

Retrospectief onderzoek met behulp van medische archieven, speciaal het huisartskaartsysteem

DOOR E. W. ROSCAM ABBING*

Iedere wetenschappelijk geïnteresseerde arts kent de aantrekkingskracht van het archief. Een grote hoeveelheid gegevens vraagt als het ware om analyse. Het werk lijkt grotendeels reeds gedaan: het materiaal is verzameld. Bovendien is het materiaal niet besmet door een hypothese die zo nodig moet worden bewezen. Bij terugblikkend onderzoek vindt de hypothesevorming immers achteraf plaats.

Waarom ziet men ondanks deze voordelen zo zelden de resultaten van retrospectief onderzoek gepubliceerd? Het antwoord op deze vraag is niet nieuw. Dat desondanks wordt overgegaan tot een uiteenzetting van de beperkingen en problemen verbonden aan retrospectief onderzoek, vloeit voort uit de wens enkele algemene inzichten toe te spitsen op het kaartsysteem van de huisarts. Als basis dient enige eigen ervaring met genoemd systeem.

I Beperkingen en problemen.

A Berkson's fallacy en aanverwante problemen. De afgelopen jaren is herhaaldelijk in medische periodieken gewezen op het gevaar voor Berkson's fallacy, voor onjuiste conclusies uit ziekenhuisarchieven en sectieverslagen (*Van der Linden; Oppers; Rümke.*) Daarbij ging het om de niet te verwaarlozen omstandigheid dat patiënten met verschillende aandoeningen ook een verschillende kans hebben in een ziekenhuis te worden opgenomen of op een sectietafel te belanden. Sommige aandoeningen geven immers in een hoog percentage der gevallen onder de bevolking aanleiding tot opname of sectie (bijvoorbeeld maligne nieuwvormingen), terwijl andere dat slechts zelden doen (zoals furunkels).

De frequentie waarmee combinaties van aandoeningen in een archief worden aangetroffen zegt daarom weinig over het onder de bevolking voorkomen van dergelijke combinaties. Wanneer men bijvoorbeeld bij sectie van patiënten met een maligne nieuwvorming in een lager percentage furunculosis aantreft dan bij patiënten met benigne nieuwvormingen, dan staat daarmee nog geenszins

Samenvatting. Enkele problemen en beperkingen verbonden aan retrospectief onderzoek worden uiteengezet, in het bijzonder voor het geval dat gebruik wordt gemaakt van het kaartsysteem van de huisarts. Vervolgens wordt een aantal mogelijkheden van dergelijk onderzoek aangegeven. Ook de praktische uitvoering komt ter sprake.

vast dat bij screening van de bevolking dezelfde verhoudingen zouden worden aangetroffen. Immers de kans om aan deze aandoeningen te overlijden verschilt per aandoening (uiteraard evenals de kans op pathologisch anatomisch onderzoek). *Rümke* geeft een aantal gefingeerde rekenvoorbeelden, die duidelijk demonstreren hoezeer de getalmatige verhoudingen in ziekenhuisarchieven en sectieverslagen kunnen afwijken van die onder de bevolking waaruit de patiënten afkomstig zijn.

Men dient zich af te vragen of men bij het analyseren van gegevens uit het huisartskaartsysteem ook het risico loopt slachtoffer te worden van Berkson's fallacy. Dat dit risico inderdaad aanwezig is, mag blijken uit een voorbeeld: een huisarts gaat na of van zijn vrouwelijke patiënten binnen een bepaalde leeftijdsgroep een groter aantal over een geringe libido klaagt voor zover zij orale anticonceptiva gebruiken, dan voor zover zij op andere anticonceptiva zijn aangewezen. Deze huisarts komt dan met zijn analyse in een situatie kenmerkend voor Berkson's fallacy: in een hoog percentage consulteert men de huisarts voor orale anticonceptiva (een aantal patiënten bezoekt een consultatiebureau); in een lager percentage voor andere anticonceptiva, terwijl de patiënten met een geringe libido als primaire klacht in nog weer een ander percentage de huisarts zullen raadplegen. Als gevolg van deze ongelijke percentages treedt Berkson's fallacy op en kunnen uit het onderzoek dus geen valide conclusies worden getrokken. Dit voorbeeld demonstreert tevens een ander probleem, dat speciaal een rol speelt bij een hulpverlener als de huisarts.

