

ENKELE PROBLEMEN BIJ HET SAMENSTELLEN VAN PROEFGROEPEN VOOR WETENSCHAPPELIJK ONDERZOEK IN HET PSYCHIATRISCH ZIEKENHUIS

door Dr. A. A. FISCHER

Zoals u weet is de Tweede Psychiatrische Universiteitskliniek in het leven geroepen door de Willem Arntsz Stichting, om aan een staf van universitaire psychiatrische en klinisch psychologische werkers de mogelijkheid te bieden binnen haar muren wetenschappelijk onderzoek te verrichten.

Deze kliniek bleek de optimale mogelijkheid te verschaffen om patiënten zonodig langere tijd op te nemen en onder controleerbare omstandigheden, eventueel modificeerbare omstandigheden, te onderzoeken, te verplegen en te behandelen. Met andere woorden: deze kliniek biedt de mogelijkheden om naast het klinisch empirisch onderzoek ook het geleide experiment een plaats te geven.

Onvermijdelijk drong zich de vraag op, welke patiënten hier opgenomen zouden moeten worden.

Het was afgesproken dat het patiënten lijdende aan schizofrenie zouden zijn.

Maar op grond van welke kenmerken moest hier de diagnose schizofrenie gesteld worden?

En verder: moest iedere patiënt, die volgens deze kenmerken schizofreen genoemd kon worden, ook *opgenomen* worden of moesten hier nog verdere restricties gelden?

Brengt men verdere restricties aan, dan kan de vraag gesteld worden: waartoe?

Waarom wordt de ene schizofrene patiënt wel opgenomen en de andere niet?

U kunt zich voorstellen, dat we langzaam maar zeker steeds verder gedwongen werden in de problemen, die inhaerent zijn aan het samenstellen van proefgroepen.

Problemen dus, die misschien op het eerste gezicht nauwelijks met het onderzoeksgebied van de schizofrenie lijken samen te hangen, maar die onvermijdelijk blijken te zijn zodra men in het wetenschappelijk onderzoek de mogelijkheden van de experimentele methode wil benutten.

Het is speciaal voor een aantal *praktische* problemen, die zich voordoen wanneer men een proefgroep wil formeren, dat ik uw aandacht vraag.

Ik laat daarbij buiten beschouwing de statistische problemen die onvermijdelijk optreden wanneer men zich met het samenstellen van proefgroepen gaat bezig houden. Ik laat hier ook buiten be-

schouwing het complexe probleem van de samenhang tussen doelstellingen van een onderzoek, de theoretische concepten die eraan ten grondslag liggen en de criteria op grond waarvan men een proefgroep samenstelt. Daarover graag een andere keer. Ik wil u nu een paar praktische problemen beschrijven aan de hand van onze ervaringen in de Tweede Psychiatrische Universiteitskliniek. Of wetenschappelijk onderzoek relevant is hangt in de eerste plaats af van het feit of objectieve meetbaarheid verkregen is, of ook objectieve metingen verricht zijn en of de juiste controle is uitgevoerd.

Maar van even groot belang is of op de juiste wijze proefgroepen zijn samengesteld.

Voor een verificatie is het essentieel, dat andere onderzoekers in staat moeten zijn om het experiment te herhalen met dezelfde of met andere methoden.

Bij een goed experiment wordt juist ook op dit laatste criterium gelet.

Collega Werlemann is mij behulpzaam geweest bij het nazoeken van de jaargang 1963 van de *British Journal of Psychiatry* de jaargang 1963—'64 van de *Archives of General Psychiatry* (amerikaans)

en de jaargang 1964 van de *Nervenarzt*.

Deze jaargangen zijn willekeurig gekozen.

In deze jaargangen hebben wij de artikelen opgezocht, die de schizofrenie betroffen.

Uit deze artikelen hebben we dezulke gekozen, die een systematisch onderzoek betroffen.

En weer uit deze artikelen hebben we dat gezocht waarin de opvattingen van de onderzoeker over de schizofrenie beschreven waren.

In de *British Journal of Psychiatry* 1963 waren 21 artikelen over schizofrenie gepubliceerd, waarvan 16 een systematisch onderzoek betroffen, waarin bij 4 de opvattingen van de onderzoeker over schizofrenie beschreven waren.

In de *Archives of General Psychiatry* 1963—'64 waren 26 artikelen over schizofrenie gepubliceerd waarvan 22 systematische onderzoeken betroffen en waarbij in 1 van deze artikelen de opvattingen van de onderzoeker over de schizofrenie beschreven waren.

In de *Nervenarzt* van 1964 waren 3 artikelen over schizofrenie. Geen van de 3 betrof een systematisch onderzoek.

Slechts in 1 van al deze artikelen waren de kenmerken van het patiëntenmateriaal zodanig gesteld, dat op grond hiervan een herhaling van het onderzoek mogelijk zou zijn.

Onze kennis, ons gedrag en ons handelen zijn voor een zeer groot deel gebaseerd op steekproeven. Dat geldt zowel voor het dagelijks leven als voor wetenschappelijk onderzoek. De mening die iemand heeft over de kwaliteit van een bepaald merk auto, of over de service van een bepaald bedrijf, is meestal bepaald door enkele ervaringen in de loop van vele jaren. De toerist die op de kaasmarkt in Alkmaar een paar hapjes kaas mag nemen en op grond daarvan zich een mening vormt over de daar verkochte kaas, verschilt alleen van de keurmeester met zijn jarenlange ervaring, doordat hij zijn conclusies over de daar verkochte kaas slechts baseert op een kleinere hoeveelheid ervaring en hij zal zich waarschijnlijk minder realiseren in welke mate hij zich vergist. Onze term „steekproef” komt natuurlijk „uit de kaas”.

Maar algemener: in elke tak van wetenschap missen we de mogelijkheden om meer dan een fragment te bestuderen van de verschijnselen waarbij wij onze kennis zouden kunnen vergroten.

Tot voor kort werd relatief weinig aandacht besteed aan het probleem hoe men een goede steekproef moet nemen.

Dit doet weinig ter zake zolang het materiaal waar we de steekproeven uit nemen uniform is, zodat iedere steekproef ongeveer hetzelfde resultaat geeft.

In het laboratorium wordt vaak beslist of iemand gezond of ziek is en aan wat voor eventuele ernstige ziekte hij lijdt, op grond van de beoordeling van een paar druppels bloed.

Dit is gebaseerd op de vooronderstelling, dat het circulerende bloed altijd goed gemengd is en dat de ene druppel bloed ons dezelfde gegevens zal verschaffen als de andere.

