

Een dilemma voor beslissingsondersteuning

*Steven F. Hartkamp
Medische Informatie- en Besliskunde
Academisch Ziekenhuis Groningen*

1. Inleiding.

In de dagelijkse geneeskundige praktijk worden door zowel artsen als patiënten veel beslissingen genomen. Beslissingen die vaak ingrijpende gevolgen voor het leven van patiënten kunnen hebben. Voor dergelijke beslissingen lijkt hulp onder andere in de vorm van een beslissingsondersteunend systeem geen overbodige luxe. Voor artsen is een dergelijk systeem zelfs eerder als noodzaak te beschouwen.

Hoewel een ondersteunend systeem een noodzaak lijkt, zal niet iedereen even enthousiast zijn over de introductie van deze beslissingsondersteunende systemen. Het enthousiasme kan enigszins bekoelen door de gedachte dat een technologische tang steeds verder de menselijke ontwikkeling afknijpt en verstikt. Het schrikbeeld dat hierbij voor ogen staat, is een geneeskunde waarin klinische oordelen zijn vervangen door output op een beeldscherm of printer. Dit beeld is niet alleen een kwaadaardig visioen van verontrusten, getuige het voorstel om gesprekken met psychiaters te vervangen door 'dialogen' via een terminal met een computer, als reactie op het programma ELIZA [9]. ELIZA is een programma dat door Weizenbaum is ontwikkeld in het kader van een onderzoek naar computeranalyse van natuurlijke taal. Het programma simuleert een nondirectieve therapeut. De mensen die een 'dialoog' aangingen beschouwden de computer als zo'n nondirectieve therapeut. Weizenbaum protesteerde overigens tegen deze voorstellen om klinische oordelen door computeroordelen te vervangen. Evenals Weizenbaum wijzen ook de ontwikkelaars van medische beslissingsondersteunende systemen een substitutie af [3,4]. Zij spreken dan ook van beslissingsondersteunende systemen en niet van beslissingssystemen. Zij wijzen erop dat het zeker niet de bedoeling is om artsoordelen door computeroordelen te vervangen. De arts blijft in hun ogen verantwoordelijk. Het is echter nog maar de vraag of die verantwoordelijkheid van de arts wel is te rijmen met de beslissingsondersteuning? Met andere woorden zijn er voorwaarden te formuleren voor het gebruik van beslissingsondersteunende systemen zodat er sprake kan zijn van meer dan alleen een formele toegeschreven verantwoordelijkheid? Hier zal ik een aanzet geven voor een beantwoording van deze vragen. Hiertoe zal ik eerst de bedoelde beslissingsondersteuning nader bekijken op grond waarvan een dilemma zichtbaar wordt. Dit dilemma, zo zal worden getoond, kan niet opgelost worden door computeroordelen te beschouwen als laboratoriumuitslagen. Door systeemoordeelen op te vatten als taalhandelingen worden de contouren zichtbaar van een oplossing.

2. Beslissingsondersteuning.

Beslissingsondersteuning door systemen lijkt het eenvoudigst gerealiseerd te worden door het genereren van een oordeel door die systemen. Enkele ontwikkelde systemen zijn op deze wijze ondersteunend. Maakt een arts gebruik van een dergelijk ondersteunend systeem dan zijn er twee mogelijkheden, namelijk:

- het systeem geeft altijd dezelfde oordelen als de arts,
- het systeem en de arts hebben soms een verschillend oordeel.

Met het spreken over 'oordelen' wil ik niet impliciet beweren dat een computer kan denken. Met 'dezelfde oordelen' bedoel ik zinnen die op het beeldscherm of via de printer verschijnen en zinnen die de arts uitsprekt of opschrijft en die voor toehoorders of lezers dezelfde betekenis hebben. Laten we de twee genoemde mogelijkheden achtereenvolgens bekijken.

