Samenvattend voldoet het traditionele klinische interview niet aan de eis van voldoende betrouwbaarheid en, mede als gevolg daarvan, de eis van voldoende validiteit. Om die reden moet de nodige voorzichtigheid worden betracht bij de interpretatie van aldus verkregen diagnoses. De haalbaarheid van het klinische interview is in de meeste gevallen goed te noemen.

Protocollen voor bronnenintegratie Het meest bekende protocol voor de integratie van verschillende informatiebronnen ten behoeve van de diagnostiek van persoonlijkheidsstoornissen is de LEAD-standaard (Spitzer 1983). Het acroniem 'LEAD' staat voor: Longitudinaal, Expert en Alle Data. Deze naam is bewust gekozen om te benadrukken dat de ontwikkeling van een 'gouden' standaard voor persoonlijkheidsstoornissen tot de onmogelijkheden behoort, en dat de beste benadering hiervan als 'loden' standaard kan worden aangemerkt. In zijn artikel pleit Spitzer

voor het verbeteren van de validiteit van het gebruikelijke klinische oordeel door meer aandacht te besteden aan de betrouwbaarheid, meerdere bronnen van informatie te gebruiken, en de psychische toestand van de patiënt over de tijd heen te volgen. De diagnose wordt uiteindelijk gesteld op basis van consensus van verschillende ervaren klinici die alle aanwezige informatie naar waarde proberen te schatten. In een validiteitstudie bleken LEAD-diagnoses te worden gekenmerkt door meer temporele stabiliteit en predictieve validiteit dan diagnoses verkregen met verschillende semi-gestructureerde interview-schema's (Pilkonis e.a. 1991). Implementatie van de LEAD-standaard in de klinische praktijk wordt in de weg gestaan door de enorme tijdsbesteding die hiermee gemoeid is. Ten slotte kan worden opgemerkt dat, voorzover geen algoritme voor de integratie van eventueel tegenstrijdige informatie wordt gehanteerd, het daadwerkelijke succes van deze methode valt of staat met de objectiviteit en onpartijdigheid van de voorzitter van de consensusbijeenkomst.

RICHTLIJNEN VOOR KLINISCHE DIAGNOSTIEK

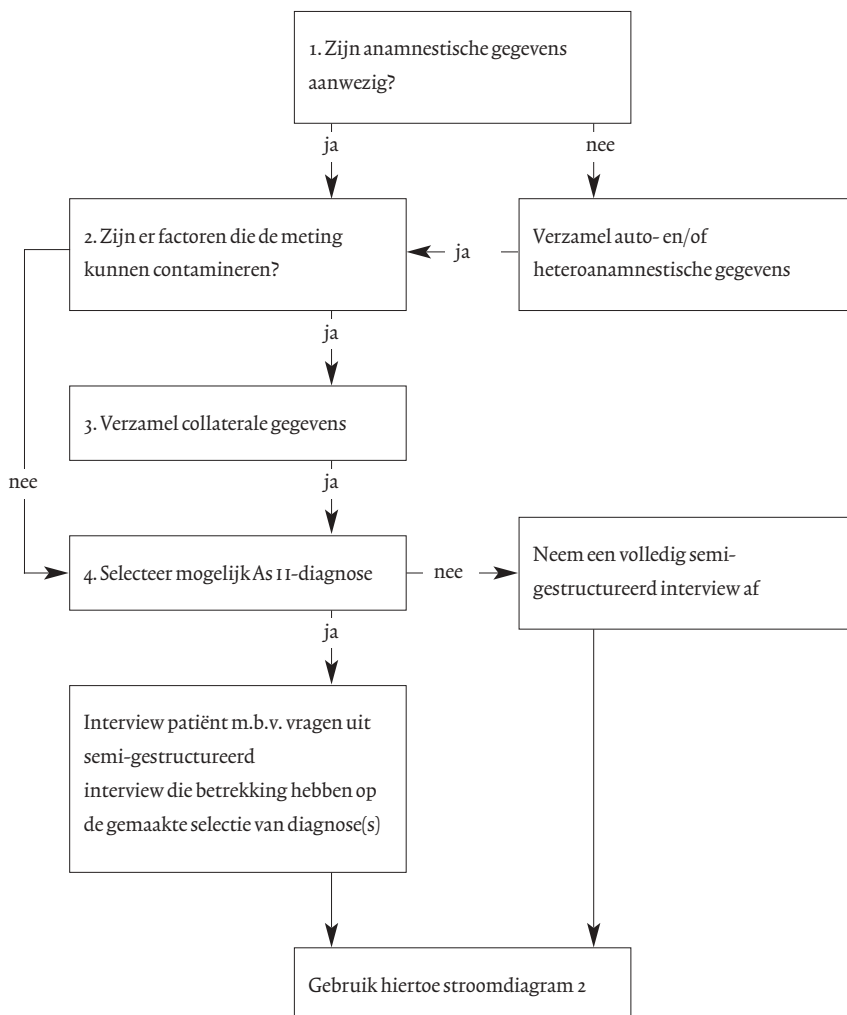
Tot dusverre is duidelijk geworden dat de bestaande methoden voor diagnostiek van persoonlijkheidsstoornissen, overigens om heel verschillende redenen, niet ideaal zijn: ze ontberen voldoende betrouwbaarheid (met name het traditionele klinische interview), voldoende validiteit (met name de zelfbeoordelingsvragenlijsten), of voldoende haalbaarheid (met name de LEAD-procedure). Hieronder worden twee stroomdiagrammen beschreven die kunnen worden gebruikt als richtlijnen voor klinische diagnostiek van persoonlijkheidsstoornissen. Bij de ontwikkeling hiervan was het belangrijkste uitgangspunt om een optimale balans tussen betrouwbaarheid, validiteit en haalbaarheid te bereiken. Het gebruik van een semi-gestructureerd interviewschema vormt de belangrijkste waarborg voor de betrouwbaarheid van de

