

Inhoud

	Voorwoord prof. dr. P.H. Dejonckere bij de eerste druk	10
	Woord vooraf bij de tweede, geheel herziene druk	12
I	Inleiding	14
1.1	Wat is evidence-based handelen?	14
1.2	Evidentie in de logopedie	16
1.3	Niveaus van evidentie	18
1.4	De PICO-vraag	21
1.5	Opbouw en gebruik van dit boek	22
2	Richtlijnen en systematische reviews	23
2.1	Inleiding	23
2.2	Richtlijnontwikkeling	23
2.3	Aanbevelingen	24
2.4	Het gebruiken van richtlijnen	27
2.5	Het beoordelen van richtlijnen	28
2.6	Systematische reviews	29
2.7	Publicatiebias	32
2.7.1	Is de vraagstelling die ten grondslag ligt aan de systematische review duidelijk?	32
2.7.2	Zijn de criteria voor de te includeren studies voor deze review expliciet beschreven?	33
2.7.3	Wordt de zoekstrategie reproduceerbaar beschreven?	34
2.7.4	Is de selectie van de te includeren artikelen onafhankelijk uitgevoerd?	35

2.8	Kwaliteitsverschillen	35
2.8.1	Is de kwaliteit van de geïnccludeerde artikelen reproduceerbaar en valide beoordeeld?	35
2.8.2	Is de kwaliteitsbeoordeling van de geïnccludeerde artikelen onafhankelijk uitgevoerd?	37
2.8.3	Is de data-extractie onafhankelijk uitgevoerd?	37
2.8.4	Is de datapresentatie eenduidig?	39
2.9	Heterogeniteit	41
2.9.1	Is de homogeniteit respectievelijk heterogeniteit van de studies gecontroleerd?	41
2.9.2	Worden de resultaten uit de afzonderlijke studies terecht gecombineerd?	42
2.9.3	Is een sensitiviteitsanalyse uitgevoerd?	43
3	Diagnostiek: beoordeling van een valideringsonderzoek	45
3.1	Inleiding	45
3.2	Validiteit van het onderzoek	48
3.2.1	Is er een valide referentietest gebruikt als gouden standaard?	49
3.2.2	Zijn de resultaten van de referentietest en de indextest onafhankelijk van elkaar beoordeeld?	49
3.2.3	Werd de indextest onafhankelijk afgenomen van andere relevante informatie over de werkelijke toestand van de patiënt?	50
3.2.4	Zijn de indextest en de referentietest beide bij alle patiënten afgenomen?	50
3.2.5	Is de referentietest toegepast voordat op basis van de resultaten van de indextest al een behandeling is gestart?	51
3.2.6	Was de selectie van patiënten voor het onderzoek valide?	52
3.3	Toepasbaarheid van de resultaten	52

3.3.1	Is de onderzochte indextest geschikt voor toepassing op uw eigen patiënten?	52
3.3.2	Is de test reproduceerbaar en kosteneffectief?	52
3.4	Belang van de resultaten	53
3.4.1	Diagnostische waarden in percentages	54
3.4.2	Diagnostische waarden in verhoudingen (ratio's)	61
3.5	Oefeningen	69
4	Therapie: beoordeling van een interventieonderzoek	72
4.1	Inleiding	72
4.2	Validiteit van het onderzoek	78
4.2.1	Heeft randomisatie plaatsgevonden?	78
4.2.2	Is sprake van een geblindeerde toewijzing van de interventie?	78
4.2.3	Zijn de groepen aan het begin van het onderzoek vergelijkbaar?	80
4.2.4	Wordt het intention-to-treat-principe gevolgd?	80
4.2.5	Zijn de uitkomsten expliciet beschreven?	81
4.2.6	Wordt de kwaliteit van de meetinstrumenten beschreven?	83
4.2.7	Is er sprake van een follow-up?	83
4.2.8	Is er sprake van blinding?	85
4.2.9	Zijn de beide groepen verder gelijk behandeld?	87
4.3	Toepasbaarheid van de resultaten	88
4.3.1	Is de in het onderzoek beschreven behandelsituatie uitvoerbaar in de eigen praktijk?	88
4.3.2	Zijn dezelfde resultaten bij de eigen patiënt te verwachten?	88
4.4	Belang van de resultaten	89
4.4.1	Grootte van het effect	89
4.5	Oefeningen	96

