

# De opbrengst van een Medline-search

## De toegankelijkheid van onderzoek naar de effectiviteit van acupunctuur II

E.A. HOFMANS

In deze secundaire analyse is allereerst nagegaan welk deel van een bestand van 67 artikelen over de effectiviteit van acupunctuur kon worden opgespoord via een aantal Medline-searches. Vervolgens is een aanvullend onderzoek uitgevoerd via de referenties bij de aldus gevonden artikelen; daarbij werd ook gezocht naar 16 artikelen die waren verschenen in niet-geïndexeerde tijdschriften. Via Medline werden 46 artikelen gevonden, waaronder 18 (van in totaal 24) 'essentiële studies'. Het aanvullend onderzoek leverde nog eens 10 in geïndexeerde tijdschriften verschenen artikelen op – waaronder 2 'essentiële studies' – en 9 artikelen in niet-geïndexeerde tijdschriften. De doeltreffendheid van de Medline-searches bedroeg 75 procent; het referentie-onderzoek voegde daar nog eens 8 procent aan toe. De doelmatigheid van de Medline-searches bedroeg – afhankelijk van de gebruikte 'gouden standaard' – 7 of 18 procent, maar de betekenis van deze twee cijfers is gering.

Hofmans EA. De opbrengst van een Medline-search. De toegankelijkheid van onderzoek naar de effectiviteit van acupunctuur II. *Huisarts Wet* 1990; 33(3): 103-6.

Drs. E.A. Hofmans, eindredacteur van *Huisarts en Wetenschap*, Postbus 49, 8200 AA Lelystad.

### Inleiding

In mijn vorige artikel kwam ik tot de conclusie dat er van publikatiebias met betrekking tot onderzoek naar de effectiviteit van acupunctuur geen sprake was: van de 89 artikelen die door *Ter Riet e.a.* waren opgespoord, bleken er 66 (74 procent) verschenen in door Index Medicus geïndexeerde tijdschriften. Belangrijker was de uitkomst dat 22 van de 24 'essentiële' studies waren gepubliceerd in geïndexeerde tijdschriften.<sup>1</sup>

De volgende vraag is natuurlijk, in hoeverre die 66 artikelen c.q. 22 essentiële studies ook werkelijk boven water komen bij een Medline-search. Uit eerder onderzoek is immers gebleken dat de 'doeltreffendheid' van een dergelijke search nogal te wensen overlaat. Afgaande op de resultaten van *Van Duijn en Dyserinck* mag verwacht worden dat langs deze weg niet meer dan 50 artikelen zullen worden opgespoord.<sup>2</sup>

Nu is een literatuuronderzoeker niet uitsluitend aangewezen op een elektronische database; de gebruikelijke volgende stap bestaat uit het doornemen van de referenties bij de in eerste instantie gevonden onderzoeksartikelen en bij eventuele literatuurstudies (voor zover die althans via Medline zijn gevonden). Daarbij gaat het niet alleen om artikelen die in Medline zouden moeten zitten, maar daar onvindbaar zijn, maar ook om artikelen in niet-geïndexeerde tijdschriften.

De vraagstelling van dit tweede onderzoek luidt:

- 1 Welk deel van de in geïndexeerde tijdschriften verschenen artikelen kan worden opgespoord via een Medline-search?
- 2 Welk deel van de ontbrekende artikelen kan worden opgespoord via de referenties bij de aldus opgespoorde artikelen (inclusief literatuurstudies)?
- 3 Wat is de opbrengst van beide strategieën in termen van doeltreffendheid en doelmatigheid?

### Methoden

1 In de Centrale Medische Bibliotheek van het Academisch Medisch Centrum van de Universiteit van Amsterdam heeft op 21

december 1989 een search plaatsgevonden met behulp van de door *Ter Riet e.a.* vermelde trefwoorden *acupuncture, clinical trials, migraine, vascular headache, headache, tobacco, smoking, asthma, rheumatoid arthritis, backache, osteoarthritis, cervical vertebrae, pain and therapy*. Op advies van de literatuuronderzoeker is bovendien een aanvullende search gedaan met behulp van het trefwoord *comparative studies* als alternatief voor *clinical trials*. Daarna is vastgesteld welke van de door *Ter Riet e.a.* beoordeelde onderzoeken via deze twee searches waren gevonden.