Maar wanneer het materiaal helemaal niet eenvormig is, zoals ook in de psychiatrie, wordt de wijze waarop een proefgroep is samengesteld en de methode die daarvoor gebruikt is, van essentieel belang.

Geen twee mensen zijn gelijk en geen twee patiënten, lijdende aan schizofrenie, zijn gelijk.

Op grond waarvan beslissen wij nu, dat een X aantal mensen tot een groep behoort? Hoe stellen wij deze groep samen?

Er bestaan verschillende manieren om proefgroepen samen te stellen.

Welke vorm van proefgroep kunnen wij het beste kiezen om met de beschikbare mogelijkheden een haalbaar doel te bereiken?

Laten we proberen om aan de hand van het schizofrenie-probleem, zoals wij dat in de Tweede P.U.K. hebben ontmoet, een aantal fundamentele problemen te belichten. Wanneer ik me niet vergis gelden deze problemen net zo voor vrijwel elk ander onderzoeksgebied in de psychiatrie.

Met de term „populatie” wordt hier aangegeven het totaal van alle gevallen, alle individuen, die beantwoorden aan bepaalde gestelde kenmerken. Zo zouden we door de specificaties „mensen” en „opgenomen in een psychiatrisch ziekenhuis” een populatie kunnen definiëren die bestaat uit alle gehospitaliseerde psychiatrische patiënten. Op een soortgelijke wijze zouden wij populaties kunnen definiëren, die bijv. bestaan uit alle patiënten die behandeld worden met psychofarmaca of alle patienten bij wie de ziekte langer duurt dan een bepaalde tijd, etc.

Door bepaalde specificaties kan een populatie deel uitmaken van een andere. Zo maakt de populatie van schizofrene patiënten, die opgenomen zijn in psychiatrische ziekenhuizen, deel uit van de populatie van alle gehospitaliseerde psychiatrische patiënten. In zo'n geval spreken we van een „sub-populatie” of van een „stratum”. Een stratum wordt bepaald door een of meer specificaties, die een populatie onderverdeelt in mutueel exclusieve delen. Zo kan bijv. een populatie van gehospitaliseerde schizofrene patiënten door het aannemen van leeftijdscategorieën tot 20 jaar, tussen 20 en 35 en boven de 35 jaar in drie verschillende strata onderverdeeld worden. Nemen we nu bijv. alle schizofrene patiënten, die opgenomen zijn in een psychiatrisch ziekenhuis en van wie de leeftijd ligt tussen de 20 en 35, dan mogen we ook in plaats van een stratum van een populatie spreken waarbij dan

1. het feit dat dit mensen zijn,
 2. die in een psychiatrisch ziekenhuis zijn opgenomen en
 3. die lijdende zijn aan schizofrenie en
 4. die een leeftijd hebben tussen de 20 en 35 jaar
- de kenmerken zijn, die deze populatie bepalen.

Het is echter in de regel ondoenlijk — om tal van praktische redenen — om alle patiënten, behorende bij zo'n populatie, te onderzoeken. Wanneer we nu een aantal patiënten selecteren met de bedoeling om iets te onderzoeken over de populatie waar zij deel van uitmaken, spreken we van een „proefgroep” of een „sample”. We hopen daarbij natuurlijk, dat wat we zullen vinden bij deze proefgroep ook op zal gaan voor de gehele populatie. Met andere woorden; wij hopen dat wat we bij deze groep van schizofrene patiënten zullen vinden ook opgaat voor de andere gehospitaliseerde schizofrene patiënten in Nederland tussen de 20 en 35 jaar. Het blijft echter de vraag of dit het geval is. In hoeverre de informatie die wij op deze wijze krijgen overeenkomt met wat we zouden vinden wanneer we alle patiënten van de populatie zouden onderzoeken, hangt voor een groot deel af van de wijze waarop de proefgroep is geselecteerd.

Zo zouden we misschien graag willen weten hoeveel patiënten van deze populatie (dus van alle gehospitaliseerde schizofrene

patiënten) een nieuwe aanpassing aan het leven buiten het ziekenhuis zouden kunnen vinden. Maar het aantal mensen uit de proefgroep wie dit lukt hoeft in het geheel niet in verhouding te staan tot hun overeenkomstig aantal in de populatie. Het zou best mogelijk kunnen zijn dat in de proefgroep een onevenredig aantal patiënten met concentratiestoornissen is opgenomen, waardoor ze eenvoudig hun weg in het verkeer niet kunnen vinden.

Of: het is mogelijk dat we op zoek naar leverfunctiestoornissen ongemerkt een aantal langdurig gehospitaliseerde patiënten in onze proefgroep zouden hebben opgenomen, die in de oorlog door slechte voedingstoestanden leverbeschadigingen hadden gekregen, met alle gevolgen van dien voor de uitkomsten van het onderzoek. Het is heel moeilijk om uit te vinden in hoeverre een proefgroep op juiste wijze de populatie representeert. We kunnen nooit garanderen dat de uitkomsten door middel van een proefgroep verkregen, opgaan voor de populatie, tenzij we een volledig vergelijkbaar census-onderzoek doen. We kunnen echter „proefgroep plannen” bedenken, die, wanneer ze goed worden uitgevoerd, kunnen garanderen dat, indien we een onderzoek met een aantal verschillende proefgroepen, geselecteerd uit een bepaalde populatie, zouden herhalen, onze bevindingen niet méér van de echte populatiepercentages zouden verschillen dan een bepaald percentage.

Hoe we een proefgroep samenstellen en hoe het plan voor het onderzoek moet zijn, hangt nauw samen met ons aspiratieniveau wat betreft de nauwkeurigheid en de betekenis van de uitkomst die we willen vinden. (Afgezien dus nog van de praktische uitvoerbaarheid waarover we nog komen te spreken).

Wanneer we met proefgroepen gaan werken is dus een vereiste, dat er een „representatief plan” is (representative sampling plan) om te verzekeren dat de kans groot genoeg is, dat de geselecteerde proefgroep, met het oog op het voorgenomen doel, voldoende representatief is t.a.v. de populatie. Het gebruik van zo’n representatief plan levert niet alle mogelijke verzekering, dat onze uitkomsten niet al te misleidend zullen zijn.

Er zijn ook andere garanties nodig, zoals bijv. het vereiste dat een zekere verscheidenheid van patiënten (diverse elementen uit de populatie) opgenomen zijn.

