

Schouderaandoeningen

Schouderaandoeningen in de huisartspraktijk

De Jongh AC. *Dissertatie. Rotterdam: Erasmus Universiteit Rotterdam, 1994; 241 bladzijden*

De resultaten van dit proefschrift benadrukken opnieuw de noodzaak 'evidence based medicine' tot een belangrijke speerpunt in huisartsgeneeskundig onderzoek en onderwijs te maken. Duidelijk wordt dat wijsheden van oude leermeesters niet altijd op stevige grondvesten staan verankerd.

De Jongh doet verslag van 'een poging veel voorkomende ziekten van de schouder betrouwbaar te classificeren met behulp van numerieke taxonomie' op basis van gegevens uit anamnese en lichamelijk onderzoek. Hij begint met een uitvoerige verhandeling over taxonomische begrippen en de ontwikkeling van de nosologie. Voor deze beschrijving wordt teruggegaan naar de biologische wetenschappen. De empirische geneeskunde wordt naast de conceptuele geneeskunde gezet, van belang om concepten die later in het onderzoek worden gepostuleerd te rechtvaardigen.

Een Medline-search aangevuld met relevante literatuur leverde 600 referenties op voor de literatuurstudie. De Jongh beschrijft vervolgens, aan de hand van de historische ontwikkelingen in de afgelopen eeuw, de verschillende klinische beelden, zoals subacromiale inklemming, capsulitis en de 'overige' aandoeningen van de schouder. Er blijken geen eenduidige beschrijvingen van deze beelden te bestaan. Dit brengt de Jongh tot één van zijn stellingen: 'Bestudering van de literatuur betreffende aandoeningen van de schouder over een periode van honderd jaar leert dat naschrijven en overschrijven ook in de geneeskunde velen onverdiend roem bezorgen'.

Theoretisch model

Het onderzoek wordt gebaseerd op het paradigma van Skinner, dat ervan uitgaat dat het classificatieproces een aantal stappen moet doorlopen: achtereenvolgens het formuleren van een theorie, de interne validatie (betrouwbaarheid) en de externe validatie (prognostische en therapeutische kenmerken).

De literatuur leverde dus geen betrouwbare ziekteclassificatie voor aandoeningen van de schouder op. De meeste literatuur kwam bovendien uit de kliniek. Daarom ont kwam de auteur niet aan de noodzaak een theoretisch model te ontwikkelen voor de indeling van schouderaandoeningen in de huisartspraktijk op basis van praktijkervaring en door middel

van consensus. Er werd bij deze theorievorming geen volledigheid nagestreefd; slechts veel voorkomende aandoeningen werden in de classificatie opgenomen. De keuzen werden ook hier op basis van ervaring gemaakt. Uitgangspunt was een taxonomie te construeren op basis van klinische syndromen.

Dit leidde tot het bewerken van de volgens De Jongh vier belangrijkste klinische syndromen. Hiervoor werden verklaringen vanuit basale (anatomische, pathofysiologische) laesies aan de schouder gezocht; plaatsen waar de weefselbeschadiging primair plaatsvindt en leiden tot klinische problemen. Onderscheiden worden het glenohumerale gewrichtskapsel, leidend tot het capsulair syndroom; de subacromiale ruimte met als klinische representanten het chronisch subacromiaal syndroom en het peracut optredende beeld van de acute bursitis; en het acromioclaviculaire gewricht, leidend tot het acromioclaviculair syndroom. Met deze indeling kon een empirische toetsing worden uitgevoerd, ofwel een begripsvalidatie van de theoretische classificatie.

Empirisch onderzoek

Vervolgens wordt de validiteit van het theoretische model getoetst in een prospectief patiëntgebonden onderzoek. Hiervoor zijn de gegevens gebruikt van een aselechte groep patiënten die op het spreekuur van de huisarts kwamen met schouderklachten. De patiënten werden aangebracht door 29 huisartsen. Acute beelden werden telefonisch bij de onderzoeker aangemeld en geleidelijk ontstane klachten mochten per post worden aangemeld. Zo ontstond een delay na het huisartsbezoek van 1 tot 10 dagen voordat de patiënten werden onderzocht. Het onderzoek werd uitgevoerd door zesdejaars studenten.