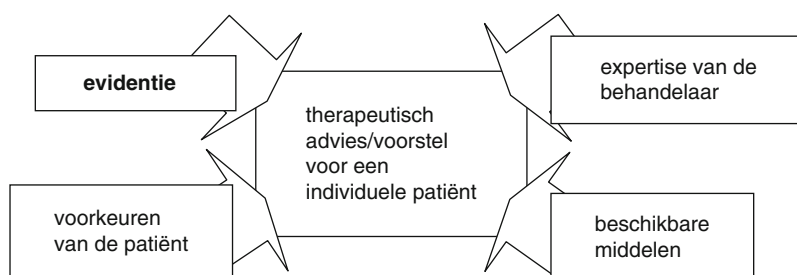
5	Beoordeling van kwalitatief onderzoek	98
5.1	Inleiding	98
5.2	De essentie van het kwalitatief onderzoeksparadigma	99
5.2.1	De sociale realiteit	99
5.2.2	Consequenties voor de kwalitatieve methodologie	99
5.2.3	Methoden van kwalitatief onderzoek	101
5.3	Beoordeling van kwalitatief onderzoek	103
5.3.1	Zijn de relevantie en het doel van het onderzoek duidelijk omschreven?	104
5.3.2	Wordt het theoretisch referentiekader adequaat beschreven?	105
5.3.3	Zijn de onderzoeksmethoden en -technieken geschikt voor het onderzoeksdoel?	105
5.3.4	Is de selectie van deelnemers adequaat voor het onderzoeksdoel?	108
5.3.5	Zijn de data op een adequate manier verzameld?	110
5.3.6	Zijn de data grondig geanalyseerd?	112
5.3.7	Worden de uitkomsten en conclusies van het onderzoek helder beschreven?	113
5.3.8	Eindoordeel. Is de kwaliteit voldoende én is de studie bruikbaar?	113
6	Het vinden van geschikte artikelen	115
6.1	Inleiding	115
6.2	Databases	115
6.2.1	Gratis databases	116
6.2.2	Databases via abonnement of bibliotheek	118
6.3	Verkrijgen van het gewenste artikel of rapport	120
6.4	Zoekstrategieën	121
6.5	Logopedie in wetenschappelijke tijdschriften	124
	Uitwerking van de oefeningen	126
	Bij hoofdstuk 3 Diagnostiek	126
	Bij hoofdstuk 4 Therapie	127

Literatuur	129
Verklarende woordenlijst	132
Over de auteurs	141
Register	142

1.1 Wat is evidence-based handelen?

In de jaren 1980 ontstond *evidence-based medicine* als toepassing van een nieuwe wetenschappelijke discipline, genaamd ‘klinische epidemiologie en biostatistiek’, in een nieuwe vorm van medisch onderwijs aan de McMaster University in Toronto (Sackett e.a., 2000). David Sackett en zijn collega’s ontwikkelden het toepassen van resultaten uit klinisch epidemiologisch onderzoek – in plaats van enkel informatie in handboeken of de ervaring van de docent – op het nemen van dagelijkse medische beslissingen.

Evidence-based medicine (EBM) is gedefinieerd als ‘het gewetensvol, expliciet en oordeelkundig gebruik van het huidige beste bewijsmateriaal om beslissingen te nemen voor individuele patiënten’ (Offringa e.a., 2008; Sackett e.a., 2000). Een andere beschrijving is: wetenschappelijk bewijs integreren met de expertise van de behandelaar en de behoefte van de patiënt met de middelen die de behandelaar ter beschikking staan; zie figuur 1.1.