meting (zie stroomdiagram 2). Om ook de validiteit zoveel mogelijk te waarborgen wordt standaard gebruikgemaakt van anamnestiche gegevens en, in gevallen waarin de clinicus dit nodig acht, van collaterale informatie ofwel informatie verkregen uit andere bronnen dan de patiënt zelf. Teneinde de haalbaarheid van de meting te bevorderen wordt, op basis van hypothesevorming, de afname van het interview beperkt tot een beoordeling van de diagnostische criteria van die stoornissen die men bij de patiënt vermoedt. Alle beslismomenten voorafgaande aan de afname van het semi-gestructureerde interview worden beschreven in stroomdiagram 1.

Stroomdiagram 1

1. Zijn anamnestiche gegevens aanwezig? Een sine qua non voor een valide meting van persoonlijkheidsstoornissen is de beschikbaarheid van gegevens over het beloop van eventuele symptoomstoornissen, middelengebruik, de behandelgeschiedenis en eventuele detentieperioden.
2. Zijn er factoren die de meting kunnen contamineren? Als sterke contra-indicaties voor het diagnosticeren van persoonlijkheidsstoornissen louter en alleen op basis van een zelfbeoordeling kunnen worden aangewezen: (1) de aanwezigheid van flo-

STROOMDIAGRAM 1: Richtlijnen voor het vaststellen welke bron(nen) van informatie dienen te worden geraadpleegd voor het verzamelen van informatie



ride psychiatrische symptomatologie, (2) een intelligentie op zwakbegaafd of lager niveau, en (3) de neiging tot sociaal-wenselijk antwoorden, instemmen of ontkennen (Widiger e.a. 1995). Voorzover deze condities kunnen interfereren met de validiteit van antwoorden op directe vragen naar de aanwezigheid van persoonlijkheidskenmerken, is het raadzaam om te beschikken over zoveel mogelijk collaterale gegevens (bijvoorbeeld patiëntdossier, heteroanamnese, informantoordeel, klinische observaties).

3. *Verzamel collaterale gegevens* Klinische gedragsobservaties gelden zonder meer als een zinvolle additionele bron van informatie die kan bijdragen aan de validiteit van de meting van persoonlijkheidspathologie. Van belang is dat de informatie goed wordt gedocumenteerd en zo mogelijk wordt gesystematiseerd. Ook kan de observerende staf worden gevraagd om een lijstje met As II-criteria te scoren. De scores kunnen dan tijdens het patiëntinterview worden vergeleken met de zelfbeoordeling.

Semi-gestructureerde interviewschema's kunnen ook worden toegepast bij individuen die de patiënt van nabij kennen, ofwel bij informanten. De validiteit van het informantoordeel is echter onbekend (Zimmerman 1994; Verheul & Van den Brink 1999). Een aantal studies hebben laten zien dat de diagnostische overeenkomst van het patiënt- en informantoordeel samenhangt met de kwaliteit van de patiënt-informantrelatie en de afwezigheid van persoonlijkheidspathologie bij de informant (o.a. Riso e.a. 1994). Deze studies suggereren dat het van belang is kritisch te zijn bij de keuze van het type informant. Conform de handleiding van de SIDP-IV, waarin expliciet wordt aangeraden om – in aansluiting op het patiëntinterview – een informant te interviewen, verdient het de voorkeur om de informant alleen die vragen te stellen die verwijzen naar observeerbare en/of interpersoonlijke kenmerken. De beoordeling van minder duidelijk observeerbare, intrapsychische kenmerken, zoals stemming en zelfper-

ceptie, is vermoedelijk van minder waarde.

