

# De Likelihood Ratio en de Unlikelihood Ratio

## Twee praktische maten voor diagnostische tests

N.P. VAN DUIJN

**Bij de individuele diagnostiek gaat het om twee praktische vragen: toont een diagnosticum met zekerheid een ziekte aan en sluit een negatieve uitslag van een test de ziekte met zekerheid uit? In deze bijdrage wordt een lans gebroken voor de Likelihood Ratio als maat voor het aantoonend vermogen van een test. Voor het uitsluiten van ziekte wordt een nieuwe maat voorgesteld: de Unlikelihood Ratio. De voordelen van deze twee maten zijn de eenvoud van interpretatie en toepassing bij de individuele diagnostiek, de congruentie met veel gebruikte maten als het Relatieve Risico en de Odds Ratio in de prognostiek en therapie, en vooral de aansluiting bij de kwalitatieve schattende aard van de dagelijkse diagnostiek.**

**Van Duijn NP. De Likelihood Ratio en de Unlikelihood Ratio. Twee praktische maten voor diagnostische tests. Huisarts Wet 1989; 32(12): 478-82.**

N.P. van Duijn, huisarts te Almere, Instituut voor Huisartsgeneeskunde, Universiteit van Amsterdam, Meibergdreef 15, 1105 AZ Amsterdam.

### Inleiding

Een maat voor de waarde van een diagnostische test kan drie doelen dienen:

- een efficiënte controle op de conclusie die men uit zijn gegevens trekt;
- een hulpmiddel bij het vaststellen van een standaard of het invoeren van een diagnosticum, bijvoorbeeld binnen een waarneemgroep;
- een handvat voor de medicus practicus die op het spreekuur moet beslissen of hij een bepaald diagnosticum toepast en wat de uitslag voor deze ene patiënt betekent.

*Knottnerus en Volovics* behandelen in hun aflevering 'Het onderscheidend vermogen van diagnostische tests'<sup>1</sup> een aantal maten vooral in het licht van de eerste twee doelen. De huisarts die wil weten wat een bepaald diagnosticum precies waard is bij een individuele patiënt, heeft daarentegen het derde doel voor ogen.

Deze bijdrage is geen pleidooi om de inhoud van *Huisarts en Wetenschap* de vorm te geven van hapklare brokken. Wel wordt de vraag aan de orde gesteld of bij een selectief gebruik van de beschikbare maten voor de waarde van een diagnosticum alle drie de doelen benaderd kunnen worden.

### Voor- en nadelen

De beschikbaarheid van een groot aantal maten kan van belang zijn voor een wetenschappelijke beoordeling, omdat elke maat zijn voor- en nadelen heeft, zoals *Knottnerus en Volovics* uitvoerig beschrijven. Het gaat bij deze beoordeling om de werkzaamheid van een diagnosticum en dan is enige abstractie geen bezwaar.

Ook diegenen die zich bezighouden met het vaststellen van standaarden, zullen zich verdiepen in de betekenis van de gebruikte maten. Bij deze beoordeling van de literatuur zijn, naast de werkzaamheid, vooral de doeltreffendheid en doelmatigheid van het diagnosticum van belang. Hiervoor zijn andere maten nodig dan voor de beoordeling van de werkzaamheid alleen. Te denken valt aan de voorspellende waarde, die eigenlijk alleen voor dit soort werk geschikt

is, of aan een ROC-curve bij meta-analyses en bij niet-dichotome tests.

Als dan uiteindelijk de werkzaamheid van een diagnosticum voldoende wordt geacht en bovendien de doeltreffendheid en de doelmatigheid groot genoeg zijn om de test aan te raden bij de beroepsgroep, dan rest het derde doel, het handvat voor de praktijk. Dit zou de doeltreffendheid op het niveau van de individuele patiënt genoemd kunnen worden.