De patiënten waren 'nieuwe' patiënten met een klacht van pijn en/of functiestoornis gelokaliseerd in het gebied van de schoudergordel en/of bovenarm. Uitgesloten werden patiënten die reeds onder behandeling van de tweede lijn of de fysiotherapeut waren (1 patiënt), dubbelzijdigheid (1) en patiënten die achteraf een incompleet registratieformulier hadden (12 patiënten). In totaal werden 89 patiënten ingesloten. De gemiddelde leeftijd was 46,3 jaar; 73 procent was vrouw.

Er werden 36 variabelen geselecteerd, afkomstig uit de volgende groepen: variabelen met nosografische kenmerken (anamnese en lichamelijk onderzoek), met symptoomaspecten (pijn en lokalisatie), met intrinsieke (invloed armbeweging) en extrinsieke kenmerken (sensibiliteit, invloed nekbewegingen). Ge-

toetst werd door middel van numerieke taxonomie, clusteranalyse en factoranalyse. Voorts werd gebruik gemaakt van een wiskundige methode van patroonherkenning, omdat 'patroonherkenningsmechanismen van de mens nogal eens tekort schieten, zeker bij multifactorieel bepaalde gegevens'. Door het bestuderen van de variabelentabel en vervolgens per cluster van alle variabelen de mediaan en het gemiddelde te berekenen, krijgt men een indruk van de specifieke eigenschappen van de zich onderscheidende clusters.

De twee grote clusters die op deze wijze ontstaan, onderscheiden zich door 'pijn' en 'bewegingsbeperking'. Deze kunnen goed worden gerelateerd aan de theoretische classificatie. Beide hoofdgroepen hebben weer onderverdelingen. De onderverdelingen zijn niet goed te relateren aan de theoretische classificatie, op twee na die overeenkomsten hebben met de bursitis en de capsulitis.

Commentaar

Bij een exercitie als deze zijn natuurlijk altijd kanttekeningen te maken, wat de auteur voor een deel ook zelf doet. Het uitgangspunt van de classificatie is de eigen waarneming, waardoor het theoretische model wordt bepaald, een en ander conform de theorieën van Cyriax. Andere uitgangspunten zijn ook mogelijk.

Hoewel een aselechte patiëntselectie plaatsvond, wordt nergens verantwoord hoe de groep patiënten precies verdeeld was. De populatie lijkt enigszins aan de jonge kant, vergeleken met de gemiddelde leeftijd van patiënten met schouderaandoeningen in de huisartspraktijk. Voor de externe validiteit is een registratie van alle stadia van de ziekte van belang.

Een tweede probleem dat niet geheel uit de verf komt, is het onderscheid tussen de classificering voor wetenschappelijke doeleinden en de classificering voor het dagelijks handelen: welke ziektesystematiek is voor de dokter van belang en welke voor de wetenschapper, die eenduidige beelden bij elkaar wil zetten.

Zeker bij de selectie van de variabelen is er een zekere mate van willekeur, en de auteur schrijft dan ook dat een andere selectie wellicht tot andere uitkomsten zou hebben geleid. Een duidelijke verantwoording voor de selectie wordt echter niet gegeven. Tevens is er geen intra-observatie gemeten.

Een ander detail is de systematiek van de literatuurstudie. Hoe de zoekstrategie is geweest, is niet terug te vinden. Als naslagwerk voor een nieuwe ontwikkeling met betrekking tot classificeren is de literatuur van groot belang. Het zou prettig voor toekomstige onder-

zoekers zijn geweest indien een cumulatieve literatuurlijst was toegevoegd.

Praktische consequenties

De resultaten van dit onderzoek hebben (nog) geen praktische consequenties. Het nodigt wel uit tot nader onderzoek en externe validering van de bevindingen, waarbij de prognostische en therapeutische betekenis van de gevonden clusters moet worden aangetoond. Symptomencomplexen zijn van elkaar te onderscheiden; nu moet nog het belang van de diagnostiek voor de behandeling worden aangetoond.