Figuur 1.1 Evidentie is een van de factoren die van invloed zijn op de voorstellen van een behandelaar aan een individuele patiënt.

Evidence-based practice is niet het genereren en uitvoeren van wetenschappelijk onderzoek, enkel het gebruiken ervan. Het wordt daarom ook wel beschreven als het overbruggen van de kloof ('bridging the gap') tussen aan de ene kant de wetenschappers die geacht worden bruikbare evidentie te produceren en te publiceren en aan de andere kant de behandelaars die geacht worden beschikbare evidentie te vinden en te integreren in hun professionele handelen.

Wat betreft de behoefte aan evidentie gaat het er niet alleen om dat we willen weten wat werkt, maar vooral ook wat niet werkt. Er zijn ongetwijfeld logopedische interventies die niet werkzaam zijn, terwijl we denken dat ze dat wel zijn. Hoewel logopedische handelingen in het algemeen niet schadelijk zijn, kunnen logopedisten wel degelijk fouten maken door een verkeerde of onnodige behandeling te geven. En evidentie dat iets niet werkt is een enorme stimulans om te gaan bewijzen dat iets anders wél werkt. Tegelijkertijd kan er evidentie zijn gepubliceerd die de praktiserend logopedist niet bereikt. Denk bijvoorbeeld aan de wetenschappelijke publicaties van logopediewetenschappers, taalwetenschappers enzovoort, die meestal uitsluitend in wetenschappelijke tijdschriften en proefschriften terug te vinden zijn.

In dit boek gebruiken we de begrippen 'evidence-based practice', EBP in de logopedie of evidence-based logopedie (EBL), dat wil zeggen logopedisten die hun professionele handelen mede baseren op wetenschappelijke evidentie.

Evidence-based handelen is een manier van werken die bestaat uit vijf stappen (Sackett e.a., 2000, Offringa e.a., 2008).

- De behoefte aan informatie omzetten in een beantwoordbare vraag.
- Zoeken naar het beste bewijs dat de vraag kan beantwoorden.
- Kritische beoordeling van het bewijs: is het valide en toepasbaar?
- Met het bewijs een beslissing nemen die past in de omstandigheden en behoeften van de patiënt.
- Evalueren: wat is het resultaat en kan deze werkwijze de volgende keer beter en efficiënter?

Voor het toepassen van deze manier van werken maakt het niet uit of het een medisch of paramedisch c.q. logopedisch probleem betreft. Deze vijf stappen worden in dit boek uitgelegd aan de hand van logopedische voorbeelden.

1.2 Evidentie in de logopedie

Evidence-based handelen is wat betreft methodiek voor elke behandelaar hetzelfde, dus voor logopedisten hetzelfde als voor artsen.

Maar een van de verschillen is toch de hoeveelheid beschikbare evidentie, die in de geneeskunde aanzienlijk groter is dan in de logopedie. Daar zijn een aantal redenen voor te bedenken.

De kwaliteit van de therapeut is medebepalend voor de effectiviteit van een therapie, volgens sommigen zelfs allesbepalend. Dat is in veel gevallen waarschijnlijk waar: een ervaren logopedist bereikt betere of snellere resultaten dan een stagiaire of pas afgestudeerde logopedist. Veel ervaring lijkt vooral van belang bij intensieve gedragsveranderende therapieën, zoals stemtherapie of stottertherapie. Het vraagt veel van de vaardigheid van de therapeut om een patiënt tot ander gedrag te krijgen, te motiveren tot oefenen en tot het uitvoeren van ongewone opdrachten. Ofwel 'een goede therapeut met een slechte methode is beter dan een slechte therapeut met een goede methode'. Maar als het resultaat van de behandeling volledig afhankelijk is van één behandelaar, is het weinig zinvol om het effect van een behandel~~methode~~ te onderzoeken. Daarom is er in de geneeskunde veel meer evidentie beschikbaar: een medicijn wel of niet geven of een routineoperatie wel of niet uitvoeren zijn duidelijk meetbare interventies. En overlijden versus overleven is een betrouwbaar vast te stellen uitkomstmaat. Zodra de invloed van de behandelaar zelf een rol gaat spelen, wordt het moeilijker. Daarom worden goede interventiestudies, juist ook medicijnstudies, 'dubbelblind' uitgevoerd; dat wil zeggen dat ook de behandelaar niet weet of de patiënt het echte medicijn of de placebo krijgt toegediend. Dus zelfs de kennis van de behandelaar over de echtheid van het medicijn – de arts doet geen oefeningen met de patiënt – wordt beschouwd als een invloed op het resultaat.