Wie wil weten hoe een testuitslag waard is bij de individuele patiënt, zal eerst een schatting moeten maken van de prior kans op ziekte bij de individuele patiënt. Bij een dergelijke schatting zal nogal eens een subjectieve beoordeling nodig zijn en worden veel moeilijk meetbare factoren meegewogen. 'Dat verschilt per patiënt' is bijna een staande uitdrukking in de geneeskunde. Het construeren van een diagnostische maat die zowel iets zegt over de prior kans als over de waarde van de test, is dan ook ondoenlijk.<sup>2</sup> De toepassing van wetenschappelijke kennis in de praktijk vraagt dus om een maat voor het diagnostisch vermogen van een test die onafhankelijk is van de prior kans.

Daarnaast zou zo'n maat aan moeten sluiten bij het diagnostisch denken. Niemand kan van alle diagnostica de precieze waarde onthouden; daarvoor verloopt het diagnostisch proces te snel. Het stellen van een voorlopige diagnose zal bijvoorbeeld niet meer dan enkele seconden duren.<sup>3 4</sup> De te onthouden maat voor de waarde van een test zou hierbij aan moeten sluiten. Veronderstellen dat huisartsen met nomogrammen voor alle symptomen en diagnostica op zak lopen, is niet realistisch, hoezeer *Sackett et al.* dat ook bepleiten.<sup>5</sup>

Een vierveldentabel is weliswaar zeer informatief, maar vergt te veel hoofdrekenswerk om toegepast te kunnen worden. Bovendien hoeft de prevalentie van de aandoening in het betreffende wetenschappelijke onderzoek beslist niet gelijk te zijn aan de prior kans op die aandoening bij de individuele patiënt. De vierveldentabellen van individuele patiënten zullen er steeds weer anders uitzien. De voorspellende waarde valt om diezelfde reden af als getal om te onthouden.

De sensitiviteit en de specificiteit zouden wel goede kandidaten kunnen zijn om als handvat te dienen in de individuele diagnostiek, omdat ze meestal voldoende onafhankelijk zijn van de prior kans op ziekte. Een probleem is dat deze twee maten niet toegesneden zijn op de twee vragen die een huisarts zich vaak stelt:

- Is een symptoom een krachtige voorspeller voor de aanwezigheid van een ziekte (bijvoorbeeld het palperen van een mammatumor)?
- Sluit het ontbreken van een symptoom een ziekte voldoende zeker uit (bijvoorbeeld het niet opgeven van purulent sputum bij het uitsluiten van een pneumonie)?

Waar wordt de sensitiviteit of de specificiteit te laag om van een goede aantoonende c.q. uitsluitende test te kunnen spreken? De meeste diagnostica zijn immers zelden én goede aantoners én goede uitsluiters.

### Likelihood Ratio

Naast de snelheid van het diagnostisch proces speelt de aard van dat proces een rol. Dit diagnostisch denken is nauwelijks kwantitatief; artsen denken veeleer in kwalitatieve termen als goed, beter en 'zegt niets'. Wie artsen vraagt kwalitatieve termen als 'frequent' en 'onwaarschijnlijk' in een kanspercentage uit te drukken, meet bij de helft van de artsen een kans van 50-75 procent voor 'frequent' en 5-15 procent voor 'onwaarschijnlijk'. De andere helft geeft zelfs hogere of lagere kansen op.<sup>6</sup>

Het kenmerkende van huisartsgeneeskundige diagnostiek is het efficiënt uitsluiten van ziekte en heel af en toe het aantonen van ziekte. Hierbij wordt eerder geschat en beoordeeld dan gemeten en berekend, en dat is maar goed ook.<sup>6</sup>

De Likelihood Ratio lijkt hier het beste bij te passen. Deze kan op een schaal van 1

tot oneindig dienen als een continue maat voor het aantoonend vermogen van een diagnosticum. Mathematisch kan de Likelihood Ratio ook een waarde tussen 0 en 1 hebben, maar als het goed is, is deze test dan niet de stations van de wetenschappelijke beoordeling en de Standaardencommissies gepasseerd. Bij een Likelihood Ratio van 1 is de test van geen waarde (het voorbeeld van *Knottnerus en Volovics* over 'het hebben van een neus voor de diagnose Diabetes'). Bij een waarde tussen de 1 en de 2 is de test een zwakke aantoner, maar we zullen het er vaak mee moeten doen, omdat veel symptomen niet veel krachtiger zijn, zoals uit de voorbeelden van *Knottnerus en Volovics* blijkt. Het heeft zin de waarde van een Likelihood Ratio te onthouden in termen van een zwakke aantoner (1-2) of een redelijke aantoner (3-10), of een krachtig aantoonende test (>10).