De Jongh verdient alle lof voor het aanpakken van een ingewikkeld en relatief onontgonnen terrein als de numerieke taxonomie. Hij heeft daarbij een belangrijke weg aangegeven voor verdere onderbouwing en theorievorming in de huisartsgeneeskunde. Het boek wordt vooral aanbevolen voor geïnteresseerden in classificatieproblematiek.

Wim Schuurman

Reactie

In de geneeskunde ligt de prioriteit maar al te vaak bij epidemiologisch en therapie-vergelijkend onderzoek; het wetenschappelijk onderzoek naar de taxonomie van ziekten, zo belangrijk voor diagnostiek en besluitvorming, staat nog in de kinderschoenen. Dagelijks worden we geconfronteerd met ziektegevallen die zich niet eenduidig laten categoriseren en telkenmale blijkt hoe gebrekkig de bestaande ziekteclassificaties zijn, ook voor dagelijks voorkomende problematiek. Deze tekortkomingen kunnen aanleiding zijn tot diagnostische misvattingen, met als gevolg ineffectief, overbodig en soms zelfs schadelijk handelen, zoals moge blijken uit het volgende voorbeeld.

Op een recente nascholingscursus over schouderaandoeningen kwam een patiënt ter sprake, die wegens een subacromiaal inklemmingssyndroom een endoscopische acromionplastiek had ondergaan, een ingreep waarbij de onderzijde van het acromion wordt gefreesd en het ligamentum coraco-acromiale wordt gekleefd. Na de ingreep verbeterden de klachten geenszins. Bij arthroscopie bleek in tweede instantie een uitgebreide scheur van het labrum glenoidale, de kraakbeenlip die verantwoordelijk is voor een belangrijk deel van de stabiliteit van de schouder.

De les hieruit is dat het klachtenpatroon van het ene ziektebeeld (subacromiaal inklemmingssyndroom) bijna zoveel gelijkenis

vertoont met het klachtenpatroon van een ander ziektebeeld (instabiliteit), dat verkeerde beslissingen zijn genomen. Indien hiervoor een eenduidig, betrouwbaar en geldig classificatiesysteem zou zijn ontwikkeld, getoetst aan de empirie, had dit wellicht voorkomen kunnen worden. Overigens: in de ziekteclassificatie van de NHG-Standaard Schouderklachten – het huidige referentiekader voor de huisarts – wordt nog niet eens gerept van instabiliteit als oorzaak van schouderklachten!

Het belang van ziekteclassificaties

Gebrekkige ziekteclassificaties zijn noch exclusief voor de huisartsgeneeskunde, noch voor het bewegingsapparaat. Wat te denken van de dyspepsieën, die nu vagelijk worden ingedeeld in 'ulcus-like', 'reflux-like' en 'motility-like'. Een goed classificatie-onderzoek zou de geldigheid van deze indeling aan kunnen tonen, alsmede de diagnostische criteria en de afbakening definitief kunnen vaststellen. Vervolgens zou men per groep een therapie-vergelijkend onderzoek kunnen doen (onderzoek naar de externe validiteit: heeft de indeling zin met het oog op het instellen van specifieke therapie).

Ieder medisch (deel)gebied kent zijn indelingsproblemen. Het is opvallend hoe gemakkelijk praktiserende artsen hierover heenstappen en zich beperken tot interesse voor de resultaten van clinical trials. Het volgende citaat gebruikte ik ook in mijn dissertatie: 'Development of epidemiological research on epilepsy has been relatively recent. The published results are often discordant, even in simple descriptive studies, because of the lack of agreement on the most basic concepts. Definitions of epilepsy, seizures and independent variables are often not stated. The classification of seizures and epileptic syndromes proposed by the International League Against Epilepsy is either not used or used incorrectly.'

