



EBM: de nieuwe kleren van de keizer?

Evidence based medicine (EBM): een term die steeds meer gebruikt wordt. Soms ook in zinnen waarvan ik me afvraag: hoe zou men dat een aantal jaar terug geformuleerd hebben, toen deze term nog geen gemeengoed was? Enkele voorbeelden:

- “Hoe evidence based is onze werkwijze en op welke manier zou je dat het beste kunnen vaststellen?”
- “Zouden we niet meer evidence based onderzoek moeten doen om onszelf echt geloofwaardig te maken en op de kaart te zetten?”
- Bij de aankondiging van een symposium: “Het programma is evidence based en gevarieerd.”
- “Hoe lang moet een bekend, werkend lijkend resultaat van een onderzoek bestaan om mee te mogen doen als bewijs voor evidence based medicine?”
- “Voor mij is het moeilijk om te beoordelen welke onderzoeken ergens op slaan (evidence based zijn) en welke niet.”
- “Dit is (niet) evidence-based en daarom doen we het (niet).”

Het doet mij denken aan het sprookje van de nieuwe kleren van de keizer. Een nieuw begrip dat te pas en te onpas wordt gebruikt en geroemd, echter zonder te zien waar het werkelijk om gaat. Nu wil ik zeker niet beweren dat EBM “niets om het lijf” heeft. Integendeel: het is een ingewikkelde materie, die veel tijd en energie kost en serieus aandacht behoeft.

Wat is er aan de hand?

Deels is het een kwestie van begripsverwarring en deels een gebrek aan kennis en vaardigheden.

De begripsverwarring betreft vooral het woord “evidence” dat veelal vrij vertaald wordt als “bewijs”. Maar voor ons woord “bewijs” gebruiken de Engelsen het woord “proof”. “Evidence” is al het beschikbare (onderzoeks)materiaal. “Evidence based” is dan ook niet hetzelfde als “bewezen effectief”.

Wat het gebrek aan kennis en vaardigheden betreft spelen er een aantal zaken. Niet iedereen beseft dat EBM duidt op het toepassen van een bepaalde methode, werkwijze. Daarnaast kent niet iedereen het epidemiologisch begrippenkader voldoende om op een juiste manier (medisch) wetenschappelijk onderzoek te beoordelen en te interpreteren. En ten slotte beheerst niet iedereen de kunst van het omgaan met cijfers, en dan doel ik vooral op de betekenis voor de praktijk, de relevantie in plaats van de statistische significantie.

In het kort komt EBM als methode erop neer dat er met een bepaalde zoekmethode (trefwoorden, databases, periode) literatuur verzameld wordt waarbij men transparant is in wat wel en niet in de beoordeling meege-

nomen wordt. De geselecteerde literatuur wordt meestal door twee onafhankelijke beoordelaars beoordeeld en gegradeerd. Resultaten worden samengevat in evidencetabellen en tekst, waarna conclusies worden geformuleerd (en soms ook gegradeerd). Bij het opstellen van een richtlijn worden vervolgens, mede op basis van overige overwegingen, aanbevelingen geformuleerd. In de kliniek neemt de behandelend arts in overleg met de patiënt een individuele beslissing.

Evidence based wil dus niets anders zeggen dan dat er op een systematische wijze is gezocht naar (onderzoeks)materiaal in databases en elders, dat dit transparant wordt gerapporteerd en dat aan de medisch wetenschappelijke informatie die uiteindelijk is geselecteerd een “level of evidence” wordt toegekend. Deze wordt vaak vertaald als “bewijskracht” maar overtuigingskracht dekt de lading beter. Sterkere evidence verdringt in principe zwakkere.

Natuurlijk is het gebruik van deze methode een hele verbetering ten opzichte van de periode daarvoor toen we nog niet zo helder en systematisch in beeld hadden wat mogelijke effecten van ons handelen zijn. Maar zoals dat vaak gaat: zodra men blindelings en zonder nadenken de systematiek volgt komen ook de nadelen in beeld.

EBM als methode is zeer arbeidsintensief: het beoordelen en samenvatten van een artikel in een evidencetabel kost al snel een uur.

Het systeem van graderen zorgt ervoor dat onderzoek vooral beoordeeld wordt op de methodologische kwaliteit, vaak met de RCT eenzaam aan top.^{1,2} Hierdoor ligt veel nadruk op de interne validiteit (is het onderzochte juist voor degenen die onderzocht zijn?) en veel minder op de externe validiteit (in hoeverre geldt het onderzochte ook voor anderen dan de onderzochten?). Weliswaar kan hier met een nieuwe manier van graderen (GRADE)³ via up- en downgraden meer nuance in aangebracht worden, maar het is en blijft mensenwerk.