De precieze getallen behoeven we niet te onthouden, want ze hebben marginaal invloed op de kwaliteit van de diagnostische beslissing. Voor de waarden van sensitiviteit en specificiteit ligt dat anders; het voorbeeld van de waarde van ronchi bij de diagnostiek van een pneumonie en de palpatie van een mammatumor laten dat zien (*tabel*). Bij de sensitiviteit en specificiteit is de *verhouding* tussen deze twee van groot belang en eigenlijk moet men mentaal steeds weer een vierveldentabel tekenen om zich een voorstelling te kunnen maken van de betekenis van deze twee getallen.

Een eenvoudige maat voor het uitsluiten van een ziekte is mij niet bekend. Wel vermelden *Knottnerus en Volovics* de Likelihood Ratio voor een negatieve uitslag van een test, wat een maat is voor het missen van de diagnose bij afwezigheid van bijvoorbeeld een symptoom. Maar dat wil een arts niet weten; hij zoekt zekerheid en geen onzekerheid in een definitie met een dubbele ontkenning. *Ridderikhoff* heeft laten zien dat de huisarts na het stellen van een voorlopige diagnose vooral argumenten vóór een diagnose zoekt.<sup>3</sup> Evenzo zal eerder gezocht worden naar argumenten vóór het uitsluiten van een ziekte, dan naar argumenten die de voorlopige diagnose onzeker maken.

Bovendien loopt deze Likelihood Ratio voor een negatieve uitslag exponentieel

Tabel De waarden van een aantal maten voor het onderscheidend vermogen van enkele diagnostica.

	Sens.	Spec.	pV+	pV-	LR-	LR	ULR
<i>Stemvorkproef volgens Rinne: pV bij een a priori kans van 30%</i>							
20 dB	69%	93%	81%	88%	0,33	9,8	3
25 dB	87%	93%	84%	94%	0,14	12,4	7,2
<i>Screening mammacarcinoom &gt;50 jaar: pV bij een a priori kans van 0,1%</i>							
Palpatie	41%	97%	0,04%	98%	0,61	14	1,6
Mammografie	95%	99%	0,1%	99%	0,05	95	20
<i>Symptomen bij een pneumonie: pV bij een a priori kans van 10%</i>							
Purululent sputum	65%	52%	13%	93%	0,67	1,4	1,5
Ronchi	19%	93%	23%	91%	0,87	2,7	1,1
Demping	4%	99%	52%	90%	0,97	4	1
Koorts	27%	94%	33%	92%	0,78	4,5	1,3
<i>Zure fosfatase, prostaatfractie bij prostaatcarcinoom: pV bij een a priori kans van 5%</i>							
Geen metastasen	38%	80%	9%	96%	0,78	1,9	1,3
Metastasen	89%	64%	12%	99%	0,17	2,5	5,8

van 0 naar 1 en een dergelijke wiskundige kromme met waarden achter de komma sluit niet erg aan bij het kwalitatieve diagnostische denken van artsen.

### Unlikelheid Ratio

In dit verband zou de Unlikelheid Ratio – specificiteit/(1-sensitiviteit) – een bruikbare maat kunnen zijn. In feite gaat het om de reciproke van de Likelihood Ratio voor een negatieve uitslag. De Unlikelheid Ratio geeft de mate aan waarin een ziekte wordt uitgesloten door het ontbreken van een symptoom of een normale uitslag van een nader onderzoek.

Ook de Unlikelheid Ratio loopt van 1 tot oneindig en heeft praktisch gezien geen betekenis bij een waarde tussen de 0 en de 1.

In de tabel zijn van een aantal symptomen en testuitslagen de gebruikelijke maten weergegeven. In de laatste twee kolommen staan de Likelihood en de Unlikelheid Ratio vermeld. Uit deze laatste twee kolommen is af te lezen welke test vooral een goede aantoner is en welke ook (of vooral) een goede uitsluiters is. De gegeven waarden bij verschillende afkappen bij de stemvorkproef van Rinne laten zien, dat het tweetal Likelihood en Unlikelheid Ratio ook bruikbaar is op een interval-schaal.