Van Praag geeft ons voor de huisartsgeneeskunde een wijze les, als hij zijn mening over het DSM-III-R-classificatiesysteem geeft: 'We zijn bij de indeling van ziekten veel te hard van stapel gelopen. We zouden de huidige ziektebeelden eens goed moeten valideren en in al hun dimensies analyseren.' Het moet toch voor een ieder duidelijk zijn dat bij gebrekkige ziekteclassificatie en nomenclatuur de praktische betekenis van clinical trials betrekkelijk is. De laatste zin van de bespreking ('Het boek wordt vooral aanbevolen voor geïnteresseerden in classificatieproblematiek') vraagt daarom om een reactie. Dit proefschrift bevat mijns inziens een boodschap die juist voor iedere praktiserende huisarts waardevol is. Het is de huis-

arts die de hoogste prioriteit stelt aan het waarnemen van symptomen en verschijnselen. Maar juist hem ontbreekt het soms aan mogelijkheden om deze waarnemingen in een goed en valide diagnostisch systeem te plaatsen. Zo'n systeem is een goede en betrouwbare (gestandaardiseerde, reproduceerbare, dus empirisch getoetste) ziekteclassificatie; zonder classificatie geen (praktische) geneeskunde, geen diagnostiek, geen behandeling. Het objectief en reproduceerbaar classificeren met behulp van numerieke taxonomie zou juist nu, met de toenemende mogelijkheden van informatie-opslag en -verwerking, en nu het vak huisartsgeneeskunde in volle ontwikkeling is, de hoogste prioriteit kunnen hebben.

Cyriax

Schuurman stelt dat door mij de vier belangrijkste klinische beelden zijn bewerkt en dat een en ander conform de theorieën van Cyriax zou zijn. Welnu, dit is een misvatting. Hoewel men bij een theoretisch model er niet aan komt bestaan theorieën als uitgangspunt te gebruiken, is het de opzet geweest bij de theoretische classificatie vooringenomenheid te beperken. In de inleiding is duidelijk aangegeven dat juist de twijfels over de 'methode Cyriax' een motief zijn geweest voor nader onderzoek. Uit de uitvoerige literatuurstudie kwam het volgende naar voren:

- een uniforme nosologie (ziekteclassificatie) van de schouder bestaat niet;
- vier 'ziektebeelden/syndromen', door diverse auteurs verschillend beschreven en benoemd, lopen als een rode draad door de nosologie van de schouder;
- de theorievorming over etiologie, pathogenese, klinisch beeld, diagnostiek, meest effectieve behandeling en de nomenclatuur is zeer divers, waarbij de mening van Cyriax vaak haaks staat op de gangbare inzichten;
- belangrijke begrippen uit de methode-Cyriax zijn zelden origineel en exclusief.

Mijn exercitie, die door het Praeventiefonds gesubsidieerd werd, was er juist op gericht de ontbrekende nosologie te ontwikkelen. Het model-Cyriax kreeg daarbij een even groot 'gewicht' als de andere modellen. De theoretische classificatie is een rationele compilatie van literatuur- en ervaringsgegevens geworden en zeker niet uitsluitend gebaseerd op eigen waarneming.

Voorts is het model onderwerp van toetsing geweest. Daarbij is het belangrijk te beseffen dat ziekten onderwerp van studie zijn geweest en niet patiënten. Het ging erom hoe ziekten (in ieder geval de bovengenoemde grote vier) zich

bij de patiënten presenteerden in de vorm van symptomen/symptoomaspecten en verschijnselen. Voor klinisch-wetenschappelijk onderzoek is het van groot belang, dat deze symptomen meet- en waarneembaar zijn. Men kan zich uiteraard afvragen of bij deze groep van 89 personen alle relevante (aan deze soort ziekten gerelateerde) symptomen waarneembaar en/of meetbaar zijn. Dat weten we niet, maar gezien de beperkte mogelijkheden van het lichaam om uitdrukking te geven aan pathologische processen, denk ik dat met de gekozen 36 variabelen het belangrijkste deel van de waarneembare fenomenen is gegeven. Overigens bevatte de oorspronkelijke onderzoeksoplijst een veelvoud aan variabelen, maar vele bleken irrelevant, zodat een weliswaar arbitraire, maar naar mijn mening relevante selectie is gemaakt. Het zal duidelijk zijn dat, indien men de gegevens van een onderzoekspopulatie voor dit doel gebruikt, een verdeling van de patiëntselectie niet aan de orde is.