Vervolgens is – conform de suggestie van *Van Duijn en Dyserinck*<sup>2</sup> – op 2 januari 1990 nagegaan onder welke trefwoorden de op 21 december gevonden artikelen waren ondergebracht. Met een deel van deze trefwoorden zijn aanvullende searches uitgevoerd. Daarna is vastgesteld welke onderzoeken hierdoor alsnog konden worden opgespoord.

2 Alle opgespoorde artikelen zijn daarna onderzocht op verwijzingen naar onderzoeken die wel door *Ter Riet e.a.* waren beoordeeld, maar niet konden worden gevonden via Medline. Daarbij zijn ook de door *Ter Riet e.a.* expliciet besproken literatuurstudies betrokken.

3 *Van Duijn en Dyserinck* hebben in hun onderzoek naar de opbrengst van een computersearch naar de diagnostiek van sinusitis onderscheid gemaakt tussen de doeltreffendheid en de doelmatigheid van de Medline-search; zij zijn daarbij uitgegaan van een gouden standaard van 16 'relevante artikelen'. De *doeltreffendheid* van een search werd bepaald als het quotiënt van het aantal gevonden 'gouden standaard'-artikelen en het totaal aantal 'gouden standaard'-artikelen; de *doelmatigheid* werd bepaald als het quotiënt van het aantal gevonden 'gouden standaard'-artikelen en het totaal aantal gevonden artikelen. De uitkomsten van beide berekeningen lopen van 0 tot 1; na vermenigvuldiging met 100 gaat het dus om percentages.<sup>2</sup>

In dit onderzoek kunnen we beschikken over twee gouden standaarden:

– het totaal aantal door *Ter Riet e.a.* gevonden onderzoeksverslagen; hier gaat het dus niet alleen om de tijdschriftartikelen, maar ook om alle overige rappor-

tages, met uitzondering van de *abstracts*.  
 – het beperkte aantal ‘essentiële studies’ (*kader*).

## Resultaten

De uitkomsten van dit tweede onderzoek maken om te beginnen een kleine rectificatie noodzakelijk. *Ter Riet e.a.* hebben in totaal 98 (en niet 99) rapportages bijeengebracht. Deze rapportages moeten als volgt worden gerubriceerd:

- 1 ongepubliceerde studie;
- 1 ‘grijze’ studie (een rapport van de Rotterdamse GGD);
- 8 abstracts, alle gepubliceerd in een geïndexeerd tijdschrift;
- 1 ingezonden stuk, gepubliceerd in een geïndexeerd tijdschrift;
- 4 publikaties in een seriewerk, dat vóór 1983 – waarschijnlijk gedeeltelijk – werd geïndexeerd, en daarna niet meer;
- 83 artikelen, waarvan er 67 zijn verschenen in een door Index Medicus geïndexeerd tijdschrift.

Dit betekent dat het percentage artikelen dat is gepubliceerd in een door Index Medicus geïndexeerd tijdschrift, 81 procent bedraagt in plaats van 74 procent. Met betrekking tot de essentiële studies is er geen verschil: 22 van de 24 studies zijn gepubliceerd in een door Index Medicus geïndexeerd tijdschrift.

**1** De search met de door *Ter Riet e.a.* vermelde trefwoorden leverde 30 artikelen op; daaronder bevonden zich 14 essentiële studies. Bovendien werd een publikatie in een seriewerk gevonden. De aanvullende search met het trefwoord *comparative studies* leverde nog eens 12 artikelen op, waaronder 4 essentiële studies. Tevens werd een door *Ter Riet e.a.* beoordeeld ingezonden stuk gevonden.

De search op 2 januari leverde vier mogelijk relevante trefwoorden op: *Evaluation studies*, *Double blind method*, *Follow-up studies* en *Prospective studies*. Via deze trefwoorden konden nog eens 4 artikelen worden opgespoord, waaronder echter geen essentiële studies.

In totaal werden langs deze weg dus 46 artikelen en twee andere verslagen gevonden, waaronder 18 essentiële studies.