De keuze voor het verzamelen van collaterale informatie kan ook afhankelijk worden gemaakt van het type persoonlijkheidspathologie dat men bij de patiënt vermoedt. Uit verschillende empirische studies blijkt bijvoorbeeld dat de narcistische persoonlijkheidsstoornis zelden op basis van patiëntinterviews kan worden gediagnosticeerd, maar wel wanneer informanten worden geïnterviewd en/of andere collaterale informatie wordt gebruikt (Riso e.a. 1994; Skodol e.a. 1988). In die situaties waarin men een narcistische stoornis vermoedt, is het daarom aan te raden om collaterale gegevens mee te wegen. Hetzelfde geldt ook voor patiënten bij wie de aanwezigheid van een paranoïde, schizoïde of theatrale persoonlijkheidsstoornis wordt vermoed (Westen 1997).

4. *Selecteer mogelijke As II-diagnoses* Het kan veel tijd besparen wanneer het interview wordt beperkt tot de diagnostische criteria van die stoornissen die men bij de patiënt vermoedt. Verschillende onderzoekers hebben gepleit voor screening aan de hand van zelfbeoordelvragenlijsten (Zimmerman 1994; Duijsens 1996). Hierbij zijn de volgende kanttekeningen op hun plaats. In de eerste plaats verschilt de sensitiviteit van screeningsvragenlijsten van stoornis tot stoornis. Duijsens (1996) vond bijvoorbeeld dat de schalen van de Vragenlijst voor Kenmerken van de Persoonlijkheid (VKP) voor alle cluster-B-stoornissen alsook de obsessieve compulsieve persoonlijkheidsstoornis een verre van optimale sensitiviteit hebben. Bovendien zijn zelfbeoordelvragenlijsten waarschijnlijk gevoeliger voor vertekeningen door het toestandsbeeld en antwoordtendenties (Widiger e.a. 1995). Op dit moment lijkt de beste strategie om, in het kader van tijdsbesparing bij het semi-gestructureerde interview, tot onderzoekshypothesen te komen vanuit een combinatie van screening, klinische indrukken en eventueel collaterale informatie.