We moeten dan ook waar maken, dat met de aantallen van deze verschillende soorten patiënten rekening is gehouden met de aantallen waarin ze voorkomen in de totale populatie.

Wanneer we hier nauwkeurig op letten (waarbij dan eigenlijk ook nog hoort dat we nauwkeurig rekening moeten houden met de te volgen meetmethode) kan een proefgroep vaak betere resultaten geven dan een grootscheeps onderzoek.

Wij beschikken eenvoudig niet over voldoende mensen en outillage, met de benodigde speciale capaciteiten, om in een dergelijk *census*-onderzoek ook maar de meest oppervlakkige informatie te krijgen. Een onderzoek op *kleinere* schaal maakt het praktisch mogelijk meer tijd te besteden per patiënt, en dus ook meer en diepgaander informatie te krijgen, terwijl de betrouwbaarheid en waarde wat betreft de geldigheid voor de gehele populatie net zo groot kan zijn, mits een „representatief plan”, een goede samenstelling van de proefgroep en goede meetmethoden worden gebruikt.

Er wordt in de moderne theorieën over proefgroepformatie een essentieel verschil gemaakt tussen „probability sampling” en „non-probability sampling”.

Hiermee wordt een fundamenteel probleem aangegeven, wat ook belangrijk is wanneer wij bijv. proberen om een proefgroep van schizofrene patiënten samen te stellen.

Het bepalende kenmerk van „probability sampling” is, dat men voor ieder van de elementen (patiënten) van een populatie de waarschijnlijkheid kan berekenen dat het in een proefgroep zal worden opgenomen.

In het eenvoudigste geval moet dan *iedere* schizofrene patiënt, die opgenomen is in een psychiatrisch ziekenhuis in Nederland, dezelfde kans hebben om opgenomen te worden in onze proefgroep. Om allerlei praktische en organisatorische redenen is dit in feite niet zo.

Maar dit is ook niet noodzakelijk.

Wel is noodzakelijk, willen wij tenminste een „probability sample” samenstellen, dat er *enige*, maar dan ook specificieerbare waarschijnlijkheid is, dat een patiënt tot de proefgroep *kán* behoren. Pas dan zijn wij in staat om aan te geven hoe groot de kans is dat onze bevindingen, gevonden bij de proefgroep, niet meer dan een bepaald percentage van de echte populatiewaarde af liggen. Tegelijkertijd bestaat er dan ook een garantie, dat voldoende gevallen zijn geselecteerd uit ieder stratum, iedere laag van de populatie, dat wil zeggen ieder stratum dat ook relevant is voor ons onderzoek.

Wegens tijdnoed en ook omdat ik u niet wil vermoeien met statistische problemen, laat ik hier verder achterwege op welke wijze te berekenen is, hoe groot de kans is, dat de bevindingen die bij zo'n proefgroep gevonden worden, overeenkomen met de bevindingen wanneer men de hele populatie zou onderzoeken.

We spraken eerder van een „representatief plan” voor het samenstellen van een proefgroep.

Het probleem is nu, dat deze „probability sampling” de enige methode is die dit mogelijk maakt. Omdat dit de enige methode is waarbij de populatie en de proefgroep in bekende verhoudingen ten opzichte van elkaar staan. Het is de enige methode waarmee een onderzoeker kan vaststellen in welke mate zijn bevindingen, gebaseerd op onderzoek van de proefgroep, overeenkomen met of afwijken van de populatiegegevens.

Het is echter niet moeilijk om vast te stellen, dat deze methode van „probability sampling” voor ons schizofrenie-onderzoek onbruikbaar is.

Het is nog wel mogelijk om te weten te komen hoeveel patiënten van een bepaalde leeftijdsgroep onder de diagnose schizofrenie in de Nederlandse ziekenhuizen verpleegd worden.

Er is ook nog wel statistisch uit te komen, dat deze populatie niet constant is en zelfs per dag sterk kan veranderen.

Door de moderne behandelingsmethoden wisselt de ziekenhuisbevolking immers tegenwoordig snel.

Het grote struikelblok vormt de diagnose schizofrenie zelf. Dit probleem is ons maar al te zeer bekend.

K. Conrad schrijft in 1958 over het probleem van de diagnostiek in de psychiatrie:

„Der Fall eines Beziehungswahns bei Verstimmung nach einer Entlobung bei einer Frau mit einer Struma würde vielleicht — ich übertreibe hier absichtlich in unerlaubter Weise — in Göttingen als beginnender schizophrener Schub diagnostiziert, während er in Tübingen mehrdimensional als sensitiver Beziehungswahn bei schizoider Konstitution und Basedowoid aufgefasst würde, in Heidelberg vielleicht als Untergrunddepression, in Berlin Ost als „affektvolle Paraphrenie” im Sinne einer scharf definierbaren heredodegenerativen Einheit, in Zürich als endokrine Psychose bei Schilddrüsenerkrankung, in Bonn als paranoid gefärbte endoreaktive Dysthymie, in Hamburg als cyclothyme Depression mit paranoiden Wahneinlagen und in Frankfurt als Folge einer gestörten Daseinsordnung, also als eine Form des Scheiterns auf dem Lebensweg.”

En dit geldt dan alleen nog maar voor het duitse taalgebied. Naar het lijkt zijn deze tegenstellingen in ons land niet zo groot en zeker niet zo bloemrijk verwoord, maar ze bestaan. Hoe kunnen we dan in deze methodische zin over een populatie spreken en hoe zouden we een proefgroep kunnen samenstellen als er zoveel, meestal onuitgesproken, verschil van mening bestaat over welke patiënt schizofreen genoemd moet worden en welke niet?

We moeten hier nadrukkelijk twee problemen niet met elkaar verwarren. Het bestaan van het bovengenoemde probleem wil in het geheel niet zeggen, dat de diagnose schizofrenie in de dage-

lijkse praktijk van de klinische psychiatrie geen bruikbaarheid zou hebben. Iedere ervaren onderzoeker weet precies het „zeer speciale schizofrene” te herkennen.

Dit wil ook alweer niet zeggen, dat er geen twijfelgevallen zijn, maar we kennen allemaal de patiënten waarbij we op grond van een bepaalde evidentie, een klinische evidentie, met grote waarschijnlijkheid kunnen zeggen, dat we hier met een schizofrene patiënt te maken hebben.

Het is echter een evidentie, het is (nog) niet onder woorden te brengen. Rümke spreekt over „het zeer bepaalde in een veranderde relatie tot de werkelijkheid”.