2.1. Identieke oordelen.

De eerste mogelijkheid waarbij het systeem en de arts altijd dezelfde oordelen geven doen zich voor als:

1. het systeem een soort echo is van de arts en datgene herhaalt wat de arts zegt, óf
2. de arts een soort echo is van het systeem en de oordelen van het systeem direct overneemt, óf
3. de arts en het systeem onafhankelijk van elkaar tot hetzelfde oordeel komen.

Een systeem dat als echo fungeert zal men, hoewel de materiële verwerking ervan geen geringe technische prestatie is, geen medisch expertsysteem noemen. Men kan het systeem in zekere zin beslissingsondersteunend noemen. Het is namelijk niet uitgesloten dat een arts zich gesteund voelt als het eigen oordeel nog eens herhaald wordt. Het systeem is dan beslissingsondersteunend zoals een "minor tranquillizer" dat ook kan zijn, door gevoelens van onzekerheid te verminderen. We beschouwen echter deze middelen nu niet bepaald als beslissingsondersteunende middelen. Blijkbaar staat ons bij beslissingsondersteunende systemen iets anders voor ogen dan het verminderen van gevoelens van onzekerheid. Het lijkt me dat de ondersteuning die met computersystemen wordt beoogd, bestaat uit het leveren van extra evidentie voor een oordeel. Als de computer het oordeel geeft dan is dat, tenminste dat is de bedoeling, extra evidentie voor de juistheid van dat oordeel en is dit een extra reden om het oordeel aan te nemen. Een systeem dat als echo fungeert is in dit opzicht niet beslissingsondersteunend, want het geeft geen extra reden om het oordeel aan te nemen.

Het tweede geval waarin de arts en systeem altijd hetzelfde oordeel hebben, doet zich voor als de arts het systeemoordeel herhaalt. Deze situatie zal zich voordoen als de arts al enige tijd met het systeem heeft gewerkt en heeft geleerd dat indien hijzelf een oordeel zou geven dat oordeel minder goed zou zijn dan dat van het computersysteem. Deze situatie is secundair aan die waarin systeem- en artsoordeel soms van elkaar verschillen. De expertsysteemontwikkelaars zullen van mening zijn dat dit niet de bedoeling is en zij zullen dit echogedrag afwijzen. Ik zal deze situatie hier niet verder bespreken, maar er in de volgende paragraaf op terugkomen.

In het derde geval komen het systeemoordeel en artsoordeel altijd overeen en komen de arts en het systeem onafhankelijk van elkaar tot hetzelfde oordeel. Er is in dit geval sprake van een perfecte simulatie van het medisch oordelen van de arts. Vanuit "artificial intelligence"-oogpunt moet hier gesproken worden

van een groot succes, maar deze simulatie lijkt me niet het doel van beslissingsondersteunende systemen. Dergelijke simulerende systemen zijn veel meer te vergelijken met echosystemen. Het enige verschil is dat de artsoordelen niet herhaald maar voorspeld worden. Maar het voorspeld worden van een oordeel geeft geen extra evidentie voor de juistheid van dat oordeel. Simulerende systemen zijn dan ook geen beslissingsondersteunende systemen.

Samenvattend kunnen we zeggen dat in het geval de systeemoordelen altijd overeenkomen met artsoordelen, er geen sprake is van beslissingsondersteuning, of dat er sprake is van een ontwikkeling die niet wenselijk wordt geacht, namelijk echogedrag. Laten we de tweede mogelijkheid bekijken waarbij de systeemoordelen en artsoordelen van elkaar verschillend kunnen zijn.

2.2. Verschillende oordelen.

Stemmen artsoordeel en systeemoordeel niet altijd overeen, dan komt het soms voor dat er een verschil in oordeel is. Wat moeten we in zo'n geval doen? Het ligt voor de hand om af te gaan op de voorspellende waarde van zowel het systeem als de arts, ook als ikzelf die arts ben.