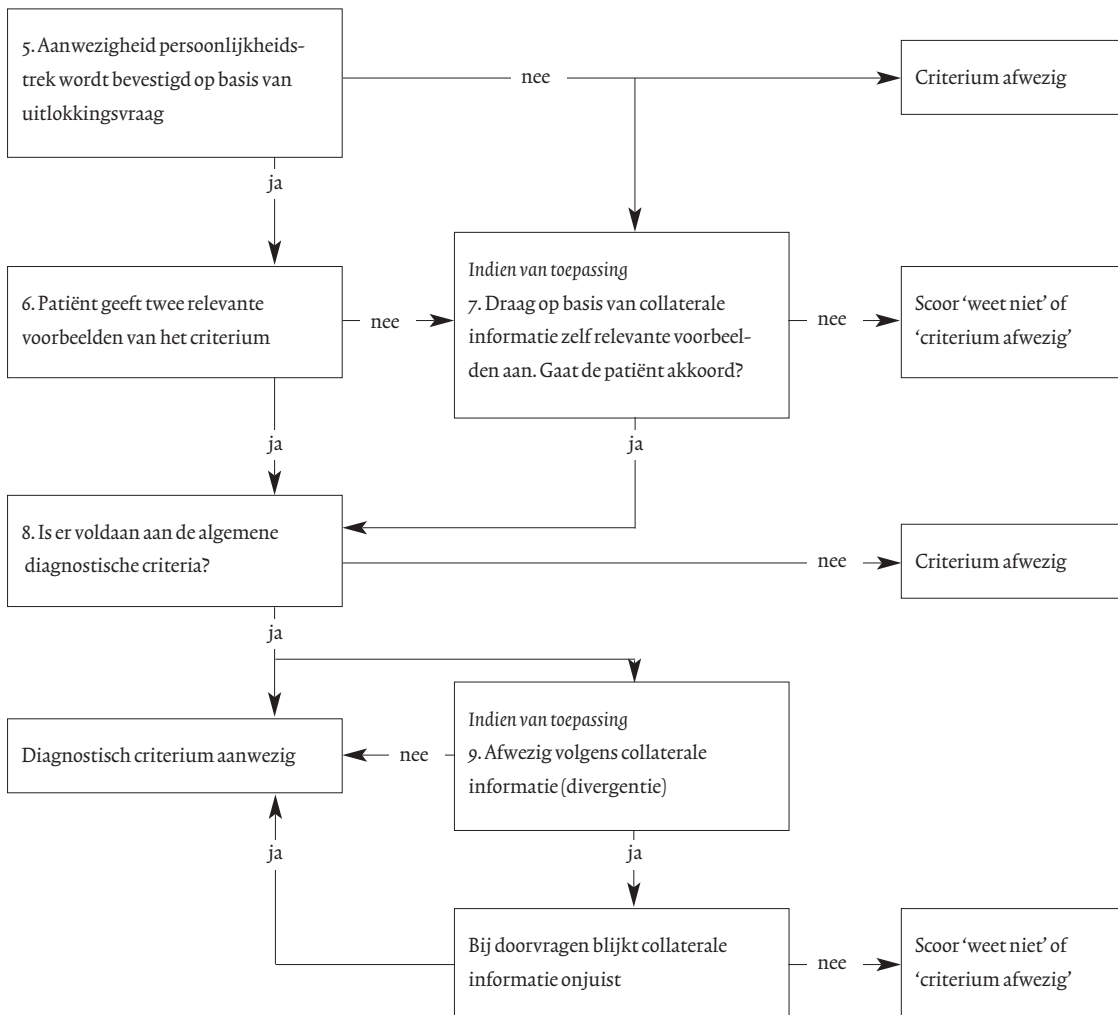
Stroomdiagram 2

Alle punten hieronder hebben betrekking op ieder van de diagnostische criteria die worden uitgevraagd met een semi-gestructureerd interviewschema. Stroomdiagram 2 dient dus voor ieder criterium opnieuw te worden doorlopen. Dit is een complexe en veeleisende taak, die slechts na intensieve training en bij voldoende interviewervaring op een meer routineuze wijze kan worden volbracht. Ook wanneer gebruik wordt gemaakt van een semi-gestructureerd interviewschema dat voorziet in expliciete

(gedrags)indicatoren en criteriumspecifieke scoringsinstructies, zal de diagnosticus vaak meer of minder arbitraire keuzen moeten maken tussen scoringsalternatieven. Onderstaande richtlijnen kunnen in die situaties enig houvast bieden.

5. Aanwezigheid persoonlijkheidstrek wordt bevestigd op basis van uitlokkingsvraag Ieder interviewschema hanteert andere standaarduitlokkingsvragen. De exacte formuleringen zijn dus niet heilig, hoewel het omwille van een acceptabele betrouwbaarheid raadzaam is de geboden formuleringen in eerste instantie te volgen.

STROOMDIAGRAM 2: Richtlijnen voor het vaststellen van de aanwezigheid van As II-diagnostische criteria met behulp van vragen uit semi-gestructureerd interview



6. *Patiënt geeft twee relevante voorbeelden van het criterium* Vraag u altijd af in hoeverre de voorbeelden die de geïnterviewde geeft, illustratief zijn voor het diagnostisch criterium. Bij twijfel dient doorgevraagd te worden. Dit is dikwijls het meest tijdrovende onderdeel van het interview-schema. Niet zelden zal de patiënt zich weigerachtig tonen voorbeelden van de persoonlijkheidstrek in kwestie te geven. Als uitgangspunt geldt dat men respect toont voor de onderliggende motieven (betrokkene vindt het bijvoorbeeld moeilijk of pijnlijk om over het onderwerp te praten). Dat hoeft evenwel niet te betekenen dat men de vraag overslaat. Een gereserveerde houding bij de geïnterviewde kan een reden zijn om (opnieuw) aandacht te besteden aan de kwaliteit van het contact met de geïnterviewde. Het is een veelgehoorde fabel dat het gebruik van semi-structureerde interviewschema's een goed contact in de weg zou staan. Ervaren interviewers rapporteren doorgaans dat geïnterviewden zich bij afname juist serieus genomen voelen en dat het contact heel plezierig kan verlopen. Voorwaarde is wel dat het interview op adequate wijze wordt geïntroduceerd en dat men, indien nodig, als het ware een alliantie met de patiënt aangaat om de taak te volbrengen.

7. *Draag op basis van collaterale informatie zelf relevante voorbeelden aan. Gaat de patiënt akkoord?* Bij geïnterviewden die maar moeilijk voorbeelden willen of kunnen geven, dient de interviewer alsnog te proberen herkenning van het criterium bij de patiënt op te roepen door hem of haar voorbeelden voor te leggen. Deze voorbeelden kunnen zijn ontleend aan de collaterale informatie. Hierbij dient de vraagstelling zoveel mogelijk open van aard te blijven.