**Tabel 1** De doeltreffendheid van de opeenvolgende zoekstrategieën. Afgeronde percentages.\*

Zoekstrategie	Alle verslagen (n=90)	Essentiële studies (n=24)
Basistrefwoorden**		
+ a) clinical trials	34	58
+ b) comparative studies – a	14	17
+ c) evaluation studies, double blind method, follow-up studies, prospective studies – (a+b)	4	–
Referentie-onderzoek	21	08
Totaal	74	83

\* Teller: aantal gevonden ‘gouden-standaard’-verslagen/studies; noemer totaal aantal ‘gouden-standaard’-verslagen/studies. Het quotiënt is met 100 vermenigvuldigd.

\*\* Acupuncture, migraine, vascular headache, headache, tobacco, smoking, asthma, rheumatoid arthritis, backache, osteoarthritis, cervical vertebrae, pain and therapy.

**Tabel 2** De doelmatigheid van de opeenvolgende Medline-searches. Afgeronde percentages.\*

Zoekstrategie	Alle verslagen	Essentiële studies
Basistrefwoorden**		
+ a) clinical trials (n=98)	31	14
+ b) comparative studies – a (n=95)	13	4
+ c) evaluation studies, double blind method, follow-up studies, prospective studies – (a+b) (n=71)	6	–
Totaal (n=264)	18	7

\* Teller: aantal gevonden ‘gouden-standaard’-verslagen/studies; noemer totaal aantal gevonden verslagen/studies. Het quotiënt is met 100 vermenigvuldigd.

\*\* Acupuncture, migraine, vascular headache, headache, tobacco, smoking, asthma, rheumatoid arthritis, backache, osteoarthritis, cervical vertebrae, pain and therapy.

## Terminologie

De terminologie die dit in dit artikel wordt gebruikt, kan aanleiding geven tot verwarring. De volgende definities zijn gehanteerd:

- **Rapportage:** elke door *Ter Riet e.a.* beoordeelde beschrijving van een onderzoek naar de effectiviteit van acupunctuur; hierin zijn dus ook abstracts, ingezonden stukken en ongepubliceerde studies begrepen.
- **Onderzoeksverslag:** de rapportages minus de abstracts; de achterliggende overweging bij dit onderscheid is dat een

abstract per definitie niet beschouwd mag worden als een op zichzelf staand onderzoeksverslag.

- **Artikel:** een als zodanig gepresenteerde publikatie in een tijdschrift (en dus niet een hoofdstuk in een seriewerk of een ingezonden stuk).
- **Studie:** in een aantal gevallen is tweemaal gerapporteerd over hetzelfde onderzoek en zijn er dus twee rapportages. Hierbij gaat het echter om één studie. In de praktijk zijn alle ‘essentiële studies’ overigens (ook) gepubliceerd als tijdschriftartikel.

2 *Ter Riet e.a.* hebben in totaal 7 literatuuroverzichten besproken. Hiervan zijn er 5 verschenen in door Index Medicus geïndexeerde tijdschriften; al deze reviews werden gevonden bij de eerste search op 21 december. In totaal waren er dus 53 verslagen beschikbaar voor een nader onderzoek op relevante referenties. Dit aanvullend onderzoek leverde tenslotte nog eens 19 artikelen op, waaronder 2 essentiële studies.

3 De opbrengst van de hele exercitie bedraagt 65 artikelen, 1 ingezonden stuk en 1 bijdrage in een seriewerk, terwijl 18 artikelen, 3 bijdragen in een seriewerk, 1 grijze publicatie en 1 ongepubliceerde studie werden gemist. Van de 24 essentiële studies werden er 20 opgespoord. De *tabellen 1* en *2* geven een overzicht van de doeltreffendheid en doelmatigheid van de opeenvolgende stappen.

## Beschouwing

Essentieel voor het bepalen van de doeltreffendheid en doelmatigheid van een reeks zoekstrategieën is de beschikbaarheid van een gouden standaard. Zo'n gouden standaard kan echter alleen ontleend worden aan de uitkomst van een standaardliteratuuronderzoek. En als dát eenmaal is uitgevoerd, komen doeltreffendheid en doelmatigheid als mosterd na de maaltijd. Bij dit soort onderzoek gaat het dus per definitie om *secundaire analyses*; het is altijd wijsheid achteraf.