Een andere verklaring voor de nog beperkte hoeveelheid wetenschappelijke logopedische evidentie is het feit dat logopedie of het onthouden daarvan niet levensbedreigend is. Risicovolle interventies dwingen onderzoek af naar de meerwaarde ten opzichte van het risico. Het is niet toevallig dat juist op het gebied van orofaryngeale slikstoornissen het wetenschappelijk onderzoek het laatste decennium zo enorm is toegenomen. Slechte diagnostiek en inadequate behandeling van slikstoornissen kunnen namelijk leiden tot gezondheidsrisico's, zoals ondervoeding en een grotere kans op over-

lijden, en in andere gevallen juist tot onnodige restricties in voeding.

Aan het einde van de vorige eeuw verschenen publicaties die overzichten gaven van logopedische effectstudies. Enderby en Emerson (1995) schreven het boek *Does speech and language therapy work?* en in het *Journal of Speech Language and Hearing Research* verscheen in 1996 en in 1998 een artikelenreeks van reviews over therapie-effecten in verschillende logopedische stoornisgebieden. Deze beschrijvende overzichten, zogenoemde *narrative reviews*, voldoen niet aan de eisen die tegenwoordig aan *systematic reviews* worden gesteld. Inmiddels zijn er verschillende *systematic reviews* over logopedische interventies verschenen in de Cochrane-database (zie hoofdstuk 2 en 6). In 2003 verscheen *Evidence-based practice in speech pathology* van Reilly, Douglas & Oates, die de evidentie beschreven van alle belangrijke logopedische domeinen. In 2005 publiceerde het Royal College of Speech Language Therapy (RCSLT) de *evidence based Clinical Guidelines* (zie www.rcslt.org/members/publications/RCSLT_Clinical_Guidelines.pdf). En de eerste Nederlandse monodisciplinaire *evidence-based* logopedische richtlijn verscheen in 2008 over de ziekte van Parkinson (Kalf e.a., 2008).

Bij de eerste druk van dit boek in 2004 konden we nog weinig zeggen over het gebruik van *evidence-based* logopedie in Nederland. Destijds waren ook de logopedieopleidingen net gestart met het invoeren van EBL in het curriculum. Dat heeft inmiddels plaatsgevonden en zelfs tot publicaties geleid (Spek, 2010). Of praktiserende logopedisten in Nederland en België tegenwoordig inderdaad veel meer naar literatuur zoeken en zelfstandig beoordelen en interpreteren weten we niet, maar het begrip ‘*evidence-based*’ lijkt te zijn ingeburgerd in ‘*evidence-based* cursussen’, ‘*evidence-based* handboeken’ enzovoort. Kennisoverdracht die gebaseerd is op evidentie heeft tegenwoordig een erkende meerwaarde. Daarbij horen ook overgeneralisaties zoals ‘*evidence-based* literatuur’ en ‘*evidence-based* onderzoek’.