Stel nu eens dat mijn diagnostisch oordeel over het algemeen in 42% van de gevallen juist is en dat van het systeem in 91% [3]. De kans op het juiste oordeel is dan 91% als ik het systeemoordeel overneem. Deze kans is groter dan de kans op het juiste oordeel als ik op mijn eigen oordeel afga, die is namelijk 42%. Dus kan ik beter het systeemoordeel overnemen. Het systeem is beslissingsondersteunend in de zin van het geven van extra evidentie voor de juistheid van het oordeel. Deze redenering is gegeven de testresultaten toepasbaar op alle situaties waarin het systeemoordeel en artsoordeel van elkaar verschillen. Dit betekent dat de arts in al die gevallen het systeemoordeel over moet nemen. Dit is niet slechts een psychologisch gevolg van het werken met een dergelijk systeem. Het is een normatieve consequentie, er zijn goede redenen om het systeemoordeel altijd over te nemen. Juist omdat het systeem beslissingsondersteunend is met goede performatietesten, moeten we echogedrag gaan vertonen. Het niet overnemen van het systeemoordeel is tegen de regels van de redelijkheid en kan patiënten schaden. Maar dit categorisch overnemen van het oordeel was niet de bedoeling van de ontwikkelaars van het systeem.

Tegen bovenstaande redenering kan men inbrengen dat ze niet correct is en voorbijgaat aan de onzekerheid over de interpretatie van de testresultaten. Het kan bijvoorbeeld wel zo zijn dat in de geteste gevallen het systeem in 91% van de gevallen het juiste oordeel gaf, maar dit hoeft niet te betekenen dat dit ook voor alle niet geteste gevallen geldt. De kans op een juist oordeel bij het overnemen van het systeemoordeel kan best kleiner zijn dan bijvoorbeeld 42%. Dit kan zich voordoen vanwege het zogenaamde plateauverschijnsel van kennisrepresentatie. Een systeem geeft goede resultaten in een bepaald kennisgebied maar de resultaten worden plotseling slechter als er kennis wordt vereist die niet is gerepresenteerd [6]. Dit verschijnsel doet zich bij artsen in veel mindere mate voor. Er is dan ook een grotere spreiding van kansen op het juiste oordeel bij het overnemen van het systeemoordeel dan bij het vasthouden aan het artsoordeel. Voor keuzen waarin er een grote spreiding van kansen is wordt het volgende beslissingscriterium wel voorgesteld [5]:

Neem het oordeel aan afkomstig van die bron waarvan het minimum van de kans op het juiste oordeel, het grootst is.

Op grond van dit criterium is het te rechtvaardigen om het artsoordeel over te nemen in alle situaties waarin het systeemoordeel en artsoordeel verschillend zijn. Echter in dit geval kunnen we nog moeilijk spreken van beslissingsondersteuning.

Het lijkt erop dat we op een dilemma zijn gestuit namelijk:

- of het ontwikkelen van beslissingsondersteunende systemen leidt tot de onwenselijke situatie dat een arts het systeemoordeel categorisch overneemt,
- of er is geen sprake van beslissingsondersteuning.

Gegeven dit dilemma ziet het er somber uit voor de ontwikkeling van beslissingsondersteunende systemen waarbij de arts verantwoordelijkheid draagt. De vraag is hoe dit dilemma tegemoet te treden. Een weg die ik zal bewandelen is nagaan of het dilemma werkelijk onvermijdelijk is.

3. Systeemoordeel als laboratoriumuitslag.

Het geschetste dilemma wordt opgeroepen door de categorische beoordeling van systeemoordelen en artsoordelen. Met categorisch bedoel ik dat de redenen om het arts- of systeemoordeel over te nemen in situaties waarin de oordelen verschillen, er toe leiden, dat ik in alle situaties waarin zich deze incongruentie van oordelen voordoet, dezelfde beslissing moet nemen. Met een alternatief voor een dergelijke categorische beoordeling, is het dilemma wellicht te vermijden. Een mogelijkheid daartoe is, systeemoordelen opvatten als laboratoriumuitslagen [3]. Laboratoriumuitslagen zijn beslissingsondersteunend, ze verminderen onzekerheid omtrent diagnostische oordelen. Bovendien bij laboratoriumonderzoek interpreteert een arts de uitkomst en besluit al dan niet op het laboratoriumonderzoek te varen. Deze analogie zou een fraaie leidraad geven voor het beoordelen van systeemoordelen. Maar voordat we denken dat het dilemma is opgelost, moet eerst duidelijk zijn of achter de geschetste analogie geen belangrijke verschillen schuil gaan die de vergelijking van een laboratoriumuitslag met een systeemoordeel minder geslaagd maken.