Duitse schrijvers spreken vaak van het „ganz Eigentümliche” en engelse schrijvers van „very specific”.

Het probleem is echter, dat dit niet te verbaliseren is tot nu toe. Hier verkeren we wat dat betreft in een stadium van de wetenschap, dat vergelijkbaar is met een fase van de vroegere fysica, waarbij de onderzoeker in staat was zeer nauwkeurig toon- en klank-onderscheiden aan te geven, maar nog niet over de mogelijkheid beschikte om dit in frequentie en amplitudo, in maat en getal, op een voor iedereen eenduidige wijze aan te geven. Het is de vraag of we dit in het onderzoeksgebied van de psychiatrie ooit zullen bereiken, maar ons zoeken is hier wel op gericht.

Voor ons onderzoek met proefgroepen hebben we hier echter weinig aan. We moeten een populatie en een proefgroep kunnen definiëren aan de hand van voor iedereen eenduidige kenmerken. Dit „zeer bepaalde” in de evidentie van de onderzoeker, wat niet te verbaliseren is, is daarvoor niet te gebruiken.

Tenzij op de volgende wijze:

Theoretisch zou ons nog één weg openstaan om deze moeilijkheid te ondervangen.

Dat is het onderzoek met behulp van beoordelaars.

Het is mogelijk om een onderzoek te doen met een proefgroep van elementen, die op grond van onafhankelijke beoordelingen van verschillende ervaren specialisten, als behorende tot een categorie gerekend kunnen worden. We missen dan wel de eerder genoemde kenmerken, maar we kunnen dan wel het houvast hebben, dat de patiënten op grond van een aantal overeenstemmende evidenties tot een groep behoren, wanneer tenminste het beoordelingsonderzoek op de methodisch juiste wijze wordt uitgevoerd.

In ons land publiceerde Dijkhuis over het beoordelingsprobleem in verband met het psychologisch onderzoek. Kortgeleden promoveerde Beyaert op een gedragsbeoordelingsonderzoek, verricht in de Tweede Psychiatrische Universiteitskliniek.

Hier zouden we dus mensen op een diagnose moeten laten beoordelen.

Maar ook nu stuiten we weer op moeilijkheden.

Deze beoordelingsonderzoeken, die complex en moeilijk zijn, zijn bij psychiatrische patiënten al meerdere malen verricht. Een goed overzicht hiervan vindt u in het genoemde proefschrift van Beyaert.

Volgens Kreitman blijkt nu, dat de procentuele overeenstemmingen tussen beoordelaars als volgt zijn:

voor de diagnose functional psychosis 31% tot maximaal 66%

voor de diagnose psychoneurosis 12% tot maximaal 24%

voor de diagnose character disorder $\pm$ 46%

voor de diagnose organic disorder 53% tot maximaal 74%

Wanneer nu bij dergelijke zeer globale categorieën slechts een zo lage overeenkomst wordt bereikt, wat moeten we dan verwachten voor de diagnose schizofrenie?

Met andere woorden: van de methode van het beoordelingsonderzoek kunnen wij voor het samenstellen van een proefgroep van patiënten (onder de diagnose schizofrenie) ook niet veel verwachten.

Zover gekomen lijkt de conclusie duidelijk.

Om in een onderzoek met proefgroepen tot uitspraken te kunnen komen, die met enige waarschijnlijkheid voor een populatie of een bepaalde laag daarvan gelden, waarbij dus ook hoort dat bepaald moet kunnen worden hoe groot deze waarschijnlijkheid is, moet de populatie (of het stratum) waarop de proefgroep wordt getrokken, bepaald zijn door eenduidige kenmerken.

En verder moet ook voor de proefgroep bepaald kunnen worden of grootte en kenmerken representatief zijn.

Doordat we niet weten wat deze objectieve kenmerken zijn voor patiënten, lijdende aan schizofrenie, kunnen we ook niet van een populatie spreken; laat staan dat we kunnen zeggen welke grootte deze populatie heeft.

Het is dus ook niet mogelijk een proefgroep samen te stellen, die aan genoemde eisen voldoet.

Conclusie: het gebruik van proefgroepen bij het onderzoek van patiënten, lijdende aan schizofrenie, is zinloos.

Helaas is echter dit probleem niet zo eenvoudig af te doen.

Waarom niet?

Zonder in problemen te treden als bijv. die van de polariteit tussen fenomenologie en natuurwetenschappen of die van de tegenstelling tussen ideografisch en nomothetisch, kunnen we stellen dat het eigen is aan ons wetenschappelijk werk, dat we zoeken naar wetmatigheden.

In ons wetenschappelijk werk zoeken wij naar overeenkomsten. Wat ik bij de ene patiënt zie, vind ik dat ook bij de andere? Wij

gaan als vanzelfsprekend uit van het idee, dat wat ons bij de ene patiënt opvalt net zo bij andere voor kan komen.

Met een open oog voor het nieuwe proberen wij bij patiënten te vinden, te herkennen, wat wij al kennen.

Wij hebben de verwachting van de herhaling van de gebeurtenissen. Wij praten ook over *de* schizofrenie, *de* manisch depressieve psychose, etc. Met andere woorden: dit komt niet maar eenmaal voor.

Zo veronderstellen we ook op grond van onze persoonlijke ervaring en de ervaring van anderen, die ons overgedragen is, dat *de* denkstoornissen, *de* affectstoornissen etc. bij de schizofrenie regelmatig voorkomen.

We veronderstellen de steeds terugkerende gebeurtenissen.

Of anders gesteld: wij veronderstellen dat psychiatrische verschijnselen in wetmatigheden te vatten zijn. Rümke stelt ook terecht, dat het schizofrene onindividueel is. Alleen al op grond van wetenschapstheoretische vooronderstellingen moeten wij dit zonder meer aannemen.

Wij zijn gewend om individuele patiënten te onderzoeken.

Iets bepaalds treft ons daarbij, meestal op grond van theoretische concepten, die weinig expliciet zijn.

In een X aantal (meestal maar klein aantal) andere patiënten menen wij hetzelfde te herkennen en nemen al spoedig een mate van algemeenheid aan, die moeilijk is aan te tonen.

In feite doen wij dan niets anders dan een soort van proefgroep samenstellen, waarop wij algemenere wetmatigheden postuleren. Het verschil is echter, dat het hierbij vrijwel altijd gaat om een ongecontroleerd en verder oncontroleerbaar experiment op grond van overwegingen en onder omstandigheden, die weinig of niet expliciet gemaakt zijn.