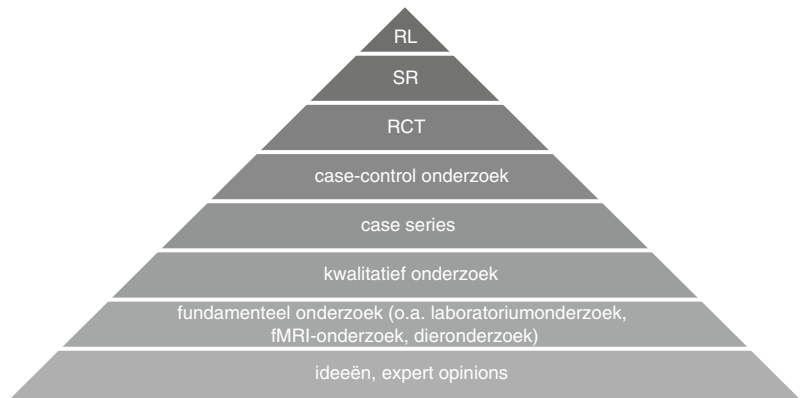
EBP is evenmin weg te denken uit de internationale logopedie. Enkele belangrijke demonstraties daarvan waren de keynote lecture van Barbara Dodd op het congres van de International Association of Logopedics and Phoniatrics (IALP) in 2007 (Dodd, 2007) en het verschijnen van het boek *Embedding evidence-based practice in speech-language therapy* van Hazel Roddam en Jemma Skead (Roddam & Skead, 2010). Zij beschrijven internationale voorbeelden rondom

vier thema's: het ontwikkelen van kennis en vaardigheden voor evidence-based handelen, het creëren van een stimulerende context voor evidence-based handelen, het toepassen van evidentie om klinische uitdagingen aan te gaan en de toekomst van evidence-based handelen in de logopedie.

1.3 Niveaus van evidentie

Een bekend fenomeen op congressen en studiedagen is dat veel presentaties fundamenteel onderzoek of kwalitatief onderzoek betreffen, zoals nieuw licht op stemplooibewegingen door gebruik van een highspeedcamera, beter begrip van hersenfuncties door gebruik van functionele MRI (fMRI), subtiele aanpassing van taalmodellen dankzij linguïstische analyse enzovoort. Logopedisten waarderen dergelijke resultaten, maar zijn tegelijkertijd geneigd die te beoordelen op therapeutische bruikbaarheid: wat moet ik ermee in de praktijk? Wat heb ik aan die informatie voor de behandeling van mijn patiënten? Zijn logopedisten daarmee wars van wetenschappelijke inzichten? Nee, voor het antwoord op deze vragen zijn kwantitatieve studies nodig die behandelingen op een gecontroleerde manier vergelijken, dus gerandomiseerde gecontroleerde trials (RCT's). Zie de evidentiepiramide in figuur 1.2 voor de opbouw van klinisch relevant bewijs. Dit zegt niets over de kwaliteit van het onderzoek, want een goed uitgevoerde fMRI-studie kan een betere publicatie opleveren dan een onbetrouwbaar uitgevoerde kleine RCT. Het laat vooral zien dat aan onderzoek naar de meerwaarde van een therapie of diagnostische meting jarenlang ander onderzoek voorafgaat. Het begint meestal met de ideeën en hypothesen van deskundigen en fundamenteel onderzoek om de hypothesen te ondersteunen of te weerleggen. Kwalitatief onderzoek is onder andere van belang om beter inzicht te krijgen in gedragsveranderingen en processen. *Case series* zijn beschrijvingen van kleine aantallen patiënten; en als patiënten (cases) op een valide manier vergeleken kunnen worden met gezonde proefpersonen (controls), ontstaat een nog beter inzicht, bijvoorbeeld in de factoren die een aandoening kunnen veroorzaken. Voor het valide vergelijken van goed omschreven behandelingen zijn RCT's nodig. Maar er zijn meerdere RCT's nodig om tot een echte uitspraak te komen of een behandeling werkt of niet en dat wordt gedaan door middel van systemati-

sche reviews. In richtlijnen ten slotte is evidentie (bij voorkeur van het hoogste niveau) bij elkaar gebracht om antwoord te geven op diverse klinische vragen.