8. *Is er voldaan aan de algemene diagnostische criteria?* Zodra twee of meer relevante voorbeelden zijn gegeven, dient geverifieerd te worden of voldaan is aan de algemene diagnostische criteria. Hieronder worden deze criteria achtereenvolgens besproken.

Volgens criterium A moet er sprake zijn van een duurzaam patroon van innerlijke ervaringen en gedragingen dat duidelijk afwijkt van de verwachtingen binnen de cultuur van betrokkene. Let wel: wat afwijkend is binnen een bepaalde (sub)culturele context kan heel goed 'normaal' zijn binnen een andere context. In veel landen kunnen bepaalde kenmerken, die hier bijvoorbeeld te boek staan als schizotypisch (bijvoorbeeld eigenaardige overtuigingen of magische denkbeelden), als heel normaal worden ervaren. En bij sommige beroepen of bedrijven wijst een tachtigjarige werkweek minder duidelijk in de richting van een overmatige toewijding aan werk en productiviteit dan bij andere beroepen of bedrijven. Ook dient men rekening te houden met de leeftijdsspecifieke en geslachtsspecifieke manifestaties van persoonlijkheidstreken.

Volgens criterium B dient de persoonlijkheidseigenschap star te zijn en zich te uiten op een breed terrein van persoonlijke en sociale situaties. Als een bepaald gedrag zich herhaaldelijk maar uitsluitend voordoet in een relatief specifieke context, dan is het onwaarschijnlijk dat dit gedrag verwijst naar een persoonlijkheidstrek. Om die reden dienen relevante voorbeelden van de criteria idealiter te worden gezocht in situaties van uiteenlopende aard.

Volgens criterium C moet het duurzame patroon in belangrijke mate resulteren in subjectief lijden of beperkingen in het sociaal en beroepsmatig functioneren of het functioneren op andere belangrijke terreinen. Bij sommige specifieke diagnostische criteria is op voorhand duidelijk dat er sprake is van lijdensdruk of functionele beperkingen (bijvoorbeeld affectabiliteit of buitensporige sociale angst), terwijl men zich er bij andere nadrukkelijk van dient te vergewissen of dit het geval is (bijvoorbeeld veel van de theatrale items).

Volgens criterium D dient het patroon stabiel en van lange duur te zijn en het begin ervan te kunnen worden teruggevoerd naar ten minste de adolescentie of de vroege volwassenheid. De DSM-IV laat bij dit criterium nogal wat ruimte

TABEL 2 Richtlijnen voor het hanteren van algemeen diagnostisch criterium D (temporele stabiliteit)

a. Personen met een leeftijd tot ongeveer 25 jaar	De adolescentie en vroege volwassenheid zijn bij uitstek levensfasen waarin iemand dramatisch kan veranderen. In veel gevallen kan daarom bij jongeren nog niet worden vastgesteld of er sprake is van een duurzaam patroon. Een goede vuistregel voor het stellen van As II-diagnoses bij jongeren is dat het gedrag ten minste drie jaar onafgebroken aanwezig dient te zijn. De code 799.9 (diagnose op As II uitgesteld) biedt een goed alternatief voor het stellen van een diagnose als er onvoldoende zekerheid bestaat.
b. Personen met langdurige hospitalisaties en/of detentieperioden	Bij personen die een groot deel van hun volwassen leven, of in elk geval van de laatste jaren, in een gecontroleerde setting hebben doorgebracht, dient men extra aandacht te besteden aan het duur-criterium. Voor een groot aantal diagnostische criteria geldt dat deze zich zelden duidelijk manifesteren in een dergelijke omgeving. Men dient in dit geval bij het navragen van relevante voorbeelden de nadruk te leggen op de perioden tussen verschillende hospitalisaties en/of detentieperioden, en alert te zijn op opmerkelijke verschillen in het gedrag bij verblijf binnen en buiten. Gedrag dat zich buiten consequent voordoet, maar zich vanwege langdurig verblijf in een inrichting de laatste jaren niet heeft gemanifesteerd, dient doorgaans toch te worden gescoord. Andersom geldt dat gedrag dat zich uitsluitend binnen manifesteert en vanwege langdurig verblijf in gecontroleerde settings in alle opzichten aan het duur-criterium voldoet, alleen mag worden gescoord indien er duidelijke aanwijzingen zijn dat betrokkene duurzaam is veranderd.
c. Personen met een leeftijd van 50 jaar en ouder	Bij ouderen dient men bedacht te zijn op opvallende gedragsveranderingen als gevolg van een bijzondere somatische of psychosociale conditie. Deze dienen in principe niet op As II te worden gediagnosticeerd. Duurzame persoonlijkheidsveranderingen als gevolg van een medische conditie kunnen onder code 310.1 worden gediagnosticeerd. Bij duurzame persoonlijkheidsveranderingen als gevolg van een psychosociale conditie kan men overwegen aan de diagnose op As II het predicaat 'in remissie' dan wel 'late aanvang' toe te voegen.