Deze twee Ratio's zijn wiskundig congruent, in tegenstelling tot de Likelihood Ratio voor een positieve en een negatieve uitslag en in tegenstelling tot het tweetal sensitiviteit en specificiteit.

Een extra voordeel is dat deze Likelihood en Unlikelheid Ratio precies dezelfde vorm en inhoud hebben als een andere veel gebruikte maat in de epidemiologie, het Relatieve Risico. Dit Relatieve Risico wordt veel gebruikt in prognostisch onderzoek en is direct toepasbaar in de spreekkamer bij het schatten van risico's. Het Relatieve Risico op longkanker bij rokers is bijvoorbeeld 10, wat betekent, dat rokers tien keer zoveel kans hebben op longkanker als niet rokers. Voor de toepassing van deze waarde van het Relatieve Risico bij voorlichting aan patiënten is in feite geen enkele berekening nodig. Ook bij onder-

zoek naar de waarde van een therapie worden vaak dit soort ratio's gebruikt.

Onlangs is een nieuwe manier beschreven om epidemiologische maten een vorm te geven die meer zou aanspreken: hoeveel patiënten moeten behandeld worden om één patiënt de ziekte te besparen?<sup>11</sup> Het gaat hier om de praktische formulering van de reciproke van de *absolute risk reduction*, ook wel *attributable risk reduction* genoemd.<sup>12</sup> In een meta-analyse van het effect van chemotherapie bij borstkanker wordt een vergelijkbare absolute maat gebruikt, zonder de reciproke te nemen, de *odds ratio reduction*.<sup>13</sup>

Op het eerste gezicht lijkt het aantrekkelijk resultaten van onderzoek te vertalen in aantallen te behandelen (of te diagnostiseren) patiënten, of in een reductie van het aantal sterfgevallen. Het nadeel is dat deze maten gegeneraliseerde incidenties bevatten en dus weer een vertaling noodzakelijk maken voor individuele patiënten die heel afwijkende prior kansen op ziekte kunnen hebben. Deze maten kunnen wel zeer geschikt zijn voor de beoordeling van de doelmatigheid van een test, bijvoorbeeld voor het vaststellen van een onderste leeftijdsgrens bij screening op cervixcarcinoom. Voor de toepassing in de individuele curatieve zorg lijken ze minder geschikt.

De door *Knottnerus en Volovics* genoemde Odds Ratio vat de Likelihood Ratio voor een positieve en die voor een negatieve uitslag samen. Voor de individuele curatieve zorg lijkt het me echter geen geschikte maat, omdat de richting van het diagnostisch vermogen verdwenen is. Een vooral krachtig aantonend diagnosticum en een tweede vooral sterk uitsluitend diagnosticum kunnen bijvoorbeeld dezelfde Odds Ratio hebben.

### Conclusie

Wellicht valt te overwegen om, naast de gebruikelijke maten voor het vermogen van een diagnostische test ten dienste van een efficiënte wetenschappelijke communicatie, ook rekening te houden met de toepassing bij de individuele diagnostiek. De Likelihood Ratio als maat voor het aantonend vermogen van een test en de Unlike-

hood Ratio voor het uitsluitend vermogen zouden een manier zijn om aan een aantal voordelen te combineren: een exacte weergave van de resultaten van het onderzoek, congruentie tussen de twee maten, congruentie met andere veel gebruikte maten als het Relatieve Risico en andere ratio's, aansluiting bij schattenderwijs en kwalitatief diagnostiseren, en vooral een onderscheid tussen het het aantonen en het uitsluiten van een aandoening. Dit laatste, het uitsluiten van ziekte, is zo zeer een kenmerk van de huisartsgeneeskunde, dat de toevoeging van alweer een nieuwe maat als de Unlikelheid Ratio misschien wel een voordeel is.