Toch hebben deze *posterior* exercities ook een praktische betekenis; zij kunnen van belang zijn bij het schatten van de *prior kansen* op een bepaalde opbrengst bij een later literatuuronderzoek. Op grond van een reeks van secundaire analyses van een bepaald welomschreven type literatuuronderzoek zouden bijvoorbeeld uitspraken kunnen worden gedaan over de te verwachten doeltreffendheid van een volgend literatuuronderzoek. Die uitspraken zullen waarschijnlijk voorlopig niet goed te kwantificeren zijn, maar het is een begin.

Wil men de doeltreffendheid van een aantal zoekstrategieën zinvol kunnen onderzoeken, dan is echter meer uniformiteit

nodig met betrekking tot de constructie van gouden standaarden.

*Van Duijn en Dyserinck* hebben voor hun gouden standaard gebruik gemaakt van een selectie van 'relevante artikelen' en de door hen aangehaalde auteurs doen dat ook. Alleen al op praktische gronden verdient deze keuze waarschijnlijk de voorkeur boven de mogelijkheid van een gouden standaard van 'alle artikelen', maar wél rijst dan het probleem van de toegepaste selectiecriteria. Die zijn bijvoorbeeld bij *Van Duijn en Dyserinck* in het geheel niet geëxpliciteerd. Verder hebben zij in hun gouden standaard bij voorbaat geen artikelen opgenomen uit niet-geïndexeerde tijdschriften (wat mij een nodeloze knieval voor de selectiebias van Medline lijkt) en beperken zij zich tot Nederlandse en Engelstalige artikelen (wat toch slecht kan leiden tot een gouden standaard die ook bruikbaar is voor internationaal vergelijkend onderzoek).

De door mij aan *Ter Riet e.a.* ontleende gouden standaard van 24 'essentiële studies' is daarentegen uitsluitend gebaseerd op kwaliteit, terwijl de gehanteerde normen gedetailleerd zijn beschreven.<sup>3</sup> Met de beperkingen van *Van Duijn en Dyserinck* zou mijn gouden standaard uit slechts 21 studies hebben bestaan, en zou de doeltreffendheid van de Medline-searches 81 procent hebben bedragen, en niet 75 procent. Dat is misschien niet veel, maar het verschil had gemakkelijk veel groter kunnen zijn, bijvoorbeeld bij een groter aantal Franse en Duitse 'essentiële studies'.

Overigens kan uit de verschillende secundaire analyses in ieder geval de conclusie worden getrokken dat de onderzoeker er goed aan doet zich niet te beperken tot een Medline-search; hij moet bijvoorbeeld ook de gevonden literatuur onderzoeken op bruikbare referenties. In mijn onderzoek nam de doeltreffendheid door deze laatste stap toe van 75 tot 83 procent; er ontbraken toen nog vier 'essentiële studies'. Wellicht hadden ook *Van Duijn en Dyserinck* langs deze weg het aantal opgespoorde 'gouden standaard'-artikelen nog wat kunnen opschroeven.

Is 83 procent voldoende? En hoe essentieel of relevant is een 'essentiële' of 'relevante' studie?

Die vragen zijn niet in algemene zin te beantwoorden. Bovendien geldt ook bij literatuuronderzoek de wet van de afnemende meeropbrengst: na een bepaald punt levert elke extra eenheid inspanning steeds minder op, terwijl omgekeerd elke extra eenheid opbrengst een steeds grotere inspanning vereist. Vroeg of laat zal de onderzoeker moeten stoppen met zijn nasporingen. En dat brengt op ons op de doelmatigheid van de hele operatie.

De *doelmatigheid* van de Medline-searches is door *Van Duijn en Dyserinck* uitgedrukt als het quotiënt van het aantal gevonden 'gouden standaard'-artikelen en het totaal aantal artikelen dat met een bepaalde search werd opgespoord. De onaangename consequentie van deze operationalisering is echter dat de doelmatigheid automatisch afneemt, naarmate men voor een kleiner aantal 'gouden standaard'-artikelen kiest; we zien dat geïllustreerd in *tabel 2*. Hierdoor wordt vergelijkend onderzoek naar de doelmatigheid van zoekstrategieën bij voorbaat vrijwel onmogelijk.