Classificatie voor de praktijk

Het is mij niet duidelijk wat bedoeld wordt met een classificatie voor wetenschappelijke doeleinden en de classificering voor het dagelijks handelen. Mogelijk speelt hier de verwarring tussen classificaties ten behoeve van de diagnostiek (taxonomieën) en classificaties ten behoeve van de administratie van gezondheidsproblemen (zoals ICD en ICPC). Wellicht ten overvloede zij opgemerkt dat het in mijn proefschrift gaat om het eerste type. Het eindprodukt van classificatie-arbeid volgens het paradigma van Loevinger en Skinner is een op wetenschappelijke wijze verkregen classificatie - dus interessant voor de wetenschapper maar evenzeer geldig en nuttig voor de praktijk. Dat men deze gebruikersvriendelijker kan maken door zeldzame syndromen, symptomen/symptoomaspecten en verschijnselen weg te laten of minder prioriteit te geven, betekent niet dat een classificatie voor de praktijk een geheel andere is. Een onderscheid tussen wetenschappelijke en dagelijkse classificatie is mijns inziens zinledig. Wel kan men een interne validiteit en een externe validiteit van een classificatie onderscheiden en onderzoeken.

Het resultaat, de nieuwe, getoetste ziekteclassificatie, is in het proefschrift bewust niet weergegeven. Wellicht zal hierover op korte termijn, in samenwerking met het NHG, gediscussieerd en gepubliceerd worden. Classificeren is een proces en 'clusteranalysis is the beginning of a classification project and not the end'. De gevonden taxonomie moet verklaard worden vanuit de pathologie; vervolgens kan

de externe validiteit getoetst worden aan de hand van het beloop en reactie op therapie.

Achteraf bezien ware het voor de betrouwbaarheid van de metingen inderdaad beter geweest, als de inter-observervariatie gemeten zou zijn. De belangrijke, meetbare variabelen zijn echter wel vastgesteld met een gevalideerd meetinstrument, de EDI 320 hoekmeter. Een meetfout van enkele graden zal nauwelijks consequenties hebben voor de interpretatie van de resultaten. Het vaststellen van binaire variabelen, zoals wel of niet een vast eindgevoel, wel of niet beperking van de passieve horizontale adductie, laat meer ruimte voor eigen interpretatie. Hier is gezorgd voor deskundige instructie van de onderzoekende studenten vooraf en lijkt de kans op fouten klein.

Een cumulatieve literatuurlijst is in verband met de omvang van het drukwerk achterwege gelaten; exemplaren zijn bij mij op aanvraag verkrijgbaar. De systematiek van de literatuurschrijver is vooral gericht geweest op het elimineren van niet-oorspronkelijke artikelen. Zoals mijn stelling al aangeeft, worden veel beweringen over schouderaandoeningen over- en nageschreven. Het vinden van onderzoeken waaruit oorspronkelijke beweringen zijn voortgekomen, is een eerste doel geweest. In mijn historisch overzicht zijn deze besproken. Op de combinatie van de sleutelwoorden 'schouderklachten/ziekten/symptomen' en 'general practice' werden bij Medline geen relevante artikelen gevonden. De meest bruikbare referenties kwamen altijd uit de oorspronkelijke artikelen en uit de hand- en leerboeken.

Vervolgonderzoek

Het is mijn hoop dat dit proefschrift heeft bijgedragen aan hernieuwde belangstelling voor het onderwerp. Het 'relatief onontgonnen terrein van de numerieke taxonomie' kan in samenwerking met de biologische wetenschappen het wetenschappelijk onderzoek in de huisartsgeneeskunde een nieuwe impuls geven. Een eerste vereiste lijkt mij om hetzelfde onderzoek in een grotere populatie te herhalen en zo meer zicht te krijgen op de stabiliteit van de clusters. Daarbij zou nog eens kritisch kunnen worden gekeken naar de variabelenlijst en aandacht besteed kunnen worden aan de inter-observervariatie. Ook het nader onderzoeken van gevonden clusters op beloop en reactie op therapie staat bij mij hoog in het vaandel. Misschien dat dan de interesse is gewekt om een ander onderwerp op dezelfde wijze onder de loep te nemen.

A.C. de Jongh