De frequentie, waarmee over bepaalde zaken wordt geklaagd door patiënten, is namelijk mede

* Arts-medewerker aan het Instituut voor Sociale Geneeskunde van de Universiteit van Amsterdam.

afhankelijk van de mogelijkheid tot klagen, dus van het contact met de hulpverlener. Een patiënte die op gezette tijden haar huisarts bezoekt in verband met anticonceptie kan bij die gelegenheid een groot aantal klachten uiten die op zichzelf voor haar onvoldoende reden zijn haar huisarts te bezoeken. Men zou van bijgevoegde klachten kunnen spreken, klachten die slechts ter sprake komen, omdat de patiënt toch al bij zijn arts is. Voorbeelden van dergelijke klachten zijn: adipositas, slapeloosheid, spanningen. Een en ander treft men dikwijls aan op de kaarten van patiënten die voor regelmatige controle komen. Men heeft dan te maken met een gevolg van het veelvuldige contact. Berekent men echter de klachtenfrequentie per contact, dan dient men daarbij de mogelijkheid te verdisconteren dat een intensief contact met de arts het uiten van bijgevoegde klachten vergemakkelijkt.

De hier geïntroduceerde bijgevoegde klachten dient men theoretisch te onderscheiden van de verborgen hoofdklachten. Bij de laatstgenoemde soort gaat het om klachten die pas in de loop van het gesprek of pas bij het afscheid worden vermeld, maar die niettemin de reden voor het bezoek vormen. Deze onderscheidingen zijn bij een retrospectief onderzoek niet met zekerheid uit het materiaal af te leiden. Het is echter nodig bij potentieel bijgevoegde klachten zich te hoeden voor premature conclusies.

Aan Berkson's fallacy en aanverwante problemen wordt daarom zoveel aandacht besteed, omdat deze problemen de laatste jaren frequent aan de orde zijn. De wetenschappelijke belangstelling van de huisarts verbreedt zich. Onderzoek in de psychosociale sfeer is niet meer uitzonderlijk. Pogingen worden gedaan statistisch een samenhang vast te stellen tussen ziekte en psycho-sociale achtergronden. Een voorbeeld van een dergelijke poging is het geprecodeerde registratieproject van de groepspraktijk Ommoord te Rotterdam (*Lamberts, 1970*). Middels een geavanceerde techniek zoals factoranalyse hoopt men inzicht te verkrijgen in de afhankelijkheidsstructuur van een groot aantal variabelen.

Daarbij is men zich bewust van het gevaar voor Berkson's fallacy blijkens een recente publikatie (*Lamberts, 1972*). Juist psycho-sociale problemen zijn lang niet altijd een aanleiding de huisarts te raadplegen. De huisarts ziet dus een selectie die beslist niet uitsluitend op grond van de ernst van de klachten is samengesteld. Tal van andere factoren bepalen het klaagedrag, onder andere de sociale herkomst, die mede bepaalt of iemand al dan niet voor een probleem zijn arts zal consulteren. Een uitgebreide analyse van het klaagedrag vindt men bij *Freidson*. Hij wijst op een mogelijk verband tussen enerzijds variabelen zoals leeftijd, geslacht, etnische groep, sociale laag en anderzijds de schroom of het

gemak waarmee men voor bepaalde problemen een arts raadpleegt. Indien dit verband inderdaad bestaat kan men niet eerder conclusies trekken uit de statistische afhankelijkheidsstructuur van praktijkgegevens dan nadat men dit verband heeft gekwantificeerd. Het is echter algemeen bekend hoe moeilijk het is de discrepantie vast te stellen tussen klacht en objectieve behoefte. Laatstgenoemd begrip laat zich slecht definiëren en operationaliseren.