Overigens is recent door *Goodman* in een geheel andere context een vergelijkbare suggestie gedaan.<sup>14 15</sup>

### Literatuur

- Knottnerus JA, Volovics A. Het onderscheidend vermogen van diagnostische tests. *Huisarts Wet* 1989; 32: 338-46.
- Knottnerus JA. Interpretatie van diagnostische gegevens [Dissertatie]. Maastricht: Rijksuniversiteit Limburg, 1986.
- Ridderikhoff J. Decision-making strategies in the general practice [Dissertatie]. Rotterdam: Erasmus Universiteit Rotterdam, 1986.
- Barrows HS, Norman GR, Neufeld VR, Freightner JW. The clinical reasoning of randomly selected physicians in general medical practice. *Clin Invest Med* 1982; 5: 49-55.
- Sackett DL, Haynes RB, Tugwell P. *Clinical epidemiology*. Boston, Toronto: Little, Brown and Company, 1985.
- Axelrod S, Makao MA. How medical professionals evaluate expressions of probability. *N Engl J Med* 1987; 316: 549-50.
- Browning GG, Swan IRC. Sensitivity and specificity of Rinne tuning fork test. *Br Med J* 1988; 297: 1381-2.
- De Vries GP. A study of prostatic acid phosphatase in serum and tissue of patients with prostatic carcinoma [Dissertatie]. Amsterdam: Universiteit van Amsterdam, 1987.
- Roebuck EJ. Mammography and screening for breast cancer. *Br Med J* 1988; 297: 1381-2.
- Diehr P, Woord RW, Bushyhead J, et al. Prediction of pneumonia in outpatients with acute cough – a statistical approach. *J Chron Dis* 1984; 37: 215-25.
- Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. *N Engl J Med* 1988; 318: 1728-33.
- Friedman GD. *Primer of epidemiology*. 3rd

ed. New York: McGraw Hill, 1987.

- <sup>13</sup> Early Breast Cancer Trialists' Collaboration Group. Effects of adjuvant Tamoxifen and of cytotoxic therapy on mortality in early breast cancer. *N Engl J Med* 1988; 319: 1681-92.
- <sup>14</sup> Goodman SN. Meta-analysis and evidence. *Controlled Clinical Trials* 1989; 10: 188-204.
- <sup>15</sup> Zucker D, Salim Y. The Likelihood Ratio versus the p value in meta-analysis: where is the evidence? *Controlled Clinical Trials* 1989; 10: 205-8.

## Naschrift

**De kern van het interessante betoog van Van Duijn is van nogal technische aard, en daarom zal het aan mij gevraagde commentaar in eerste instantie ook nogal technisch zijn. De achtergrond van zijn voorstel is een zekere onvrede met bestaande maten voor het onderscheidend vermogen van diagnostische tests, en daarop zal ik meer beschouwend ingaan.**

### Technisch commentaar

Van Duijn acht de Likelihoodratio (het aannemelijkheidsquotient) van een positieve testuitslag LR+ om diverse redenen een aantrekkelijke maat, en dat zullen velen van harte met hem eens zijn. Hij heeft echter moeite met de Likelihoodratio van een negatieve testuitslag (LR-), waarbij ik overigens niet al zijn argumenten even duidelijk vind. Wellicht ervaart hij het nog het meest als hinderlijk, dat het waardenbereik van LR+ (van 1 tot oneindig, indien tenminste de positieve uitslag voor ziekte pleit) zo verschilt van dat van LR- (van 1 tot 0, indien de negatieve uitslag tegen ziekte pleit). Zo'n verschillend waardenbereik kan inderdaad belemmeringen opwerpen voor een vlotte interpretatie van de betekenis van de betreffende grootheden en het is goed om op zo'n probleem te wijzen. Een vlotte, haast intuïtieve interpretatie is immers juist vaak de behoefte van de arts.

Van Duijn draagt ook een alternatief aan. Hij is van mening dat het invoeren van de zogenaamde 'unlikelighoodratio' of ULR (die niets anders is dan 1/LR-) een vooruitgang zou zijn. Dit nu betwijfel ik.