voor interpretatie over, met als gevolg dat we bij verschillende instrumenten een wildgroei van verschillende richtlijnen zien (Verheul & Van den Brink 1999). De vraag is of deze meer expliciete scoringsrichtlijnen altijd recht doen aan de werkelijkheid en, belangrijker nog, of zij iets toevoegen aan het DSM-IV-criterium (Widiger e.a. 1995). Wij stellen voor om criterium D te blijven gebruiken zonder te proberen dit, althans in algemene termen, verder te operationaliseren. Een uitzondering geldt wellicht voor criteria die betrekking hebben op gedragingen die maar heel weinig voorkomen, zoals suïcidaal gedrag of hardnekkig zoeken naar een andere relatie als een intieme relatie tot een eind komt. Bij dit soort criteria bestaat een groot risico op onderdiagnostiek wanneer het duurcriterium te letterlijk wordt genomen en een duurzame en hoogfre-

quente manifestatie als voorwaarde voor scoring wordt gehanteerd. Ten slotte behoeft de beoordeling van het stabiliteitscriterium in een aantal gevallen extra aandacht (zie tabel 2).

Volgens criterium E dient het patroon niet eerder te kunnen worden toegeschreven aan een uiting of een consequentie van een andere psychische stoornis, en volgens criterium F dient het patroon niet het gevolg te zijn van de directe fysiologische effecten van een middel (bijvoorbeeld drug, geneesmiddel) of een somatische aandoening (bijvoorbeeld schedeltrauma). In tabel 3 worden enkele globale regels gegeven voor de scoring van persoonlijkheidskenmerken bij patiënten die tevens lijden of hebben geleden aan een As I-stoornis of somatische aandoening.

TABEL 3 Richtlijnen voor het hanteren van algemene diagnostische criteria E en F

Relatie	Code	Overwegingen
1. Beloop van gedragspatroon valt samen met intermitterend beloop van andere conditie.	Geen diagnose op As II (v71.09)	Er is geen sprake van een blijvende persoonlijkheidsverandering.
2. Beloop van gedragspatroon valt samen met chronisch beloop van andere conditie.	Persoonlijkheidsstoornis (301.xx)	Hier is sprake van comorbiditeit. De aandoeningen op As I en As II kunnen niet worden ontrafeld.
3. Maladaptief gedragspatroon is begonnen tijdens andere conditie en is verdwenen binnen vijf jaar na remissie van andere conditie.	Geen diagnose op As II (v71.09)	Er is geen sprake van een blijvende persoonlijkheidsverandering.
4. Maladaptief gedragspatroon is begonnen tijdens andere conditie en duurt voort na remissie van andere conditie.	Diagnose op As II uitgesteld (799.9)	Indien remissie van andere aandoening minder dan vijf jaar geleden en/of indien er onvoldoende evidentie is voor een blijvende persoonlijkheidsverandering.
	Persoonlijkheidsstoornis (301.xx)	Indien het nu maladaptieve gedragspatroon reeds voor aanvang van de As II-stoornis in subklinische vorm aanwezig was.
	Persoonlijkheidsstoornis (301.xx), late aanvang	Indien premorbide persoonlijkheid duidelijk afwijkt van huidige kenmerken en indien andere aandoening psychosociaal van aard is.
	Persoonlijkheidsverandering door een somatische aandoening (310.1)	Indien premorbide persoonlijkheid duidelijk afwijkt van huidige kenmerken en indien andere aandoening somatisch van aard is.

9. *Afwezig volgens collaterale informatie (divergentie)*
Indien op grond van de gegeven voorbeelden aan het specifieke criterium lijkt te zijn voldaan en alle algemene criteria zijn geverifieerd, kan de collaterale informatie naast het patiëntoordeel worden gelegd. Bij divergerende bronnen van informatie kan men besluiten meer vragen te stellen aan de patiënt, dan wel om het criterium te scoren op basis van de meest betrouwbare en valide informatiebron.

EPILOOG

Betrouwbare en valide diagnostiek van per-

soonlijkheidsstoornissen vereist een passende opleiding (bijvoorbeeld die van psychiater of gezondheidszorgpsycholoog), aangevuld met uitgebreide en intensieve training in het afnemen van een semi-gestructureerd interview. Wat de haalbaarheid betreft, weten we uit eigen ervaring dat een goed getrainde diagnosticus in gemiddeld 60 minuten (range 30-90 minuten) de volledige procedure volgens de richtlijnen kan voltooien.