B Onafhankelijke hypothesevorming. De hypothese en het materiaal waaraan moet worden getoetst, dreigen elkaar op twee wijzen te besmetten: 1 de hypothese wordt gevormd op grond van ervaringen met het materiaal waaraan moet worden getoetst; 2 hypothese en praktijkvoering stammen uit dezelfde theoretische achtergrond.

ad 1 Als huisarts krijgt men zo nu en dan vermoedens over een samenhang van waargenomen verschijnselen. Een bevestiging van dergelijke vermoedens door een kwantitatieve analyse van de eigen praktijkgegevens is dan echter van geringe waarde. Het is heel wel mogelijk dat slechts een toevallig in de praktijk aanwezig fenomeen wordt opgespoord. Waren daarentegen de vermoedens in geheel anders samengestelde praktijken bevestigd, dan had men een uitkomst verkregen van veel algemenere betekenis. Het gaat hier dus om de volgende beïnvloeding: materiaal → hypothese.

ad 2 De theorie beïnvloedt vanzelfsprekend de praktijk en dus ook de anamnese. Wie heeft geleerd, dat maagulcus en obstipatie dikwijls samengaan, zal bij maagklachten naar obstipatie vragen. Wil hij echter dit verband nog eens statistisch bewijzen dan zal hij bij een retrospectief onderzoek op de kaarten van zijn ulcuspatiënten relatief vaak de vermelding obstipatie aantreffen, omdat hij bij deze patiënten naar de ontlasting heeft gevraagd. Slechts een prospectieve opzet voorkomt dit zeer fundamentele bezwaar. De hier besproken beïnvloeding ziet schematisch als volgt eruit: theorie → materiaal
→ hypothese

C Representativiteit van het patiëntenbestand. De patiënt kiest veelal zijn arts, er is dus sprake van selectie. Topografisch overlappen huisartspraktijken elkaar en vormen zij geen aselechte steekproef uit de binnen een bepaald gebied wonende populatie. Zo nu en dan is het echter mogelijk de nagenoeg volledige populatie van een gemeente in een onderzoek te betrekken door alle praktijken uit die gemeente samen te voegen. Uiteraard zijn dan de uitkomsten van een dergelijk onderzoek niet generaliseerbaar naar andere gemeenten. Wil men toch de uitkomsten van een onderzoek kunnen veralgemenen, dan dient men de te onderzoeken praktijken volgens een

aantal criteria zoals urbanisatiegraad, leeftijdsopbouw en dergelijke goed gespreid over dat gebied te kiezen, hetgeen een zeer moeilijke opgave is (*Oliemans*). Het werken met meer dan een praktijk bij een retrospectief onderzoek geeft overigens dikwijls problemen in verband met de verschillende notatiegewoonten van de betreffende artsen (zie ID 1b).

D *Het materiaal*. Achtereenvolgens worden behandeld de volledigheid van het materiaal (1) als patiëntenadministratie (a) en als medische registratie (b), de eenduidigheid (2) en de leesbaarheid (3) van het materiaal.

1a De volledigheid van het materiaal als patiëntenadministratie betreft zaken zoals: hoeveel patiënten hebben nog geen kaart en zijn de kaarten van inmiddels overleden of tot andere praktijken behorende patiënten nog niet verwijderd?

Deze vragen zijn van belang omdat dergelijke administratiefouten de uitkomsten van het onderzoek zullen beïnvloeden. De patiënt van wie ten onrechte de kaart een jaar geleden niet werd verwijderd, wekt bij de onderzoekers de indruk het laatste jaar niet te hebben geklaagd. In een Engels onderzoek (*Morrell en medewerkers*) wordt een aanzienlijk verschil aangetoond tussen het feitelijke patiëntenbestand en het administratieve. Daarbij blijkt dit verschil te variëren naar leeftijd en geslacht. Deze bevindingen zijn voor de bekende epidemioloog *Holland* reden te concluderen dat huisartsmateriaal ongeschikt is voor epidemiologisch onderzoek. Overigens dient men zich af te vragen in welke mate deze problematiek ook in de Nederlandse verhoudingen een rol speelt. De administratie van de ziekenfondsen maakt correcties voor een belangrijk deel van de praktijk mogelijk.

1b De volledigheid van het materiaal als medische registratie betreft vragen zoals: welke spontaan geuite en welke anamnestic en bij onderzoek gevonden klachten en symptomen werden genoteerd; hoe consequent werd de registratie volgehouden; werd de registratie gecontinueerd tijdens vakantie en waarnemingen?