Willen we werkelijk tot geldende en bruikbare wetmatigheden komen, dan zullen we wel met proefgroepen moeten gaan werken. Dat wil zeggen: we zullen wel *moeten* gaan werken met groepen patiënten, waarvan wij zover als maar enigszins mogelijk is moeten specificeren waarom wij denken, dat zij in één groep thuishoren.

En dan op zodanige wijze specificeren, dat een ander onderzoeker in staat is eenzelfde proefgroep samen te stellen en het onderzoek te herhalen.

Aanvaarden we deze konsekwentie niet, dan moeten wij ons postulaat van wetmatigheden in het veld van de psychiatrie loslaten. Maar dan mogen we ook niet spreken van *de* schizofrenie, *de* schizofrene denkstoornissen etc.

Dit zijn allemaal termen die wetmatigheden veronderstellen, op grond waarvan wij gewend zijn regels te hanteren voor therapie,

prognose etc. Diagnostiek, ongeacht of dit is in termen van symptomen, syndromen of ziektebeelden, bestaat dan niet meer.

We moeten dus wel met proefgroepen werken in het wetenschappelijk onderzoek in de psychiatrie en van de schizofrenie in ons geval. Maar hoe?

Er bestaat nog een andere mogelijke methode voor het samenstellen van proefgroepen, maar daarbij zullen wij onze eisen veel minder hoog moeten stellen.

Dat is de methode van de „non-probability sampling”, die ik u reeds terloops noemde.

Bij „non-probability sampling” is, zoals het woord aangeeft, geen enkele waarschijnlijkheid te bepalen, die iedere patiënt van de populatie heeft om in de proefgroep te komen.

Dus ook niet of hij enige kans heeft.

Bij de „probability sampling” was de *populatie* bepaald door een aantal eenduidige kenmerken.

Doordat de kenmerken van de elementen van de proefgroep in overeenstemming waren met de populatie kenmerken, kon een waarschijnlijkheidsberekening gemaakt worden.

Bij de „non-probability sampling” bestaat geen populatie, die beantwoordt aan eenduidige kenmerken.

Een populatie in deze zin kan alleen worden vermoed.

Wel blijft echter de eis voor de proefgroep, dat deze bepaald moet zijn door een aantal eenduidige kenmerken, omdat anders geen herhaling van het onderzoek met een gelijke proefgroep kan plaatsvinden, en nooit een uitbreiding van het onderzoek tot een populatie-onderzoek mogelijk is.

Door het ontbreken van de populatie kenmerken is nu echter geen waarschijnlijkheidsberekening mogelijk.

Nemen wij ons onderzoek in de Tweede Psychiatrische Universiteitskliniek nu weer als voorbeeld:

De populatie van patiënten, lijdende aan schizofrenie, bestaat niet wanneer wij de eis van eenduidige kenmerken stellen.

Wij beschikken niet over populatie kenmerken.

Onze proefgroep zal echter wel door eenduidige kenmerken bepaald moeten zijn.

Wij zijn daarbij van de overweging uitgegaan, dat schizofrenie als verzamelbegrip voor de hebefrenie, katatonie, dementia paranoïdes en misschien ook parafrenie, zeker niet verder te onderzoeken is.

Afgezien van het academisch probleem of deze beelden wel bijeen horen, is een dergelijke ordening te heterogeen om er duidelijke kenmerken voor te vinden.

Het beeld dat het duidelijkst aanwijsbaar is, het meest voorkomt

en — niet toevallig — in de literatuur het best en het meest beschreven is, is dat van de hebefrenie.

Besloten werd dus om speciaal op deze categorie de aandacht te richten.

Met welke eenduidige kenmerken is nu deze groep te omschrijven?

I. PROEFGROEPCRITERIA

A. *Insluitende kenmerken*

1. *Leeftijd van het manifest worden van de stoornis.*

Tussen 14 en 25 jaar.

Kenmerk op grond van de anamnese.

Maatstaf is de datum van het eerste contact met psychiater, medisch opvoedkundig bureau of de datum van eerste opname.

(De hetero-anamnestiche gegevens over het tijdstip waarop door de omgeving voor het eerst het abnormaal gedrag is waargenomen zijn onvoldoende betrouwbaar).

2. *Opname in een psychiatrisch ziekenhuis is noodzakelijk gebleken.*

(Indicatief voor het feit, dat de tolerantiegrens van het milieu waarin de patiënt thuishoort is overschreden.

Het evenwicht is op beslissende wijze verstoord.

Ook indicatief voor het bestaan van gedragsstoornissen, maar niet indicatief voor de ernst van deze stoornissen. Het is denkbaar dat bij een lage tolerantie een lichte mate van gedragsstoornissen reeds voldoende is voor een beslissende verstoring van het evenwicht. Omgekeerd: bij een grote tolerantie kan ernstige pathologie soms niet tot opname aanleiding geven).

3. *De veranderde relatie tot de realiteit, in de zin van psychose.*

Beoordelingscriterium.

Alleen geldend als opnamekenmerk.

Alleen in het gedrag waarneembare stoornissen die zeer duidelijk zijn.

(Bijvoorbeeld: verbaal: formele denkstoornissen

waanvoorstellingen

hallucinaties

niet verbaal: op zeer opvallende wijze zich terugtrekken uit de omgeving, opvallende in-activiteit, initiatieven die onmiskenbaar niet in overeenstemming zijn met de mogelijkheden van de situatie, agressieve handelingen zonder motivering).

B. *Uitsluitende kenmerken.*

(Voornamelijk op basis van aantoonbare organische afwijkingen).

4. *Leeftijdsgrenzen waarbinnen de patiënt in de Tweede Psychiatrische Universiteitskliniek opgenomen kan blijven.*

a) ondergrens: 21 jaar.

(Uitsluiting van integratie- en regulatie-stoornissen van de puberteitsfase).

(Zie ook bij stratificatie-kenmerken).

b) bovengrens: 40 jaar.

(Uitsluiting van storende invloeden door het climacterium of daarop volgende fasen).

(Zie ook stratificatiekenmerken).

(Tevens wordt met deze grens gesteld, dat patiënt wat betreft leeftijd nog acceptabel is voor het werkmilieu. Voorts kan hier-

mee geacht worden, dat patiënt wat betreft zijn levensfase nog toekomstverwachtingen kan hebben, afgezien van de invloed van de bestaande stoornissen).