In het bovenstaande hebben we ons geconcentreerd op richtlijnen voor categoriale diagnostiek van persoonlijkheidsstoornissen. Het is echter nog maar de vraag of *categoriale* diagnostiek een goede representatie van de werkelijkheid

oplevert. Een groot aantal taxonometrische studies heeft laten zien dat er geen duidelijke grenzen bestaan tussen een normale persoonlijkheid en persoonlijkheidsstoornissen (Widiger & Costa 1994). Bovendien levert een categoriale benadering een aantal classificatiedilemma's op, bijvoorbeeld wanneer een patiënt twee of meer persoonlijkheidsstoornissen heeft, en bij klinisch betekenisvolle heterogeniteit binnen categorieën. Steeds meer onderzoekers en klinici pleiten daarom voor een gefaseerde diagnostische procedure waarbij voor communicatieve doeleinden in de eerste fase gebruikgemaakt wordt van de verschillende diagnostische categorieën, terwijl de tweede fase een (multi)dimensionale beschrijving of kwantificering van de psychologisch- en biologisch-disfunctionele componenten van de pathologie omvat (Van Praag 1990; Verheul 1998). Bij een borderline persoonlijkheidsstoornis omvat de tweede trede bijvoorbeeld een precieze beschrijving of kwantificering van stoornissen in de impulscontrole, de regulatie van stemming en agressie, de identiteitsintegratie en het cognitief-perceptueel functioneren (Verheul 1998). Deze beschrijving is relevant voor een op de specifieke problemen van de individuele patiënt toegespitste indicatiestelling en/of behandelplanning.

LITERATUUR

- Akkerhuis, G.W., Kupka, R.W., Groenestijn, M.A.C., e.a. (1996). *PDQ-4+ Vragenlijst voor Persoonlijkheidskenmerken*. Lisse: Swets & Zeitlinger.
- Blashfield, R.K., & Herkov, M.J. (1996). Investigating clinician adherence to diagnosis by criteria: A replication of Morey and Ochoa (1989). *Journal of Personality Disorders*, 10, 219-228.
- Diekstra, R.F.W., Duijsens, I.J., & Eurelings-Bontekoe, E.H.M. (1996). *International Personality Disorder Examination (IPDE)*, Ned. vertaling. Lisse: Swets & Zeitlinger.
- Duijsens, I.J., Eurelings-Bontekoe, E.H.M., Diekstra, R.F.W., e.a. (1993). *Dutch VKP. Vragenlijst voor kenmerken van de persoonlijkheid. Voorlopige handleiding*. Lisse: Swets & Zeitlinger.
- Duijsens, I.J. (1996). *Assessment of personality disorders: Construction, reliability and validity of the VKP self-report* (academisch proefschrift). Lisse: Swets & Zeitlinger.
- Jong, C.A.J. de, Derks, F.C.A., Oel, C.J. van, e.a. (1996). *Gestructureerd interview voor de DSM-IV persoonlijkheidsstoornissen (SIDP-IV)*. NOVADIC-publicatie.
- Morey, L.C., & Ochoa, E.S. (1989). An investigation of adherence to diagnostic criteria: Clinical diagnosis of the DSM-III personality disorders. *Journal of Personality Disorders*, 3, 180-192.
- Pilkonis, P.A., Heape, C.L., Ruddy, J., e.a. (1991). Validity in the diagnosis of personality disorders: The use of the LEAD-standard. *Psychological Assessment*, 3, 46-54.
- Praag, H.M. van (1990). Two-tier diagnosing in psychiatry. *Psychiatry Research*, 34, 1-11.
- Riso, L.P., Klein, D.N., Anderson, R.L., e.a. (1994). Concordance between patients and informants on the Personality Disorder Examination. *American Journal of Psychiatry*, 151, 568-573.
- Schotte, C., & Doncker, D. de (1994). *Assessment DSM Persoonlijkheidsstoornissen - versie DSM-IV (ADP-IV)*. Niet uitgegeven vragenlijst, Universitair Ziekenhuis Antwerpen.
- Skodol, A.E., Rosnick, L., Kellman, D., e.a. (1988). Validating structured DSM-III-R personality disorder assessments with longitudinal data. *American Journal of Psychiatry*, 145, 1297-1299.
- Spitzer, R.L. (1983). Psychiatric diagnosis: Are clinicians still necessary? *Comprehensive Psychiatry*, 24, 399-411.
- Spitzer, R.L., & Fleiss, F.L. (1974). A re-analysis of the reliability of psychiatric diagnosis. *British Journal of Psychiatry*, 125, 341-347.
- Verheul, R. (1998). *Classificatie en differentiële diagnostiek*. In W. van Tilburg, W. van den Brink & A. Arntz (red.), *Behandelingsstrategieën bij de borderline persoonlijkheidsstoornis*. Houten/Diegem: Bohn Stafleu Van Loghum.
- Verheul, R., & van den Brink, W. (1999). *Persoonlijkheidsstoornissen*. In J. Ormel, W. van den Brink & D. Wiersma (red.), *Leerboek Psychiatrische Epidemiologie*. Utrecht: Uitgeverij De Tijdstroom BV.
- Weertman, A., Arntz, A., & Kerkhofs (1996). *Gestructureerd klinisch interview voor DSM-IV persoonlijkheidsstoornissen* (Ned. bewerking van First e.a. SCID-II, version 2.0, 1996). Maastricht: Universiteit Maastricht, Vakgroep Medische Psychologie, Maastricht.
- Westen, D. (1997). Divergences between clinical and research methods for assessing personality disorders: Implications for research and the evolution of axis II. *American Journal of Psychiatry*, 154, 895-903.
- Widiger, T.A., & Costa, P.T. (1994). Personality and personality disorders. *Journal of Abnormal Psychology*, 103, 78-91.
- Widiger, T.A., Mangine, S., Corbitt, E.M., e.a. (1995). *Personality*

