

AAN: Onze Lieve Vrouwe Gasthuis (OLVG)
Medische Bibliotheek
Postbus 95500
1090 HM Amsterdam

fax 020-5993837

PPN: 301751420,165740221
TSG : tijdschrift voor gezondheidswetenschappen, 2007. - Utrecht : V&W.
ISSN 1876-8776

GEWENST: 2006 vol. 84 afl. 3 K E
i450 ; ; 1998 V76 1 - 2010 V88 8
Artikel : FORUM de cijfers de baas
Bladzijden : 187-188

NCC/IBL PERIODIEK 24-02-2012
A096243945 / Mariska Zwartsenbur
di: 24-02-2012 / dp: 24-02-2012
Beh. voor: 20-04-2012 Rek.: JA
PPN: 301751420,165740221

NCC/IBL PERIODIEK 24-02-2012
A096243945 / Mariska Zwartsenbur
di: 24-02-2012 / dp: 24-02-2012
Beh. voor: 20-04-2012 Rek.: JA
PPN: 301751420,165740221

Rekening aan:
Onze Lieve Vrouwe Gasthuis (OLVG)
Medische Bibliotheek
Postbus 95500
1090 HM Amsterdam

Rekening aan:
Onze Lieve Vrouwe Gasthuis (OLVG)
Medische Bibliotheek
Postbus 95500
1090 HM Amsterdam

1200783

De gewenste publicatie (aankruisen wat van toepassing is):

(1)	☐	origineel gestuurd	(6)	☐	niet beschikbaar	-----
(2)	☐	kopie gestuurd	(7)	☐	uitgeleend	
(3)	☐	lokaal / overige	(8)	☐	wordt niet uitgeleend	
(4)	☐	nog niet aanwezig	(9)	☐	bibliografisch onjuist	
(5)	☐	niet aanwezig	(0)	☐	bij de binder	-----

3. SIGNIFICANTIE EN RELEVANTIE

Willy-Anne van Stiphout, Jan Burema

Wanneer beschouwen we een bepaalde variabele als een relevante determinant of risicofactor voor een bepaald gezondheidsprobleem? Deze vraag raakt aan de kern van het epidemiologisch onderzoek, dat immers, naast het tellen, primair erop gericht is om verbanden aan te tonen. De maat van verband in de epidemiologie is het relatieve risico (RR). Dit kan een ratio zijn van twee prevalentiecijfers, incidentiecijfers of cumulatieve incidenties of afgeleid worden, onder bepaalde voorwaarden (waarover een volgende keer meer) uit een odds ratio.

Waarom is dat relatieve risico zo geliefd? Wel, het geeft niet alleen aan dat er een verband is, maar ook hoe sterk dat is. Zoals bekend is epidemiologisch onderzoek vooral gericht op het vinden van mogelijke oorzaken van gezondheidsproblemen. Tegelijkertijd is het zo dat men door middel van observationeel onderzoek nooit kan bewijzen dat een bepaalde factor ook werkelijk een oorzaak is. Voor het aantonen van de 'pathofysiologische werking' van een determinant is over het algemeen laboratorium- of dierexperimenteel onderzoek noodzakelijk. Vandaar ook dat in de ogen van veel natuurwetenschappers de epidemiologie maar een 'softe business' is.

Echter, er zijn een aantal aspecten die 'oorzakelijkheid' in observationeel epidemiologisch onderzoek meer of minder aannemelijk maken. Zo is er de absolute voorwaarde dat de determinant in de tijd vooraf moet gaan aan de ziekte. Een dosis-respons relatie (hoe meer blootstelling aan de determinant, hoe groter de kans op ziekte) is een aspect dat oorzakelijkheid aannemelijk maakt. Dat geldt ook voor de bevinding dat het wegnemen van de determinant samengaat met een afname van het risico op ziekte. En, last but not least: hoe sterker het verband (lees: hoe groter het relatieve risico) hoe meer aannemelijk dat er sprake is van oorzakelijkheid. Die gerichtheid op het relatief risico is dus heel begrijpelijk, al weten we inmiddels dat de sterkte ervan ook afhankelijk is van de verdeling van noodzakelijke en voldoende oorzaken van een gezondheidsprobleem in een bepaalde populatie. En dat ook zwakke verbanden 'oorzakelijk' kunnen zijn. RR 's van 10-20, zoals ooit gevonden voor het verband tussen roken en longkanker (een zeer sterk verband en ook causaal gebleken) worden maar hoogst zelden gevonden. Tegenwoordig moeten we ons doorgaans tevreden stellen met tamelijk zwakke verbanden en RR 's tussen 0,5 en 2. Om ze wat meer te doen lijken worden ze soms als percentage risicotoename gepresenteerd. Zo wordt bijvoorbeeld een RR van 1,3 'vertaald' naar 30% meer en een RR van 0,8 naar 20% minder risico. Of het voor de praktijk een relevante toe- of afname van het gezondheidsprobleem is hangt, zoals we de vorige keer gezien hebben, ook af van het absolute risico.