5. *Uitsluiting van stoornissen, ook gekenmerkt door een veranderde relatie tot de realiteit in de zin van een psychose, maar waarvan de aetiologie reeds bekend is.*

Verschillende beelden door Rümke beschreven onder pseudoschizofreniën.

6. *Uitsluiting van intercurrente ziekten.*

(Astma, diabetes, etc.).

7. *Uitsluiting van eerder doorgemaakte ziekten, die mogelijk aanleiding hebben kunnen geven tot cerebraal organische stoornissen.*

(traumata capitis, intensieve elektro-shockbehandeling, langdurig gebruik van medicijnen waarvan een mogelijke invloed bekend is, bijv. nidadongebruik bij tuberculose, „minimal cerebral damage”).

8. *Uitsluiting van manifeste neurologische functiestoornissen.*

9. *Uitsluiting van constitutionele factoren voor zover ze gepaard gaan met manifeste functiestoornissen.*

(Een status dysraficus zonder manifeste neurologische afwijkingen wordt niet uitgesloten. Een vroegtijdig behandeld congenitaal hartgebrek met geslaagde compensatie evenmin.

Bij een patiënt met een congenitaal hartgebrek kan men zich afvragen hoe de specifieke invloed hiervan op de opvoeding is geweest, wat de invloed hiervan is op de cerebrale circulatie, en voorts kan deze afwijking een belangrijke invloed hebben op de mogelijkheden van revalidatie. De patiënt kan bijv. niet meedoen met bewegingstherapie).

C. *Factoren die bewust niet als in- of uitsluitende kenmerken gebruikt zijn.*

1. *Heredität.*

(Uit de literatuur bekende onderzoeken kunnen niet bewijzen dat deze factor van invloed is).

2. *De duur van de stoornis.*

(Deze is op zichzelf niet beperkend. Door het stellen van een leeftijds-grens op grond van eerdergenoemde motieven wordt hier wel een grens aan gesteld).

3. *Het verloop.*

(Bij de kenmerken wordt geen onderscheid gemaakt tussen schubartig verloop, langzaam progressief verloop, defect toestand, etc. Dit is een beoordelingsprobleem waarbij geen voldoende betrouwbaarheid is te verkrijgen).

Wanneer u nu kijkt naar de kenmerken, die onder het hoofd „proefgroepcriteria” zijn vermeld, ziet u dat op grond hiervan in de Tweede Psychiatrische Universiteitskliniek patiënten worden opgenomen die

tussen de 21 en 40 jaar oud zijn,

bij wie de stoornis manifest werd tussen de 14 en 25,

waarbij in het gedrag zeer duidelijk waarneembare stoornissen voorkomen, die wijzen op een veranderde relatie tot de realiteit in de zin van een psychose,

voor wie opname in een psychiatrisch ziekenhuis noodzakelijk was gebleken en bij wie ondubbelzinnig aantoonbare somatische afwijkingen uitgesloten konden worden.

Terugkomend op wat wij over „non-probability sampling” bespraken is duidelijk, dat wij op grond van deze kenmerken niet kunnen stellen, dat we hier hebefrene patiënten onderzoeken. Hoezeer we ook hebben gepoogd om met eenduidige kenmerken zo dicht mogelijk te benaderen wat wij klinisch onder hebefrenie verstaan.

Voorts is nog niet bekend wat de populatie is, die aan deze kenmerken voldoet en hoe groot deze is.

Exacter is wanneer we zeggen, dat populatie en proefgroep nog identiek zijn. Toch stellen wij een proefgroep samen in het besef (terecht of ten onrechte) dat deze betrekking heeft op een grotere populatie.

De vooronderstelling van wetmatigheden dus.

Ons probleem van het samenstellen van een proefgroep is hier echter helemaal nog niet mee opgelost.

Aan de hand van enkele technieken voor het samenstellen van „non-probability samples” zou ik hier verder op in willen gaan.

Bij de eenvoudigste vorm van „non-probability sampling”, de z.g. „accidental sampling” wordt eenvoudig genoeg genomen met elke patiënt, die voldoet aan deze kenmerken tot een bepaald aantal bereikt is.

Wij zouden bijv. elke patiënt, die ons door de Sociaal Psychiatrische Dienst wordt aangeboden, en die aan deze kenmerken voldoet, kunnen opnemen.

Maar stel nu dat de Sociaal Psychiatrische Dienst ons voornamelijk patiënten zou sturen, die nog in de opvoedingssituatie thuishoren, terwijl er mogelijk een groot verschil is in het gedrag van patiënten, die wel nog in de opvoedingssituatie thuishoren en die niet meer in de opvoedingssituatie thuishoren.

In dat geval zouden de uitkomsten van ons onderzoek een misleidend beeld geven.

Het kan dan belangrijk zijn volgens het principe van de „quota sampling” rekening te houden met dergelijke mogelijke verschillen door het formeren van subgroepen. Het doel van „quota sampling” is selectie van een proefgroep waarin alle facetten, die mogelijk van belang zijn in relatie tot de populatie, vertegenwoordigd zijn, maar dan in bekende procentuele verhoudingen. We zouden bijv. 99% opvoedelingen en 1% niet-opvoedelingen kunnen opnemen. Welke procentuele aantallen aanvaardbaar zijn hangt echter af van het totaal aantal elementen van de proefgroep.

Er is een groot verschil in relevantie van de bevindingen van de 1% niet-opvoedelingen wanneer het totaal aantal elementen van de proefgroep 500 of 100.000 is.

1% van 500 is 5 en 1% van 100.000 is 1000.

Een algemenere uitspraak over bevindingen bij 5 patiënten is onaanvaardbaar, over 1000 patiënten eventueel wel.

Wanneer de procentuele verhoudingen voor bepaalde subgroepen bekend zijn en dan ook nog verantwoord, is later altijd een berekening voor de totale populatie mogelijk.

Maar hier doet zich een nieuw praktisch probleem voor.

In de regel is het niet mogelijk om in een psychiatrisch ziekenhuis te beschikken over een proefgroep van 500 mensen, laat staan over een die nog groter is.

Het doen van uitspraken over subgroepen, al zijn die dan ook in procentuele verhoudingen bekend, is dus niet verantwoord.

Het gebruik van het principe van „quota sampling” is daardoor voor proefgroepen in psychiatrische ziekenhuizen vrijwel onmogelijk.

De Tweede Psychiatrische Universiteitskliniek heeft de beschikking over 50 bedden.

In dergelijke omstandigheden verdient het de voorkeur volgens het principe van „stratificatie” een verdere beperking aan te brengen.