Deze vragen leggen een aantal zwakke plekken bloot van het huisartskaartsysteem als materiaalbron voor wetenschappelijk onderzoek. Indien de praktijkvoerder niet kan aangeven wat hij over het algemeen wel en wat hij doorgaans niet noteert of indien hij inconsequent (zonder vaste gewoonten) bij de registratie te werk gaat, zijn de kaartsysteemgegevens voor wetenschappelijk onderzoek waardeeloos met als uitzondering de vermelding van feiten die vanwege hun zeer grote medisch-praktische betekenis altijd worden genoteerd (diabetes, epilepsie en wat dies meer zij).

2 Eenduidig is het materiaal van een ongecodeerd-

de registratie veelal niet. De gebruikte omschrijvingen in het kaartsysteem zijn dikwijls bij gebrek aan een zekere diagnose vaag en voor velerlei uitleg vatbaar, althans bij interpretatie door een ander dan de registrerende arts. Indien echter codenummers in plaats van uitgebreide omschrijvingen worden gebruikt, bestaat weliswaar geen onduidelijkheid meer, maar moet reeds terstond bij het eerste contact met de patiënt voor een diagnose uit de code-lijst worden gekozen met uiteraard een grote kans op fouten.

3 Dat de leesbaarheid van in continue tijdnood gemaakte aantekeningen dikwijls voor buitenstaanders, maar ook voor insiders niet groot is, zal bij niemand verbazing wekken, al zal deze omstandigheid menige onderzoeker tot wanhoop brengen. In dit verband is een gecodeerde registratie zonder meer te prefereren.

Na deze lange litanie over haast onoplosbare problemen valt het niet moeilijk ook ten aanzien van het huisartskaartsysteem de stelling van *Mainland* te onderschrijven: „... that many collections of data do not contain nearly so much usable information as their gross bulk would suggest”.

Ofschoon kennis van onmogelijkheden en problemen nuttig is ter voorkoming van vergeefse pogingen, wordt een dergelijke kennis pas werkelijk nuttig indien deze leidt tot het opsporen van mogelijkheden en tot oplossingen van problemen.

II *Mogelijkheden en uitvoeringsadviezen.*

A *De vraagstelling*. Geschikt voor retrospectief onderzoek zijn vragen naar de praktijkvoering, bijvoorbeeld hoe staat het met de consult- en verwijfs-frequentie bij de verschillende leeftijden, geslachten etcetera van patiënten in het algemeen en bij particuliere tegenover fondspatiënten?

Voorts kan men onderzoek doen naar het vóór-komen van bepaalde aandoeningen en eventueel naar een onderlinge samenhang tussen die aandoeningen. Daarbij is het echter noodzakelijk zich te beperken tot die aandoeningen die steeds worden genoteerd en die ook in een zeer hoog percentage van de gevallen tot een bezoek aan de huisarts leiden; dit ter voorkoming van problemen als onder IA genoemd. Hoe aantrekkelijk het voor de huisarts ook lijkt zijn kaartsysteem te benutten voor onderzoek naar een verband tussen klachten als adipositas, spanningen, slaapstoornissen enerzijds en sociale achtergronden anderzijds, een dergelijk onderzoek strandt op methodologische problemen. Het verdient aanbeveling hypothesevorming en toetsing niet in dezelfde praktijk te laten plaatsvinden. Om hypothesen van elders te kunnen betrekken heeft men echter een soort ideeënmarkt nodig.

In die behoefte voorzien de medische tijdschriften enigszins, al mag men betwijfelen of nog slecht onderbouwde hypothesen inderdaad ter publikatie worden aangeboden. Voor de ideeënuitswisseling zijn gesprekken met collegae waarschijnlijk belangrijker.

B De praktijk. De voorkeur verdienen praktijken die sterk gelijken op „de gemiddelde Nederlandse praktijk” ten aanzien van criteria zoals leeftijdsverdeling, verdeling naar geslacht en naar sociale afkomst. Het is echter onhaalbaar „de gemiddelde Nederlandse praktijk” in alle opzichten te benaderen. Een praktijk zal immers zelden zowel urbaan als ruraal zijn, zowel exclusieve als „volksbuurt”-kenmerken hebben.