Evidence-based wil weliswaar zeggen: er is (goed) naar gekeken, maar het is geen keurmerk voor de waarheid, laat staan voor directe toepasbaarheid in de praktijk.

De methode is vooral geschikt voor het beoordelen van interventies en diagnostiek. Maar ook op deze terreinen is niet alles goed genoeg of direct toepasbaar onderzocht. Onderzoeksomstandigheden kunnen nogal eens afwijken van de dagelijkse praktijk. Vaak betreft het een ander soort populatie dan die bij de behandelaar op bezoek komt, soms zijn de interventies niet goed vergelijkbaar en heel vaak betreft het andere uitkomstmaten dan de gewenste. Zoals Einstein al zei: “Not everything that can be counted counts, and not everything that counts can be counted.”

Maar er is vooral veel meer aandacht nodig voor de grootte van het effect en de relevantie voor de praktijk. Veel uitkomsten van onderzoek worden in termen van statistische significantie en relatieve maten gegeven. Dat de redactie van dit tijdschrift daar geen voorstander van is en streeft naar absolute uitkomstmaten, omdat deze beter te interpreteren zijn voor de praktijk, is duidelijk te lezen onder het kopje “statistiek” in de “Richtlijnen voor auteurs” die onlangs aangevuld zijn.⁴

Daarbij moeten we ook, wat ik maar even noem, “cijferwijs” worden. Het is goed ons te realiseren dat geen enkele interventie bij 100% van de doelgroep aanslaat, dat de alternatieve aanpak of het achterwege laten van enige interventie ook succes kan hebben, dat geen enkele risicofactor in 100% van de gevallen ziekte voorspelt en de afwezigheid daarvan evenmin duidt op een kans van 0%. Het gaat altijd om een beetje meer of een beetje minder. Hoeveel meer en hoeveel minder is iets waar we meer bij stil zouden moeten staan, alvorens we besluiten tot de ene of andere handelwijze. En het kan geen kwaad om daar vooraf, voor we een artikel lezen, eens goed bij stil te staan.

Anders is het risico groot dat we ons laten leiden door hetgeen een auteur ons vertelt. Zinsneden als “De meeste patiënten hebben baat bij ...”, “Interventie A heeft een gunstig effect op ...” of “De behandeling is effectief ...” moeten ons niet alleen nieuwsgierig maken naar wat, bij wie, onder welke omstandigheden is onderzocht, maar vooral naar de onderliggende getallen. Laten we, in plaats van kritiekloos de aanbevolen weg te volgen, toch vooral zelf blijven nadenken.

De remedie tegen de “nieuwe kleren van de keizer” luidt dan ook: maak ze zichtbaar!

Benoem:

- de soort kleding: wat is er precies onderzocht?
- de kleur: bij wie en onder welke omstandigheden is dit onderzocht?
- het materiaal: wat was de uitkomstmaat?
- de pasvorm: wat is de relevantie voor mijn praktijk?

En of ze mooi zijn? Dat mag u zelf bepalen! Want alleen over smaak valt te twisten.

Willy-Anne H.J. van Stiphout

LITERATUUR

1. www.cbo.nl/thema/Richtlijnen/Samenwerkingsverbanden/EBRO-platform/
2. www.cbo.nl/thema/Richtlijnen/EBRO-handleiding/PDF-versie/
3. www.gradeworkinggroup.org/
4. www.verenigingvenw.nl/TSG%20auteursinstructie%20wetenschappelijke%20artikelen.pdf (zie de site van V&W: www.verenigingvenw.nl en klik op het tijdschrift)

Nieuwe hoofdredacteur

Al kort na het vertrek van Niek Klazinga heeft TSG een nieuwe hoofdredacteur. Johan Polder komt de hoofdredactie versterken. Johan kent het werkveld van de public health als geen ander door zijn werk bij het RIVM ten behoeve van de Volksgezondheid Toekomst Verkenningen en zijn positie als bijzonder hoogleraar gezondheidseconomie bij TRANZO, Universiteit van Tilburg. Hij kent ook TSG al goed omdat hij al negen jaar deel uitmaakt van de redactie van ons tijdschrift.

TSG heeft in Johan Polder een prima opvolger gevonden van Niek Klazinga

Joost van der Gulden, namens de redactie van TSG