In de gebruikelijke situatie, waarbij LR+ en LR- worden bepaald of verstrekt,

weet je direct dat waarden groter dan 1 de kans op ziekte groter maken, en dat waarden lager dan 1 deze kans kleiner maken. Vergissingen zouden gemakkelijker kunnen plaatsvinden, indien LR+ en 1/LR- (ofwel de ULR volgens Van Duijn) worden gebruikt, beide met eenzelfde waardenbereik. Vervelender is dat 'positieve' testuitslagen soms onverwacht eerder tegen dan voor ziekte lijken te pleiten, zoals de afwijkende auscultatie in het artikel van Muris.<sup>1</sup> In de gebruikelijke benadering krijg je dan een LR+ met een waarde tussen 0 en 1, en derhalve een LR- met waarden groter dan 1. Indien het voorstel van Van Duijn gevolgd wordt, krijgt men in zo'n geval naast een LR+ <1, ook nog een ULR (1/LR-) <1, hetgeen de zaak onbedoeld gecompliceerd maakt.

Het belangrijkste bezwaar van de introductie van de 'LR+ en ULR constellatie' vind ik echter, dat deze niet te veralgemeniseren is naar niet-dichotome tests, en dat zijn er vele. De 'LR+ en LR- constellatie' is dat wel: men spreekt dan alleen niet van twee, maar van meer of soms zelfs oneindig veel mogelijke testuitslagen. Gedacht kan worden aan respectievelijk de indeling volgens Papanicolaou of de hoogte van het ST-segment. Het gaat dan niet meer om LR+ of LR-, maar om LR<sub>x</sub>. Per definitie loop de schaal dan van 0 naar oneindig, met als 'indifferente' uitslag die testwaarde(n) waarbij LR<sub>x</sub> = 1. Het is onlogisch en verwarrend om alleen in gevallen waarin een duidelijk afkappunt tussen positieve en negatieve uitslagen wordt gekozen, met ULR (1/LR-) in plaats van met LR- te gaan werken.

Ik ben het overigens wel met Van Duijn eens dat het een voordeel is dat zijn twee maten voor een positieve respectievelijk negatieve uitslag op eenzelfde schaal vergelijkbaar zijn, maar dat weegt mijns inziens niet op tegen de genoemde problemen.

Zonder verder op details in te gaan, zou ik nog willen opmerken dat ik Van Duijn niet volg als hij een parallel ziet tussen de 'LR+ en ULR' constellatie' en het gebruik van het relatief risico. Het relatief risico kent immers naast waarden  $\geq 1$  wel degelijk waarden tussen 0 en 1, bijvoorbeeld in het geval dat een bepaalde factor met een

verlaagd risico verbonden is. Er is dus wat het waardenbereik betreft, eerder een parallel te trekken met de 'LR+ en LR- constellatie'. Overigens is het, afgaand op de weergave in een vierveldentabel, nog meer op zijn plaats om het relatief risico (in de etiologie) te zien als parallel van de verhouding tussen de voorspellende waarden van de positieve respectievelijk de negatieve uitslag (in de diagnostiek).

### Achtergrond

De achtergrond voor het denkwerk van Van Duijn acht ik zeer belangrijk, en dit wordt nog eens geaccentueerd door het bovenstaande – helaas noodzakelijkerwijs – ingewikkelde verhaal. Vaak zijn diagnostische associatiematen moeilijk te interpreteren en niet goed te implementeren in de dagelijkse zorg, en schieten ze over het doel van de medicus practicus heen. Bovendien – en dat hebben wij in ons artikel laten zien<sup>2</sup> – is er een veelheid van mogelijke diagnostische maten, elk met eigen mogelijkheden en beperkingen, en dat maakt het er niet gemakkelijker op. Tenslotte is iedere maat, ook de Likelihood Ratio en de Odds Ratio, gevoelig voor allerlei, vooral selectieve invloeden en dus afhankelijk van de populatie waarin deze bepaald is. Men kan zich gemakkelijk van deze perikelen afmaken door te zeggen dat dit gebied nu eenmaal moeilijk en omvangrijk is (evenals andere terreinen in de geneeskunde) en dat problemen niet opgelost worden door er de ogen voor te sluiten. Men kan echter ook trachten meer begrijpelijke, meer aansprekende notatievormen te vinden voor de medicus practicus. Van Duijn's bijdrage schudt daartoe wakker en bevat stof tot nadenken. Ik denk dat we zijn suggestie ten aanzien van de Likelihood Ratio voorlopig in elk geval als volgt kunnen oppikken:

Bij gebruik van de Likelihood Ratio bij een dichotome test is de LR+ een maat voor het 'ziekte-aantonend' vermogen van een positieve uitslag. Wat betreft de negatieve uitslag is de LR- beschikbaar als maat voor het 'ziekte-uitsluitend' vermogen van een negatieve uitslag. Indien men het lastig vindt dat LR+ en LR- een ander, moeilijk vergelijkbaar waardenbereik hebben, dan kan men dit rekentechnisch oplossen door van de LR- de reciproke te nemen: men

## Acupunctuur

kan dan direct aflezen of er sprake is van een diagnosticum dat:

- of ziekte beter kan aantonen dan uitsluiten (LR+>1/LR-);
- of ziekte beter kan uitsluiten dan aantonen (LR+<1/LR-);
- of indifferent is op dit punt (LT+=1/LR-).

In dit geval lijkt het me het niet nodig en wellicht zelfs extra verwarrend om ook nog een nieuwe term voor 1/LR- in te voeren.

J.A. Knottnerus

<sup>1</sup> Muris JWM, Vaessen MHJ, Sturmans F. De waarde van auscultatie bij de diagnostiek van bronchus obstructie in de huisartspraktijk. *Huisarts Wet* 1987; 30: 272-274.

<sup>2</sup> Knottnerus JA, Volovics A. Het onderscheidend vermogen van diagnostische tests. *Huisarts Wet* 1989; 32: 338-46.

## De twijfel aan het oordeel

In het redactioneel commentaar bij de artikelenreeks van *Ter Riet e.a.* over acupunctuur - c.q. hun weinig positieve bevindingen dienaangaande - plaatst *Van Veen* een aantal vraagtekens bij het voorheen aan acupunctuur verleende 'voordeel' van de twijfel.<sup>1</sup> Een dergelijke opinie lijkt in de rede te liggen. De studie van *Ter Riet e.a.* bestaat uit een grondige en zorgvuldige literatuuranalyse die, conform de normen voor onderzoek binnen het natuurwetenschappelijke model, zonder meer als wetenschappelijk verantwoord valt te kwalificeren en die deswege alle waardering verdient. Toch is het de vraag of aan de hand van een dergelijk onderzoek inderdaad een verantwoord oordeel over acupunctuur mogelijk is.

Ten aanzien van die vraag staat vooreerst ter discussie of met de argumentatie van *Ter Riet e.a.* (i.e. dat 'er redelijk veel onderzoek is verricht', respectievelijk dat 'de effectiviteit van acupunctuur daalt, naarmate de kwaliteit van het onderzoek stijgt') het door hen aangegeven 'inductieprobleem' van de empirische wetenschapsmethode voldoende is afgedaan, zodat generalisatie van de bevindingen gerechtvaardigd is.<sup>2</sup> Hiertoe is in ogenschouw te nemen dat de betreffende studie eenzijdig uitgaat van een westerse, natuurwetenschappelijke optiek, zonder op enigerlei wijze acht te slaan op de specifieke geaardheid van de Chinese Geneeskunst ('Traditional Chinese Medicine' - TCM), respectievelijk het fundamentele verschil met de westerse geneeskunde. *Ter Riet e.a.* hebben de grondbeginselen en basisbegrippen van de TCM veelal te summier of onvolledig weergegeven, terwijl daarnaast een aantal elementaire TCM-principes in het geheel niet aan de orde komt. Daarenboven zijn de respectievelijke interpretaties primair binnen de westerse geneeskunde georiënteerd en als zodanig in eerste instantie gericht op het onderkennen van een (pathologisch-)anatomisch substraat. Het naarstig zoeken van een substraat voor de acupunctuurmeridianen is hiervan een goed voorbeeld, met name omdat daarmee volgsagen voorbij wordt gegaan aan de conceptuele context van de TCM-entiteiten en het wezenlijke bestek van de TCM.