Een laboratoriumonderzoek naar de waarde van een variabele Var met als uitslag ' $Var=x$ ', bijvoorbeeld ' $\gamma\text{-GT}=60 \text{ U/l}$ ', geeft op verschillende manieren evidentie, namelijk:

- de uitslag ' $Var=x$ ' is evidentie voor de juistheid van het oordeel dat $Var=x$.
- de uitslag ' $Var=x$ ' is evidentie voor oordelen van de vorm: 'patiënt heeft ziekte X.' of 'heeft baat bij therapie Y'.

Doelt men met de analogie tussen systeemoordelen en laboratoriumuitslagen op de eerste vorm van evidentie, dan is de analogie zo goed als volledig, echter ze biedt geen goede leidraad voor systeembeoordeling. Een oordeel gegenereerd door een beslissingsondersteunend systeem geeft evidentie voor de juistheid van dat oordeel. Dit is gelijk aan de eerste vorm van evidentie waarbij een laboratoriumuitslag ' $Var=x$ ' een reden vormt voor de juistheid van het oordeel dat inderdaad $Var=x$. Echter in de praktijk beschouwen artsen, ondanks hun kennis over betrouwbaarheid, laboratoriumuitslagen van de vorm ' $Var=x$ ' als evidentie voor oordelen van de vorm $Var=x$. Bijvoorbeeld, een uitslag ' $\gamma\text{-GT}=60 \text{ U/l}$ ' wordt als evidentie gezien voor het oordeel dat inderdaad $\gamma\text{-GT}=60 \text{ U/l}$. De analogie

zou dan leiden tot de categorische richtlijn om het systeemoordeel direct over te nemen, hetgeen juist niet de bedoeling was van de analogie.

De betrouwbaarheid van een laboratoriumuitslag 'Var=x' speelt niet in de eerste plaats een rol bij de vraag of Var=x inderdaad het geval is (deze vraag is meestal niet interessant), maar ze speelt samen met andere factoren een rol bij het bepalen van de mate van evidentie van de uitslag voor diagnostische en therapeutische oordelen. Soms is een uitslag extra evidentie, soms niet. Op deze tweede vorm van evidentie zal men doelen indien men aan een analogie tussen laboratoriumuitslagen en systeemoordelen een leidraad wil ontleenen. Maar achter deze analogie gebaseerd op de tweede vorm van evidentie gaat een relevant verschil schuil tussen laboratoriumuitslagen en systeemoordelen.

De evidentierelatie tussen laboratoriumuitslagen en diagnostische en therapeutische oordelen is gebaseerd op een pathofysiologische theorie die empirisch van aard is. De evidentierealtie tussen systeemoordeel en de juistheid van het oordeel is niet gebaseerd op zo'n pathofysiologische theorie maar op een informatie-verwerkingstheorie. Dit is een theorie waarin is weergegeven wat een correcte wijze is om informatie te verwerken en conclusies te trekken. Dit is geen empirische theorie over hoe artsen concluderen, maar het is een theorie met een normatief karakter en geeft weer hoe artsen zouden moeten concluderen. Een arts kan twijfels hebben over de evidentie die een laboratoriumuitslag γ -GT=60 U/l geeft voor de diagnose alcoholisme op grond van een pathofysiologische theorie die verklaart waarom deze waarde zo hoog is. Voor het in twijfel trekken van de juistheid van het systeemoordeel is er niet zo'n theorie voor handen. Hoe een arts dat moet beoordelen is niet helemaal duidelijk. Dit verschil in de aard van de theorie welke de laboratoriumuitslagen c.q. systeemoordelen evidentie doet zijn voor diagnostische oordelen, maakt dat de analogie met laboratoriumuitslagen niet goed bruikbaar is om een leidraad te geven voor het evalueren van systeemoordelen. Wel wordt duidelijker in welke richting een evaluatie zal gaan. Een evaluatie van systeemoordelen zal betrekking hebben op de informatieverwerkende theorie en de aard van de interpretatie van het oordeel. Het plaatsen van systeemoordelen in het perspectief van een theorie over taalhandelingen biedt wellicht een beter zicht op interpretatie van en verantwoordelijkheid voor systeemoordelen.