Representativiteit voor de Nederlandse situatie zal men beter kunnen benaderen door met meer dan een praktijk te werken. Daarbij zal uiteraard met het toenemen van het aantal praktijken de kans op representativiteit groeien, maar tevens zullen de problemen toenemen ten gevolge van de verschillende notatiegewoonten van de deelnemende artsen. Het werken met meer dan een praktijk beperkt het aantal in aanmerking komende vraagstellingen sterk. Slechts die variabelen kunnen in het onderzoek worden betrokken, die in gelijke mate door de deelnemende artsen worden genoteerd. Het zal daarom dikwijls uitsluitend kunnen gaan om ernstige somatische aandoeningen.

C De praktijkvoerder. Voor de keuze van een praktijk is de praktijkvoerder van minstens zo groot belang als de patiëntensamenstelling. Diens houding ten aanzien van de registratie is doorslaggevend. Slechts artsen die met vaste regelmaat in hun kaartstelsel aantekeningen maken en daarbij consequent te werk gaan, komen in aanmerking. Het is duidelijk dat door deze restrictie zeer veel praktijken voor onderzoek afvallen. Toch zijn er nog meer beperkingen noodzakelijk.

Het kaartstelsel dat voortdurend door meer dan een arts wordt bijgehouden, laat zich slecht analyseren, aangezien twee notatie-stijlen door elkaar heenlopen. Ook in de gewone eenmanspraktijk wordt soms tijdens waarnemingen het kaartstelsel voorzien van „vreemde” aantekeningen. Indien het echter om niet meer dan enkele weken vakantie gaat, kan de onderzoeker soms — afhankelijk van de vraagstelling — deze aantekeningen herkenbaar aan het handschrift of de consultdatum buiten beschouwing laten.

Dat de praktijkvoerder een enigszins leesbaar handschrift moet hebben indien hij het onderzoek niet zelf doet, heeft geen betoog. Problemen rond leesbaarheid en interpretatie spelen echter geen rol, wanneer het kaartstelsel codecijfers bevat. Helaas zullen er echter weinig artsen zijn die hiervoor de E-

lijst van het Nederlands Huisartsen Genootschap inderdaad benutten. Het voortdurend moeten zoeken naar de passende code maakt deze werkwijze onaantrekkelijk. Daar tegenover staat, dat na een zekere inwerkperiode deze notitiewijze niet veel tijd hoeft te kosten, omdat men de codenummers dan uit het geheugen kan putten. Men beschikt dan over een soort steno, dat slechts voor de persoonlijke gegevens moet worden aangevuld met conventionele notities. Het eerder genoemde bezwaar dat een coderingssysteem dwingt tot het stellen van een voorbarige diagnose is te ondervangen door een nummer op te nemen voor twijfelgevallen of door achter het nummer van de diagnose een vraagteken te plaatsen.

Naast de houding van de arts ten aanzien van de registratie is soms ook diens houding in medisch-ethisch opzicht van belang. Het heeft bijvoorbeeld weinig zin om ter bepaling van de behoefte aan abortus bij een in deze conservatief denkend arts na te gaan hoe dikwijls hij met verzoeken om abortus wordt geconfronteerd. Zijn patiënten zullen, zijn standpunt kennend, niet met dergelijke verzoeken komen. De ontvankelijkheid van de arts voor bepaalde klachten is steeds van belang. De arts die zich weinig thuis voelt in de psycho-sociale problematiek, zal zijn patiënten (onbewust) niet stimuleren tot het uitspreken van dit soort problemen.

D De onderzoeker. Uit het voorgaande zal duidelijk zijn geworden dat aan het samengaan in één persoon van de kaartstelsel-eigenaar en de onderzoeker vele voordelen zijn verbonden zoals bijvoorbeeld bij problemen rond interpretatie en leesbaarheid. Bovendien spelen het beroepsgeheim en de bereikbaarheid van het kaartstelsel eveneens een rol.

Een bezwaar is echter dat de huisarts die voor een toetsend onderzoek zelf zijn kaartstelsel bestudeert, de hypothese aan een ander moet ontleen (zie IB1). Dit is een zwaarwegend psychologisch bezwaar, dat men evenwel zal moeten accepteren, wil men niet stuiten op een groot aantal praktische problemen verbonden aan onderzoek van anderen kaartstelsel.

