

## Hoofdstuk 7 (blz. 208)

1. Sinds midden jaren negentig van de vorige eeuw kunnen zowel privépersonen als instellingen zoals universiteitsbibliotheken gebruik maken van internet. De belangrijkste applicatie daarbij is Entrez PubMed, waarvoor geen gebruikskosten worden aangerekend, een faciliteit die destijds bedongen werd door Al Gore (Nobelprijs voor de Vrede, 2007). Dit instrument heeft zich sindsdien ontwikkeld tot een van de krachtigste informatica antwoorden op de informatielawine.
2. De Medical Subject Headings (MeSH) is een aan PubMed gekoppeld hiërarchisch geordend woordenboek dat bestaat uit ongeveer 25000 kern termen (zogenaamde descriptors) en ongeveer 10 maal zoveel daaraan gekoppelde ingangstermen (zogenaamde supplementary concept terms en ook entry terms). Zoekacties op basis van eigen termen worden op voorhand gestructureerd en geordend door MeSH. De toegang tot dit woordenboek is zeer eenvoudig te verkrijgen via het centrale werkblad van PubMed.
3. Men moet keuzes kunnen maken op basis van kwaliteit in plaats van bij toeval. Kwaliteit staat daardoor centraal bij iedere zoekoperatie. Kwaliteit is maar ten dele te definiëren met getallen zoals de impactfactor. Relevantie is ook een kwaliteitskenmerk, maar het is nauwelijks te vatten in een algoritme voor de gebruiker. Hulp bij literatuuronderzoek door meer ervaren onderzoekers is daarom geen luxe.
4. De auteur van dat artikel kan dit pas doen twee jaar na verschijnen van zijn/haar artikel. Op dat moment wordt het aantal citaties van dat specifieke artikel bekend. Dit aantal kan vergeleken worden met het gemiddeld aantal citaties dat een artikel van dat tijdschrift genereerde in een observatieperiode van twee jaar, zijnde de (meest recente) impactfactor van dat tijdschrift.
5. Omdat er slechts een minderheid van tijdschriften met hoge impact bestaat en de verdeling van impacten volgens de rangorde van tijdschriften per domein bijgevolg niet lineair is (zie leerboek blz. 171, Fig. 26), is de kans erg groot dat de som der impacten van de twee lager gerangschikte tijdschriften toch nog kleiner is dan de impact van dat ene toptijdschrift. Het beste advies zou dus moeten zijn om toch de studie naar dat toptijdschrift te sturen.
6. Deze opdracht is in een aparte tekst in detail geformuleerd.
7. PMC of Pub Med Central. Als limiet is deze database af te raden, omdat de geselecteerde groep tijdschriften helaas nog te klein is. Het bijzondere is evenwel dat het hier vaak om toptijdschriften gaat, en niet om tweederangs tijdschriften.
8. Men kan de auteur benaderen met een gerichte vraag naar het specifieke artikel; in de mate van het mogelijke zal de auteur hier altijd positief op reageren, want iedere auteur wil gelezen worden.
9. De informatiedichtheid van een tabel is over het algemeen hoger dan die van een (eenvoudig) grafisch product. Moderne toptijdschriften genereren tegenwoordig grafisch materiaal dat zich volledig onttrekt aan deze algemene stelregel. Hier is de kwaliteit echter zeer hoog en dit is zelden haalbaar voor basisgebruikers van grafische applicaties.

10. Veel informatie per oppervlakte eenheid met zo weinig mogelijk data inkt.
11. Meet het volledig grafisch effect en vergelijk dit met het effect op de data. Heel vaak zijn de data te herleiden tot één dimensie. Het grafisch effect drukt die data veelal uit in twee of zelfs drie dimensies. Als een oppervlak in bijvoorbeeld een tijdsreeks vier keer groter wordt en het achterliggend getal slechts twee keer groter, dan is de leugenfactor vier, gedeeld door twee, dus twee (zie bijvoorbeeld blz. 186, fig. 33).
12. Onderzoek rapporteren vereist heel wat afwegingen. Discussie over wat wel en wat niet kan is de beste oplossing wanneer beslissingen om met data om te gaan lastig worden. Een eerste auteur houdt veel van zijn/haar data en dat is niet altijd het beste uitgangspunt (liefde maakt blind; dat geldt ook voor de onderzoeker). Iemand die in isolement werkt, heeft niet langer de kans om even wakker geschud te worden en nuchter naar de eigen data te kijken. Dit soort isolement is dus een risico voor een al te subjectieve manier van data verwerken. Ere-auteurs geven de valse indruk dat er wat dit betreft geen probleem is; ze zijn daarom een omstandigheid die wetenschappelijk falen in de hand werkt.
13. De eerste functie van ieder uitdrukingsmiddel van data, tabellen, figuren of grafieken is het zelf begrijpen wat de data betekenen. Uitdrukken en samenbrengen van gegevens gaat dus primair om verwerken van gegevens, niet ompubliceren.
14. De rol van kwantitatieve technieken is eveneens om er achter te komen wat de gegevens betekenen. Deze technieken zijn vooral een hulpmiddel om het soms verkeerde verwachtingspatroon te doorbreken; een statistische techniek toont bijvoorbeeld aan dat een verschil, waarvan men gedacht had dat het significant zou zijn, dit uiteindelijk niet is. De betekenis van mogelijke verschillen wordt beïnvloed door het verwachtingspatroon; er zijn objectieve (vaak kwantitatieve) methoden nodig om aan te tonen of die verwachtingen al of niet kloppen.
15. Het tijdschrift van tweede keuze heeft heel vaak een lagere impact. Dit betekent dat de statistische kans op een groot aantal citaties lager uitvalt, maar dat verhindert niet dat dit artikel nog veel geciteerd kan gaan worden, meer zelfs dan de impact van het tijdschrift zou voorspellen. Herinner u de uitspraken van eredoctoer Martin Raff tijdens zijn lezing hierover.
16. De nefaste rol van ere-auteurschappen is in het antwoord op vraag 12 aan de orde geweest. Men kan deze vormen van auteurschap herkennen aan de vlotte verwijzing naar de eerste auteur, bijvoorbeeld op congressen, in geval van vragen over het artikel. Als louter de ere-auteur aanwezig is in geval dat het artikel in het openbaar wordt aangevallen, dan is het ere-auteurschap extra pijnlijk en een vraagteken achter de reputatie van de onderzoeksgroep.