4. Systeemoordeel als taalhandeling.

De theorie van taalhandelingen vindt zijn oorsprong in de ideeën van L. Wittgenstein en J.L. Austin [2]. De achterliggende gedachte is dat we met het spreken ook handelingen verrichten en effecten tot stand brengen. Hierbij wordt niet op alle effecten gedoeld maar op die effecten welke direct samenhangen met de betekenis van de uitspraken. Er wordt bijvoorbeeld niet bedoeld dat ik door middel van het spreken via een geluidssensor een lampje laat branden. Het standaardvoorbeeld van wat wel wordt bedoeld, is het effect van het uitspreken van "Hierbij doop ik dit schip tot Stalin." waarmee het schip wordt gedoopt. Een ander voorbeeld is het effect van "Hierbij verklaar ik u tot man en vrouw." Men spreekt hier van illocutionaire effecten. Dit zijn effecten die direct door middel van de betekenis van de uitspraken tot stand komen. Deze theorie is uitgewerkt door J.R. Searle [8]. Hij geeft onder andere een omschrijving van een taalhandeling waarmee we een ander iets duidelijk willen maken, dat wil zeggen een communicatieve handeling. Onder een communicatieve handeling, 'iets zeggen en dat ook bedoelen', verstaat Searle:

Spreker S uit de zin T en bedoelt die ook (dat wil zeggen: bedoelt letterlijk wat hij zegt) en:

- S heeft met de uiting U van T de intentie (i-l) om hoorder H te doen weten (herkennen, beseffen) dat zich standen van zaken voordoen die door (bepaalde van) de regels van T worden gespecificeerd (Dit effect noemen we het illocutionaire effect IE)
- S heeft de intentie om IE met U teweeg te brengen door middel van de herkenning van i-l.
- S heeft de intentie dat i-l op grond van (door middel van) kennis van H van (bepaalde van) de regels die (de elementen van) T beregelen, zal worden herkend.

Vatten we nu systeemoordelen op als communicatieve handelingen waarmee iets wordt meegedeeld, dan impliceert dat, dat er een intentie moet zijn van een 'spreker', waarmee tevens de verantwoordelijkheid voor het oordeel bekend is.

Een belangrijk verschil tussen een oordeel uit een boek of van een geraadpleegde consultant wordt hiermee duidelijk. Een schrijver en consultant hebben de intentie om een stand van zaken met een taaluiting mee te delen. In het geval van computersystemen is de output niet zonder meer door iemand geïntendeerd. Want ook al volgt de output op de input dan mag men nog niet concluderen dat iemand die de input verzorgt, veelal de gebruiker, de output intendeert. Intenties kan men niet zonder meer zo uit elkaar afleiden. Men kan namelijk niet zeggen dat indien Jan de intentie heeft om uit te vinden of de avondster de morgenster is, en in werkelijkheid de avondster de morgenster is, dat Jan de intentie heeft om uit te vinden of de avondster de avondster is [7]. Het is namelijk niet uitgesloten (ja, zelfs zeer aannemelijk) dat Jan niet weet of de avondster de morgenster is. Evenzo kan aan een gebruiker niet de intentie voor het systeemoordeel toegeschreven worden ook al volgt de output uit de ingegeven input. Willen we dat wel kunnen, dan moeten we er redelijkerwijs vanuit kunnen gaan dat de gebruiker weet wat op de input volgt. Dat wil zeggen dat de gebruiker moet weten hoe het systeem werkt.