Afgezien daarvan is het gebruikte doelmatigheids-quotiënt geen realistische afspiegeling van de inspanning die men zich voor een bepaald resultaat moet getroosten: sommige titels die een Medline-search oplevert, kan men zonder meer terzijde leggen, andere vereisen bestudering van het *abstract*, en weer andere kunnen pas beoordeeld worden door het eigenlijke artikel te raadplegen. Maar al deze titels wegen even zwaar in de noemer van het quotiënt. En dan laat ik nog buiten beschouwing dat deze operationalisering al bij voorbaat ongeschikt is voor het bepalen van de doelmatigheid van aanvullend referentie-onderzoek.

Dat betekent echter niet dat doelmatigheid geen zinvol begrip zou zijn bij het evalueren van een bepaald literatuuronderzoek. Integendeel: de onderzoeker doet er goed aan periodiek de balans op te maken en daarbij ook de doelmatigheid van zijn nasporingen te betrekken. Hij zou zich na een aantal Medline-searches en aanvullend referentie-onderzoek bijvoorbeeld kunnen afvragen hoe groot de kans is dat verdere nasporingen nieuwe – en dan met name 'essentiële' – studies zullen opleveren, en

hoeveel inspanningen hij zich daarvoor zal moeten getroosten.

Het antwoord op deze vraag zal hij natuurlijk nooit zeker weten. Maar is het niet juist de huisarts die moet leren leven met schattingen en onzekerheden?

### Dankbetuiging

Met dank aan mw H.C. Dyserinck en mw drs. S. Bakker, respectievelijk literatuuronderzoeker en bibliothecaris bij de Centrale Medische Bibliotheek van het Academisch Medisch Centrum van de Universiteit van Amsterdam, voor hun hulp bij uitvoeren van de searches, en aan G. ter Riet voor het ter beschikking stellen van fotocopies van de opgespoorde artikelen.

### Literatuur

<sup>1</sup> Hofmans EA. Publikatiebias – Realiteit of mythe. De toegankelijkheid van onderzoek naar de effectiviteit van acupunctuur. *Huisarts Wet* 1990; 33(1): 14-5.

<sup>2</sup> Van Duijn NP, Dyserinck H. Literatuuronder-

zoek via een computersearch. Doeltreffendheid en doelmatigheid van een search naar sinusitis. *Huisarts Wet* 1990; 33(1): 9-13.

<sup>3</sup> Ter Riet G, Kleijnen J, Knipschild P. De meta-analyse als review-methode. *Huisarts Wet* 1989; 32: 176-81. ■

### Abstract

**Hofmans EA. Publication bias – reality or myth. The accessibility of research on the effectiveness of acupuncture. *Huisarts Wet* 1990; 33(1): 14-15.**

**Hofmans EA. The results of a Medline search. The accessibility of research on the effectiveness of acupuncture II. *Huisarts Wet* 1990; 33(3): 103-6.**

In 1989 a report was published in *Huisarts en Wetenschap* of a meta-analysis of randomized controlled research on the effectiveness of acupuncture. In collecting the required material the authors did not confine themselves to a Medline search but aimed for completeness; in this context they indicated the possibility of publication bias. This assumption was further investigated in a secondary analysis of the collected material.

The meta-analysis involved 98 research reports: 1 unpublished study, 1 'grey' study, 8 abstracts, 1 letter to the editor, 4 publications in a series and 83 journal articles. 67 of these 83 articles appeared in journals indexed in Index Medicus. No fewer than 22 of the 24 studies which scored the highest in the meta-analysis were published in indexed journals.

It was subsequently shown that 46 of the 67 indexed articles could actually be traced via Medline, including 18 of the highest scoring studies.

An additional search was carried out via the references of the articles found. This yielded a further 10 articles appearing in indexed journals, including 2 of the highest scoring studies, and 9 articles in non-indexed journals.

The conclusions of this study are therefore: there are no indications of publication bias in this area; a Medline search failed to reveal a number of studies; an additional search led to a very satisfactory final survey of the literature.

**Key words** Acupuncture, Medline, Meta-analysis.

**Correspondence** E.A. Hofmans, PO Box 49, 8200 Lelystad, The Netherlands.