Figuur 1.2 Evidentiepiramide. Hoe hoger in de piramide, hoe groter de invloed van de evidentie op het therapeutisch handelen. RCT = randomized clinical/controlled trial; SR = systematische review; RL = richtlijn.

Dichotome uitkomsten versus continue uitkomsten

Dichotome gegevens zijn twee elkaar uitsluitende waarden: ja/nee, wel/niet, vrouw/man enzovoort. Het hebben van dichotome uitkomsten betekent dat van alle patiënten na afloop van het experiment bekend is of zij wel of niet de stoornis hebben, wel of geen verbetering hebben doorgemaakt enzovoort.

Continue uitkomsten zijn veranderingen die uitgedrukt zijn binnen een serie van waarden, bijvoorbeeld een bepaalde testuitslag met daarbij een percentielscore. Met een afkappunt zijn continue waarden overigens weer te dichotomiseren in afwijkend/niet afwijkend of verbeterd/niet verbeterd.

Klinische relevantie versus statistische significantie

Statistische significantie (de p-waarde is kleiner dan 0,05) betekent dat de kans dat een gevonden verschil of effect op toeval berust kleiner is dan 5%. Dat is een belangrijke bevinding, maar niet hetzelfde als een klinisch relevant resultaat. Het gaat in evidence-based handelen met nadruk om de meerwaarde van een interventie voor de

patiënt; met andere woorden: het moet relevant zijn, echt iets uitmaken in relatie tot de geleverde fysieke, mentale en financiële inspanning. Een effect kan statistisch significant zijn (vooral als grote groepen patiënten zijn onderzocht), maar toch klein van omvang en klinisch niet de moeite van het toepassen waard. Klinische relevantie wordt in valideringsonderzoek vooral uitgedrukt in de voorspellende waarden van een test (zie hoofdstuk 3) en in interventieonderzoek, vooral in een *number needed to treat* (NNT; zie hoofdstuk 4). Dat kan alleen met dichotome uitkomsten, zodat berekening met 2x2-tabellen mogelijk is. Vandaar dat bij het beoordelen van artikelen de klassieke statistiek buiten beschouwing blijft en een eenvoudige rekenmachine of excelspreadsheet voldoende is.

Maar uit het oogpunt van zorgvuldigheid moeten we eraan toevoegen dat ook eenvoudig te berekenen uitkomsten (zoals het percentage patiënten dat is verbeterd) vergezeld zouden moeten gaan van een betrouwbaarheidsinterval (BI). Van elk experiment kun je je immers afvragen in hoeverre de uitkomsten hetzelfde zijn als je het precies zo nog een keer zou uitvoeren. De conventie is een BI van 95%, dat wil zeggen de laagste tot de hoogste uitkomst waar 95 van de waarden tussen zouden liggen als hetzelfde experiment honderdmaal herhaald zou worden. Hoe kleiner het aantal patiënten – die altijd van elkaar verschillen – hoe verder de uitkomsten bij herhaling uit elkaar liggen (een breder betrouwbaarheidsinterval) ofwel een minder betrouwbare weergave van de werkelijkheid. Dat betekent overigens niet dat het aantal patiënten in een studie altijd zo groot mogelijk moet zijn, maar dat is een methodologische discussie die buiten het bestek van dit boek valt. Zie box 1.1 voor een voorbeeld. De betrouwbaarheidsintervallen worden automatisch berekend in de rekenhulp die op de website (<http://extras.springer.com>) te vinden is.

Box 1.1

De Lee Silverman Voice Treatment (LSVT) is internationaal de meest onderzochte logopedische behandeltechniek voor parkinsonpatiënten. In een van de deelstudies (Baumgarter et al., 2001) is onder meer gekeken bij hoeveel patiënten een afname van de heesheid met ten minste 60% was bereikt. De controlebehandeling was ademtherapie. Het ging om twintig patiënten en acht van de dertien (62%) die met LSVT waren be-

handeld hadden een betere stemkwaliteit tegenover één van de zeven (14%) die de ademtherapie kregen. Het absolute verschil is 47%, dus ongeveer de helft van de behandelde parkinson-patiënten is minder hees na behandeling met LSVT (zie verder hoofdstuk 4). Het 95% BI daarvan is 28 tot 66%. Bij dezelfde uitkomsten met tweehonderd patiënten zou het 95% BI 41 tot 53% geweest zijn.