Volgens het principe van „stratified sampling” wordt de populatie verdeeld in 2 of meer strata.

De strata kunnen gebaseerd zijn op een enkel criterium (zoals bijv. de opvoedingssituatie of de intelligentie) of op een combinatie van twee of meerdere criteria.

Door de proefgroep op deze wijze onder te verdelen ontstaan een aantal subgroepen.

In plaats van onderzoek te doen met een „globale” proefgroep, kan nu één subgroep genomen worden.

De uitkomsten van een dergelijke, door meerdere kenmerken omschreven proefgroep, zijn van een beperkter geldigheid, maar de kansen dat deze uitkomsten minder zullen afwijken van een overeenkomstig deel van de *populatie* worden groter.

Deze stratificatie heeft echter alleen maar zin wanneer men redelijkerwijs kan verwachten, dat de verschillen in de uitkomsten van de onderscheiden categorieën groot zullen zijn.

Anders kan men deze stratificatie beter achterwege laten.

Als voorbeeld noem ik de stratificatiekenmerken, die in de Tweede Psychiatrische Universiteitskliniek voor verdere beschrijving van de proefgroep worden gebruikt.

STRATIFICATIEKENMERKEN

1. Patiënt mag geen „opvoeding” zijn.

Beoordelingskenmerk.

(Patiënt mag ook geen scholier meer zijn).

2. Patiënt moet, wat betreft leeftijd en levensfase, geacht kunnen worden eigen verantwoordelijkheid te dragen.
(Afgezien van het feit of hij dit onder invloed van de stoornissen ook werkelijk kan).
Beoordelingskenmerk.
3. De patiënt moet minstens de lagere school goed doorlopen hebben.
(Als maatstaf voor intelligentie, mogelijkheid tot leren, aanvaardbare werkinstelling, aangepast functioneren vóór het manifest worden van de stoornis. Dit is geen indicator voor normaliteit).
4. De familie waaruit patiënt komt moet minstens sinds 2 generaties (vóór hem) in Nederland woonachtig zijn.
(Gedragsvormen, gebonden aan een ander cultuurpatroon, kunnen de aanpassingsstoornis op onduidelijke wijze beïnvloeden. Daarnaast geeft deze factor aparte problemen in de revalidatie. Hiermee wordt ook uitgesloten dat de ouders van verschillende culturele achtergronden komen, wat specifieke relatieproblemen kan geven).
5. „Ziekteconfrontatie”. Patiënt moet zodanig met zijn aanpassingsstoornissen geconfronteerd zijn geweest, dat hij de opname aanvaardt.
(Dus niet van belang is het z.g. „ziekte-inzicht”).
6. De patiënt mag vóór zijn komst in de Tweede Psychiatrische Universiteitskliniek niet langer dan totaal 5 jaar gehospitaliseerd zijn geweest.
7. De familie moet de opname kunnen accepteren. Ze hoeven geen „ziekte-inzicht” te hebben.
8. Direct contact met de familie moet mogelijk zijn.

Bij de stratificatie is geen rekening gehouden met het sociale milieu waaruit de patiënt afkomstig is.

Evenmin is rekening gehouden met de sekse.

Hoewel te verwachten is dat er een groot verschil is in uitkomsten wat betreft een hernieuwde aanpassing tussen vrouwelijke en mannelijke patiënten, hebben we tenslotte besloten bij onze stratificatie daarmee geen rekening te houden en zowel vrouwen als mannen in gelijke verhoudingen in de kliniek op te nemen. Het motief daarvan is duidelijk.

We verwachten dat er juist ten aanzien van een hernieuwde aanpassing een gunstige invloed uitgaat van het samenbrengen van mannen en vrouwen in één leefgemeenschap.

Dit aspect brengt mij tot het laatste probleem, dat ik u vandaag nog zou willen noemen: het belang van de *omstandigheden*.

In een wetenschapsgebied zoals bijv. de fysica is het vanzelfsprekend, dat bijv. bij een experiment met een monster gas rekening wordt gehouden met factoren zoals de temperatuur en de druk. Deze kunnen immers van bepalende invloed zijn op de uitkomsten. In de literatuur heb ik slechts enkele verslagen van experimenten met proefgroepen van psychiatrische patiënten kunnen vinden, waarin de omstandigheden werden vermeld waaronder het onderzoek werd verricht.

Om nog even bij ons voorbeeld te blijven van de doelstelling van de hernieuwde sociale aanpassing: het spreekt vanzelf, dat het

van enorme invloed op de uitkomsten van een onderzoek is of de patiënten van een proefgroep opgenomen zijn op een afdeling met gesloten of met open deuren.

Ook om een experiment te kunnen herhalen moeten deze omstandigheden bekend zijn.

Toch is het niet toevallig, dat juist over deze omstandigheden meestal gezwezen wordt.

Het is heel moeilijk om deze te doorzien. Een vis is zich het minst bewust van het water waarin hij zwemt.

Nog moeilijker is het om deze omstandigheden te objectiveren.

De komst van een nieuwe dokter, of een nieuw lid van de verpleegstaf, of het weggaan van een van hen, geeft vaak merkbare veranderingen op de afdeling.

Toch is het noodzakelijk deze omstandigheden zo nauwkeurig mogelijk te specificeren willen de uitkomsten betekenis hebben. Niet alleen op de uitkomsten hebben de omstandigheden invloed, maar ook op de samenstelling van de proefgroep.

Om nog eens een vergelijking met de fysica te trekken: een gas dat bij een bepaalde druk vloeibaar wordt kan voor een experiment onbruikbaar zijn.

Zo bleek voor onze proefgroep een beperkende factor te zijn of een patiënt de omstandigheden in de groep verstoort of niet, waarbij niet noodzakelijk is of hij zich hierbij (al) kan aanpassen.

Wanneer voor één patiënt de deuren op slot moeten, worden hierdoor de omstandigheden sterk veranderd.

In het volgende is geprobeerd om deze omstandigheden enigermate te specificeren.

OMSTANDIGHEIDSKENMERKEN

Specifiek voor de situatie in de Tweede Psychiatische Universiteitskliniek.