Als we met een perfect systeem van doen hebben dat correcte oordelen geeft, dan moeten we dus weten wat juiste oordelen zijn en wat een goede wijze van informatieverwerken is. Juist deze kennis ontbreekt veelal. Er wordt maar weinig onderzoek gedaan naar de elementen van een goede beslissing in de medische praktijk. Slechts als over deze elementen meer duidelijkheid bestaat kunnen we met enige redelijkheid intenties toeschrijven aan artsen, de gebruikers van deze systemen. Dit betekent dat we de arts, de

gebruiker, verantwoordelijkheid toe kunnen schrijven als deze op de hoogte is van een correcte informatie-verwerkingstheorie en het systeemoordeel gebaseerd is op die theorie. Is zo'n theorie niet bekend, dan kunnen we de systeemoordeelen beter onder een ander type taalhandeling rangschikken dan het communicatieve type. Ze komen dan meer overeen met het type dat Andersson de geformaliseerde sociale handelingen noemt [1]. In deze gevallen hoeft degene die het oordeel uit, het illocutionaire effect niet te intenderen. Het is voldoende dat in de taalgemeenschap wordt aanvaard dat voor het optreden van het illocutionaire effect slechts nodig is dat het oordeel wordt geuit, bijvoorbeeld bij een formele ceremoniële handeling als het openen van een weg, waarbij iemand op een of ander manier moet zeggen "Hierbij is de weg geopend". Voor medische konteksten betekent dit dat men medische oordelen als formele ceremoniële handelingen ziet, een visie die naar ik vrees, hier en daar valt te herkennen. De introductie van beslissingsondersteunende systemen past probleemloos in deze medische ceremonie. Het lijkt me dat men juist tegen dat ceremoniële toneelspel protesteert indien men artsoordelen niet wil vervangen door computeroordelen. Wil men waken tegen een ceremoniële geneeskunde en een gedeelte van het dilemma voor beslissingsondersteuning, het echogedrag van artsen, vermijden, dan moet men veel meer onderzoek verrichten naar een juiste informatieverwerkingstheorie waarin aandacht wordt geschonken aan de rechtvaardiging van medische oordelen. Dit betreft niet alleen de formeel logische component van klinische oordelen maar vooral de praktische component. Dit vereist onderzoek naar de menselijke waarden en keuzen welke in het systeem gerepresenteerd worden c.q. gerepresenteerd moeten worden.

Referenties:

- 1 J.S. Andersson. How to define 'Performative'. Uppsala, 1975.
- 2 J.L.Austin. How to do things with words. Oxford, 1975.
- 3 F.T. de Dombal. Computer-aided diagnosis of acute abdominal pain: the British experience. In: Professional judgment. J. Dowie & A. Elstein (eds.), Cambridge, 1988, p. 190-199.
- 4 M. Evenblij. Medische waakhonden. In: Intermediar, 1988,(24), nr. 10, p.47-51.
- 5 P. Gärdenfors & N.E. Sahlin. Unreliable probabilities, risk taking, and decision making. In: Decision, probability, and utility. P. Gärdenfors & N.E. Sahlin (eds.), Cambridge, 1988.
- 6 P.F. Lucas & A.R. Janssens. Medische expertsystemen: hulpmiddel bij diagnose en therapie. In: Nederlands Tijdschrift voor Geneeskunde, 1985,(129), p. 168-165.
- 7 P.A.M. Seuren. Tussen taal en denken. Utrecht, 1975.
- 8 J.R. Searle. Taalhandelingen. (vert.) Utrecht, 1977.
- 9 J. Weizenbaum. Computer power and human reason. Middlesex, 1985.