Tegenwoordig zijn ook de wetenschapskaternen van kranten zoals de *Volkskrant* en *NRC Handelsblad* geregeld gevuld met kritische beschouwingen van het laatste biomedische wetenschappelijke nieuws. Vooral de interpretatie van getallen kan lezers makkelijk op het verkeerde been zetten. Bijvoorbeeld het op veel websites te vinden bericht ‘Veel koffiedrinken beschermt tegen het krijgen van de ziekte van Parkinson’. De koffiedrinker denkt: ‘Mooi, dat scheelt weer.’ De theedrinker denkt: ‘Hoeveel koffie, hoeveel kleiner is die kans dan en is dat relevant genoeg om voortaan koffie te gaan drinken?’ Voor kritische lezers en consumenten heeft wetenschapsjournalist Hans van Maanen (2009) het boek *Goochelen met getallen – cijfers en statistiek in krant en wetenschap* geschreven. Ook zeer de moeite waard voor logopedisten en studenten die hun rekenvaardigheden in een breder gebied willen leren toepassen.

1.4 De PICO-vraag

Een vraag naar specifieke evidentie (stap 1) wordt in de systematiek van evidence-based handelen geformuleerd als een PICO-vraag (Offringa e.a., 2008). Ook dat is kenmerkend voor het zoeken naar klinische relevantie uit kwantitatief onderzoek. PICO staat voor *patient-intervention-comparison-outcome* en dat houdt het volgende in.

Patient	=	de patiëntengroep voor wie evidentie wordt gezocht.
Interventie	=	de interventie waarvan de effectiviteit (therapie) of diagnostische waarde (diagnose) wordt onderzocht, meestal de indexbehandeling of indextest genoemd.
Comparison	=	de interventie of test waarmee wordt vergeleken: de referentie- of controlebehandeling, of de referentietest.
Outcome	=	de gewenste uitkomst, bij voorkeur als dichotome maat (zie boven).

Het formuleren van PICO-vragen en het zoeken naar literatuur op basis van de PICO-onderdelen wordt verder besproken in hoofdstuk 6. Het toetsen van artikelen op geschiktheid met behulp van de PICO-onderdelen komt aan de orde in de diverse hoofdstukken.

1.5 Opbouw en gebruik van dit boek

Zoals aan het begin van dit hoofdstuk is beschreven, bestaat evidence-based handelen uit vijf stappen: vraag stellen, informatie zoeken, informatie kritisch beoordelen, informatie toepassen en het resultaat en het proces zelf evalueren. In dit boek beschrijven we de verschillende onderdelen in de omgekeerde volgorde. Vóór het lezen van de beschrijving van het formuleren van vragen en het zoeken naar informatie in databases, is namelijk al enige kennis van epidemiologische begrippen nodig. Daarom vonden we het nuttiger om eerst het kritisch beoordelen van artikelen te behandelen en zo de lezer vertrouwd te maken met begrippen uit de klinische epidemiologie (hoofdstukken 2 t/m 5). Daarna volgt het zoeken naar en selecteren van geschikte artikelen in wetenschappelijke databases, op basis van een goed overdachte vraag (hoofdstuk 6). Op de website zijn de beoordelingsformulieren te vinden die in de hoofdstukken worden besproken. Voor het maken van de berekeningen die in hoofdstuk 3 en 4 worden beschreven, is een eenvoudige rekenmachine voldoende. Op de website staan bovendien de 2x2-tabellen geautomatiseerd in een rekenprogramma.