Toch lijken in veel wetenschappelijke publicaties de auteurs wat minder nadruk te leggen op vermelding van de grootte van het relatieve risico dan op het feit of het 'statistisch significant' is. Statistische toetsen houden rekening met toevalsvariantie en beantwoorden de vraag of een gevonden verschil of relatief risico groter is, dan op basis van toeval verwacht kan worden. Dat bepaalt immers hoe 'overtuigend' een conclusie omtrent een gevonden verband is. Echter, of een verschil of relatief risico 'statistisch significant' werd, was niet alleen afhankelijk van de grootte van het verschil, maar ook van de aantallen waarnemingen (hoe groter het aantal personen, hoe nauwkeuriger de schatting, en dus des te overtuigender de conclusie verbonden aan eenzelfde geschatte effect). Eenzelfde te bestuderen associatie zal dus in een klein onderzoek niet, en in een groot onderzoek wel statistisch significant worden.

Louter afgaan op statistische significantie doet daarom de praktijk, en ook de wetenschap, niet altijd recht. Enerzijds omdat relatief zwakke verbanden (relatieve risico's veel kleiner dan 2) onevenredig veel aandacht kunnen krijgen als ze in een groot onderzoek gevonden worden, omdat ze daar statistisch significant zijn. Anderzijds worden er in kleine onderzoeken soms qua grootte 'relevante' verbanden gevonden die net niet statistisch significant zijn. Wanneer daardoor de schatting van dit effect niet expliciet in het artikel genoemd wordt of, erger nog, deze interessante factor niet meegenomen wordt in de slot-analyse van het onderzoek, dan spreekt men van een fout-negatieve bevinding.

Dat laatste zouden we een 'kunstfout' willen noemen. Waarom? Bij ieder verband dat onderzocht wordt kunnen andere determinanten als versturende variabele ('confounder') de schatting van het effect beïnvloeden. Dat betekent dat de uitkomst van een univariaat onderzocht verband niet altijd de juiste interpretatie toelaat: zonder correctie kan het verband sterker lijken dan het in werkelijkheid is, maar ook juist minder sterk. M.a.w. een wel aanwezig verband kan door confounding gemaskeerd worden en dus ten onrechte niet meegenomen in een multivariaat model. Dit kan, zowel voor de wetenschap (wat zijn de determinanten en hoe sterk is hun relatie met het gezondheidsprobleem) als voor de praktijk tot foute conclusies leiden.

Een soortgelijke 'interpretatie fout' kan ook optreden als na correctie ('adjustment') de grootte van het RR ongeveer gelijk blijft, maar het betrouwbaarheidsinterval wijder wordt. Wanneer het betrouwbaarheidsinterval aanvankelijk de waarde van de nulhypothese ($RR=1$) juist niet bevatte, maar daarna wel, dan lijkt het of een significant resultaat opeens niet meer

significant (en dus ook niet meer interessant) is. Vaak wordt dan gesteld dat er geen verband meer is. Maar doordat er meer variabelen in de analyse betrokken zijn is slechts de schattingsnauwkeurigheid toegenomen. En alleen daardoor is de significantie met een sprongetje gewijzigd, ook al is de sterkte van het verband en dus de mogelijke relevantie nauwelijks veranderd.

Hoe zien we het graag in het TSG? Zoals veel auteurs inmiddels al doen wordt rondom de verband- of associatiematen het 95%-betrouwbaarheidsinterval gegeven in plaats van alleen de p-waarde. De p-waarde geeft de kans aan dat het gevonden verschil louter op basis van toeval gevonden had kunnen worden. Als deze kans kleiner dan 5% is ($p < 0,05$) dan is het conventie te zeggen dat het verschil statistisch significant is omdat het te onwaarschijnlijk is dat zo'n uitkomst zou optreden als er geen echt verschil was. Het 95% -betrouwbaarheidsinterval geeft de lezer de mogelijkheid om te zien wat de range van plausible waarden is. Daarnaast zien we graag dat de auteur in tabellen ook de ongecorrigeerde waarden zet en verder duidelijk aangeeft voor welke determinanten gecorrigeerd wordt en waarom voor deze determinanten gekozen is. Een inhoudelijke keuze (het zijn relevante determinanten, want na correctie verandert er wezenlijk iets) heeft daarbij de voorkeur boven een statistische. En vergeet ook de absolute risico's niet, zodat de werker in de praktijk de maatschappelijke relevantie kan bepalen.

Als laatste nog even terug naar de praktijk. Hoe bepaal je nu of een determinant of risicofactor relevant is om er actie op te gaan ondernemen? Een eerste voorwaarde is natuurlijk dat de determinant te beïnvloeden moet zijn. Dat geldt bijvoorbeeld niet voor factoren als leeftijd en geslacht. Verder moet duidelijk zijn hoe de risicofactor 'werkt', hetgeen bij sociaal economische variabelen nogal eens moeilijk in te schatten is. Natuurlijk is het belangrijk om te kijken naar de grootte van het effect van de risicofactor: hoeveel meer risico, ook in absolute zin, levert deze op en vinden we dat verschil relevant? En niet te vergeten: hoe vaak komt de risicofactor voor en is dit frequent genoeg om een relevant effect op het gezondheidsprobleem te sorteren? Het maximaal te bereiken effect is te berekenen via de zogenaamde *PAR*: het populatie attributief risico, dat aangeeft hoe groot de proportie van een bepaald gezondheidsprobleem is dat toe te schrijven is aan die betreffende risicofactor. Op de website van het RIVM (www.nationaalkompas.nl) zijn daar voor de grote ziektebeelden diverse voorbeelden van te vinden.

CORRESPONDENTIEADRES

Dr. Willy-Anne H.J. van Stiphout, vrijgevestigd sociaal geneeskundige en docent epidemiologie, www.decijfersdebaas.nl of e-mail: stip.info@xs4all.nl