Disorder Interview-IV: A semistructured interview for the assessment of personality disorders. Odessa, Florida USA: Psychological Assessment Resources.

Zimmerman, M. (1994). Diagnosing personality disorders: A review of issues and research methods. *Archives of General Psychiatry*, 51, 225-245.

AUTEURS

R. VERHEUL, psycholoog, is als senioronderzoeker verbonden aan het Amsterdam Institute for Addiction Research (AIAR) en de vakgroep Psychiatrie van het Academisch Medisch Centrum/Universiteit van Amsterdam.

Correspondentieadres: AIAR, Jacob Obrechtstraat 92, 1071 KR Amsterdam, e-mail: verheul@aiar.nl.

W. VAN DEN BRINK, arts-epidemioloog, is directeur van het Amsterdam Institute for Addiction Research (AIAR), en

als hoogleraar verbonden aan de vakgroep Psychiatrie van het Academisch Medisch Centrum/Universiteit van Amsterdam. Correspondentieadres: Psychiatrisch Centrum AMC, Tafelbergweg 25, 1105 BC Amsterdam.

PH. SPINHOVEN, klinisch psycholoog/psychotherapeut, is als hoogleraar verbonden aan de sectie Klinische en Gezondheidspsychologie van de Universiteit Leiden.

Correspondentieadres: Pieter de la Court Gebouw, Postbus 9555, 2300 RB Leiden.

R. HARINGSMA, psycholoog, is als gastmedewerker verbonden aan de sectie Klinische en Gezondheidspsychologie van de Universiteit Leiden.

Correspondentieadres: Pieter de la Court Gebouw, Postbus 9555, 2300 RB Leiden.

Het artikel werd voor publicatie geaccepteerd op 12-7-1999.

SUMMARY Guidelines for clinical assessment of DSM-IV personality disorders – R. Verheul, W. van den Brink, Ph. Spinhoven, R. Haringsma –

BACKGROUND In general, clinical diagnoses of personality disorders lack sufficient reliability and validity. There is currently no consensus about the most appropriate method for the clinical assessment of DSM-IV personality disorders.

AIMS Guidelines for the clinical assessment of DSM-IV personality disorders are proposed.

METHOD This systematic review of procedures for diagnosing personality disorders is based upon three criteria: reliability, validity and feasibility. The existing assessment methods are evaluated against these three criteria and, subsequently, guidelines are developed by aiming at finding an optimal balance in the extent to which the criteria are met.

RESULTS The reliability of personality disorder assessment benefits substantially from strategies towards structuring the clinical interview as well as standardizing the scoring of the diagnostic criteria. A semi-structured interview schedule fits in optimally with these requirements. To improve validity it is recommended to have available biographical data, collateral information, and an adequate operationalization of both the general and specific diagnostic criteria of Axis I. It is advocated to enhance the feasibility through generating hypotheses based upon self-report screening, clinical impressions and (if necessary) collateral information.

CONCLUSIONS Reliable and valid assessment of DSM-IV personality disorders is a complex task that requires both sufficient knowledge of DSM-IV and intensive training in the administration of a semistructured interview schedule.

[TIJDSCHRIFT VOOR PSYCHIATRIE 42 (2000) 6, 409-422]

KEYWORDS diagnosis, guidelines, personality disorders