Voorts kan worden gesteld dat de studie van *Ter Riet e.a.*, evenals eerdere compilaties,<sup>3 4</sup> in principe uitgaat van acupunctuur als additief-westerse behandelwijze voor 'ziekte'. Het westerse begrip van ziekte verschilt echter fundamenteel van de beschouwing omtrent ziekte in de TCM. Dit hangt direct samen met het verschil van visie op 'gezondheid'. Anders dan de westerse geneeskunde, waar gezondheid slechts een afgeleide van de status van afwezigheid van pathologie voorstelt, onderkent de TCM ge-

zondheid als een conditie op zichzelf, die specifiek bepaald wordt door intrinsieke krachten vanuit de mens zelf, op geleide van diens dynamische evenwicht met de omgeving. Ziekte wordt primair gezien als een functionele verstoring van de gezondheid, bij wijze van een tekortschieten van de intrinsieke, gezondheidbepalende krachten, in het kader van een falende adaptatie aan een gebeurtenis in de omgeving. De ziekte - in de zin van pathologisch-anatomisch substraat - wordt hierbij als een secundair fenomeen van deze functieverstoring aangemerkt. In tegenstelling tot de westerse geneeskunde, die de etiologie van ziekte beziet in de zin van een pathologisch agens dat de 'voorheen gezonde' mens van buitenaf overkomt, wordt in de TCM de oorzaak van ziekte allereerst binnen de mens gezocht: in een interieure verzwakking, waardoor het exterieure agens 'pathogeen' heeft kunnen worden.

Studies zoals die van *Ter Riet e.a.* staan door hun eenzijdige gerichtheid op het westerse ziektebegrip compleet bezijden de context van de TCM en gaan daarmee ten enenmale voorbij aan de (gestelde) eigenlijke geaardheid van acupunctuur - 'vrij vertaald' naar westers-medische begrippen: een methode voor beïnvloeding van de 'gezondheid' respectievelijk het herstel daarvan (acupunctuur als geneswijze).

Deze eenzijdige oriëntatie op het bestek van behandeling van 'ziekte' komt ook tot uiting in het feit, dat in de door *Ter Riet e.a.* geselecteerde acupunctuuronderzoeken merendeels sprake is van 'formule-acupunctuur', een wijze van acupunctuur die primair op het westerse ziektebegrip is gericht en waarbij specifieke diagnostiek en therapie conform het TCM-concept geheel en al ontbreken.

Een ander negatief punt in dit verband is de keuze voor omliggende, specifieke indicatiegebieden (ziektebeelden, pijnsyndromen) bij de studie, hetgeen van meet af aan een grove inhoudelijke beperking inhoudt. Binnen de TCM wordt de specificiteit van een beeld gezien in termen van de aard of wijze van de verstoring van gezondheid, met een navent scala van manifestaties, waarvan een 'symptoom' als bijvoorbeeld lage rugpijn slechts één (eventueel) onderdeel is.

In soortgelijke zin gelden bezwaren ten aanzien van het in de studie gekozen uitgangspunt van placebo-gecontroleerd acupunctuuronderzoek, met name in aanmerking genomen, dat binnen de TCM het begrip 'placebo' a priori geen betekenis heeft.

Daarnaast zijn methodologische kanttekeningen te plaatsen aangaande het model van het acupunctuuronderzoek met placebo-controle. De specifieke problemen omtrent de placebo-

## RECTIFICATIE

## Statistiek

In de aflevering 'Het onderscheidend vermogen van diagnostische tests' in de serie *Huisarts, wetenschap en statistiek* (*Huisarts Wet* 1989; 32(9): 338-46) zijn enkele fouten geslopen.

• Op pagina 339, eerste kolom moeten de regels 6 t/m 9 worden vervangen worden door de volgende tekst:

De bijbehorende aannemelijkheidsquotienten (A+) zijn respectievelijk:

- $A^+ = 100\% / (100\% - 0) = 1,0$  (voor 'het hebben van een neus' ten aanzien van diabetes) en
- $A^+ = 27\% / (100\% - 67\%) = 0,82$  (voor 'een verlengd expirium' ten aanzien van bronchusobstructie).

• In de toelichting bij tabel 6 (pag. 342) zijn de verklaringen van de velden c en d verwisseld.