1. De patiënt mag geen storend gedrag vertonen dat zo ernstig is dat het de groep verstoort.
(Het is niet noodzakelijk dat patiënt zich kan aanpassen).
(Bijvoorbeeld agressief gedrag dat de patiëntengroep niet meer verdragen kan is een uitsluitend kenmerk).
2. De patiënt moet de z.g. „gemengde” verplegingssituatie in zoverre kunnen verdragen, dat hij niet tot dusdanig storend gedrag komt dat de groep dit niet meer tolereert. (Zeer ernstige seksuele ontremmingen).
3. De patiënt moet in staat zijn om op een zaal te slapen tezamen met 5 tot 10 andere patiënten.
4. Patiënten die een voortdurende directe zorg nodig hebben, worden uitgesloten.
(Bijv. bij ernstige suicidaliteit, ernstige agressiviteit, langdurige voedselweigering, langdurige bedverpleging).
De activiteiten van de verplegenden zijn primair gericht op de behandeling van de hele groep en de patiënt als lid van de groep. Patiënten, die *te veel* persoonlijke aandacht vragen kunnen, gezien de structuur van de situatie, niet behandeld worden.
5. Een patiënt moet een mate van bewegingsvrijheid kunnen verdragen, die be-

paald wordt door het feit, dat de deur van de afdeling nooit op slot is. De deur van het ziekenhuis zonodig wel.

U zult met mij eens zijn dat de hier genoemde omstandigheden-kenmerken een nog zeer onbevredigend beeld geven.

Dit probleem heeft onze voortdurende aandacht en dat zal zo ook moeten blijven.

Een grote moeilijkheid daarbij is, dat we deze omstandigheden niet geheel in de hand hebben.

Voor al de wisseling in de samenstelling van de verpleegstaf en het algemeen bekende probleem van het gebrek aan verplegenden, geven ons voortdurend grote zorgen.

Ik heb gepoogd aan de hand van onze ervaringen bij het samenstellen van proefgroepen in de Tweede Psychiatrische Universiteitskliniek u een aantal problemen te beschrijven, die het samenstellen van proefgroepen in het algemeen betreffen.

Een probleem dat ik hier verder niet heb besproken is: dat de doelstellingen van het onderzoek en de theoretische concepten natuurlijk ook grote invloed hebben op de kenmerken die men kiest.

Zoals ik in het begin al stelde: Een bespreking van dit complexe probleem gaat ver buiten de doelstellingen van deze voordracht. Aan de hand van ons voorbeeld heb ik gepoogd een indruk te geven van de verschillende methoden voor het samenstellen van proefgroepen en de wijze waarop men tot het kiezen van een methode kan komen.

„Probability sampling” bleek voor ons onderzoek onmogelijk. Deze techniek zal in de psychiatrie voorlopig slechts een beperkte toepassing kunnen vinden doordat niet is te voldoen aan de eisen, aan een populatie te stellen. Wanneer wij echter in onze aspiraties bescheiden blijven en bereid zijn om omwegen te maken, is een verdere voortgang wel mogelijk.

Terwijl wij zo graag de schizofrenie zouden willen onderzoeken, zijn wij genoodzaakt genoeg te nemen met een proefgroepsamenstelling waarvan voorlopig nog zeer de vraag is of deze iets met ons probleem van de schizofrenie te maken heeft.

Toch lijkt het voor de psychiatrie al van grote waarde, wanneer uitkomsten verkregen kunnen worden die iets zeggen over bepaalde patiënten, die aan intersubjectief eenduidige kenmerken voldoen en die onder bekende omstandigheden verkeren.

SUMMARY

In this paper practical problems of sampling in the given situation of the psychiatric hospital are discussed. Theory and method underlying patient selection have far reaching consequences for the outcome of an investigation. Even if we leave out of consideration the statistical problems, and the interrelated problems

of theoretical concepts and goals of an investigation, there remain a lot of technical pitfalls of sampling itself which means: so many risks for the success of the investigation.

On the basis of the difficulties met by sampling schizophrenic patients it is demonstrated that in the situation of the psychiatric ward the principle of "probability sampling" is not applicable, when there is no well defined population. Clinical diagnosis in psychiatry is based on implicit evidence, for the greater part. Reliability of these diagnoses, as the basis for sampling and further investigation in order to get reliable information, is insufficient.

"Non probability sampling" is a useful tool in trying to find scientific rules. But conclusions are only valid for the patient who are in the sample and they can not be generalized to other patients unless replication with other samples is done.

This means that the criteria chosen as a basis for sampling must be objective, otherwise no replication is possible.

It is stated then, that further stratification on relevant criteria is not only inevitable, but also preferable in order to get conclusions that are useful.

It is stressed that for replication circumstances under which the investigation is done must also be stated explicitly. Only in this way, by being modest in planning and expectations, further reliable research in the psychiatric hospital is possible by means of sampling techniques. Sampling will always be an inevitable and necessary tool.

These problems are illustrated by examples taken from the research going on in the University Hospital for Psychiatry of the Willem Arntsz Stichting in Utrecht.

LITERATUUR

- Beyaert, F. H. L. — Gedragsbeoordeling in het psychiatrisch ziekenhuis. De P.U.K.-schalen voor gedragsbeoordeling. Leiden, Stenfert Kroese (1960). Proefschrift Utrecht.
- Cochran, W. G. — Sampling techniques. New York: Wiley (1956).
- Dijkhuis, J. H. — Het beoordelen in de psychologie. Utrecht, Bijleveld (1960). Proefschrift Utrecht.
- Fischer, A. A. — Nosologisch denken in de psychiatrie. Over theoretische en methodologische problemen. Utrecht, Elinkwijk (1963). Proefschrift Utrecht.
- Kish, L. — Selection of the sample. In: Festinger & Katz: Research methods in the behavioral sciences. New York: Holt (1953): p. 174-240.
- Kline, N. S. — Samples and controls in psychiatric research. Psychiat. Quart., 27 (1953): 474.
- Kreitman, N. — The reliability of psychiatric diagnosis. J. ment. Sci. 107 (1961): 876-886.
- Kreitman, N., Sainsbury, P., Morrissey, J. a.o. — The reliability of psychiatric assessment: An analysis. J. ment. Sci., 107 (1961): 887-908.
- Klinische onderscheidingen in de groep der schizofrenieën". Ned. Tijdschr. G., 102, no. 28 (1958): blz. 1361-1366.
- Rümke, H. C. — Mehrdimensionale Diagnostik und nosologische Entität. Über einmaliges und uniformes Geschehen. Festschrift zum 70. Geburtstag von E. Kretschmer. Stuttgart, Thieme (1958): blz. 1-9.
- Rümke, H. C. — Nosology and classification. Field studies in the mental disorders, edited by J. Zubin. New York: Grüne & Stratton (1961): blz. 73-84.