Het doel van dit artikel is onder andere een stimulant te vormen voor bestudering van het eigen kaartstelsel om daarmee een van een ander afkomstige hypothese te toetsen. Indien toch anderen kaartstelsel moet worden bestudeerd, is het verstandig van te voren een aantal kernwoorden (of afkortingen) door de eigenaar van het stelsel in diens haastige handschrift te laten opschrijven om zo het lezen te vergemakkelijken. In een dergelijk geval verdient het bovendien aanbeveling alvorens met het eigenlijke onderzoek te beginnen, eerst tijd te nemen om al lezend aan het handschrift te wen-

nen. Aldus voorkomt men dat in de beginfase van het onderzoek over meer relevante informatie wordt heengelezen, dan aan het einde daarvan (inter-observer-variation).

Het werken met meer dan een onderzoeker moet zo mogelijk worden vermeden, aangezien dit het risico van een inter-observer-variation met zich brengt. Het vermogen een bepaald handschrift te ontcijferen en de wijze waarop formuleringen worden geïnterpreteerd, zullen immers van onderzoeker tot onderzoeker variëren. Omdat het doorlezen van systeemkaarten echter tijdrovend is en omdat een buitenstaander dikwijls slechts kortdurend over het systeem kan beschikken, ontkomt men niet aan een bestudering door meer mensen tegelijkertijd, tenzij het onderzoek kan geschieden door de eigenaar van het systeem, die dan zijn speuractiviteiten meer kan uitsmeren over momenten die hem goed uitkomen.

E De methode. Geheel afhankelijk van de vraagstelling kiest men een methode. In dit verband kan slechts op enkele basismogelijkheden worden gewezen.

Indien men van de vraagstelling uitgaat meer te willen weten over het vóórkomen van bepaalde klachten, staan onder meer de volgende wegen open:

- 1 klachtfrequentie per tijdseenheid, bijvoorbeeld seizoen-gebonden, in de loop der jaren toenemend, daarbij rekening houdend met de aard van het patiëntenbestand;
- 2 klachtfrequentie per patiënt per tijdseenheid zoals het laatste jaar werd geklaagd door mannen dan door vrouwen, meer door bejaarden dan door jeugdige patiënten; deze methode is ongeschikt voor bijgevoegde klachten, die immers vaker worden geuit, door frequenteerders en vaste controlepatiënten;

- 3 klachtfrequentie per consult per patiënt per tijdseenheid: hadden vrouwen het laatste jaar vaker bijgevoegde klachten dan mannen en wat dies meer zij.

Dit soort beschouwingen zijn, gezien de vele mogelijke vraagstellingen, onbeperkt uit te breiden. Ik wil echter volstaan met een voor de hand liggend advies in de vorm van een variant op de bekende aanbeveling: „Dokter niet zelf, maar raadpleeg uw dokter” en wel „Raadpleeg alvorens met een onderzoek te beginnen een onderzoeksdeskundige”.

Het voorgaande maakt, naar ik hoop, niet alleen de beperkingen, maar ook de resterende mogelijkheden van retrospectief onderzoek met behulp van het huisarts-kaartsysteem duidelijk. Wellicht voorkomt het een enkele vergeefse poging en zal het daarnaast het slagen van andere pogingen bevorderen, waarmede dan het doel van dit artikel zou zijn bereikt.

Summary. Retrospective studies with the aid of medical records, particularly general practice card-index systems. Some problems and limitations of this type of retrospective study are discussed with special reference to those making use of general practice card-index systems. A number of possibilities of this type of study is outlined, and the practical procedure is discussed.

- Freidson, E. Profession of medicine. A study of the sociology of applied knowledge. Dodd, Nead and Co., New York, 1972.
- Holland, W. W. Data Handling in Epidemiology. Oxford University Press, 1970.
- Lamberts, H. (1970) huisarts en wetenschap 13, 10-14
- Lamberts, H. (1972) huisarts en wetenschap 15, 369.
- Linden, W. van der (1960) Ned. T. Geneesk. 104, 1108.
- Mainland, D. Elementary Medical Statistics. W. B. Saunders, Philadelphia and London, 1963.
- Morrell, D. C. (1970) J. roy. Coll. gen. Practit. 19, 331.
- Oliemans, A. P. Morbiditeit in de huisartspraktijk. Stenfert Kroese, Leiden, 1969.
- Oppers, V. M. (1964) Ned. T. Geneesk. 108, 1153
- Rümke, Chr. L. (1970) Ned. T. Geneesk. 114, 756.