

Meerdere risicofactoren tegelijk

De ontdekking dat je verschillende risicofactoren voor één of meerdere ziekten hebt, kan beangstigend zijn voor een individu. Zoveel bedreigende en tegenstrijdige informatie kan verwarrend werken en juist een defensieve, vermijdende reactie veroorzaken. In plaats van gezonder te gaan leven, wat toch de bedoeling is, verandert er vervolgens dus niks.

Voor veel mensen is het bovendien onmogelijk om tegelijkertijd te stoppen met roken, minder vet te gaan eten, meer te bewegen en minder alcohol te gaan drinken.

Daarom moeten prioriteiten gesteld worden, bijvoorbeeld op basis van iemands motivatie of eigen effectiviteitsverwachtingen om een bepaald gedrag te veranderen, of op de mate waarin de reductie van de risicofactor een bijdrage

levert aan het reduceren van de risico's. De vraag is of een leek zonder extra informatie een goede afweging maakt.

Maar hoe moet het dan wel?

De (huidige) site geeft aan waar je gevaren zitten en wat je kunt doen om je risico te beperken. Dat is al heel wat, want mensen worden zich op die manier bewust van hun ongezonde leefstijl. Maar veel mensen wéten inmiddels wel dat ze beter zouden kunnen stoppen met roken, meer moeten bewegen en minder vet maar meer fruit en groente moeten eten. Kennis en bewustwording alleen is niet genoeg om gedrag te veranderen. Dus zou de site juist verder moeten gaan, bijvoorbeeld door mensen te motiveren tot gedragsverandering door een op de persoon toegesneden voedings- en beweegadvies. Een dergelijk advies zou dan

gericht in kunnen gaan op de doelen die de persoon voor zichzelf stelt, feedback geven op belangrijke determinanten van gedrag en tips geven voor verandering. Bij een follow-up kan nagegaan worden of de doelen zijn bereikt, eventueel gevolgd door nieuwe feedback. Kortom: voor ontwikkelaars van een vergelijkbare Nederlandse site is er nog genoeg om over na te denken.

CORRESPONDENTIEADRES

Dr. Jascha de Nooijer, Vakgroep
Gezondheidsvoorlichting, Universiteit
Maastricht, e-mail:
j.denooyier@gvo.unimaas.nl

DE CIJFERS DE BAAS

5. CONFOUNDING

Willy-Anne van Stiphout, Jan Burema

Confounding, een verwarrend begrip? Daar proberen we in deze bijdrage dan iets aan te doen. We plaatsen het begrip in zijn context (fouten die bij onderzoek op kunnen treden), leggen uit wat ermee bedoeld wordt en hoe je er, vooralsnog alleen 'in gedachten', al redenerend, voor kunt corrigeren.

Bij het bedenken en uitvoeren van onderzoek en bij het interpreteren van de resultaten kan het een en ander mis gaan. We onderscheiden daarbij toevallige en systematische fouten (zie schema 1). Over toevallige fouten hebben we het al eens gehad: bij alles wat we meten speelt toeval een rol. Door het berekenen van een 95%-betrouwbaarheidsinterval (b.i.)

rondom onze uitkomstwaarde kunnen we aangeven hoe groot de rol van het toeval geweest kan zijn. Het betrouwbaarheidsinterval geeft immers aan welke andere waarden dan die ene uitkomst ook plausibel zijn. Bijvoorbeeld: we vinden tussen lichamelijk niet-actieven en actieven een verschil in cholesterol van 0,4 mmol/L met een 95%-b.i. liggend tussen 0,1 en 0,7 mmol/L. Met andere woorden: er is een verschil van 0,4 mmol/L gevonden, maar alle andere waarden tussen 0,1 en 0,7 mmol/L zijn ook plausibel. Dit is een zeer bruikbare kwantificering van onze onzekerheid.

Schema 1 Soorten fouten

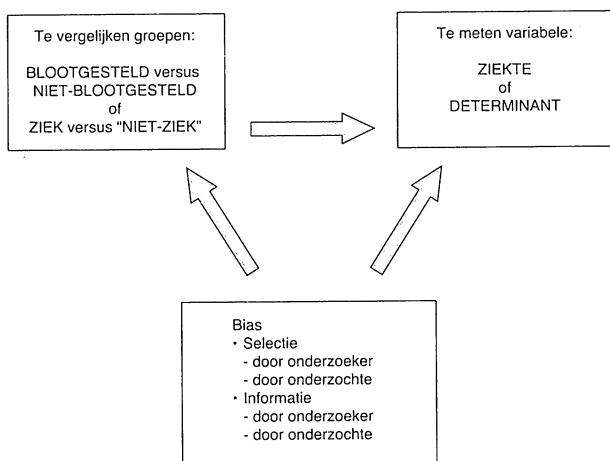
Soorten fouten	Hoe mee om te gaan?
- toevallige	Inzichtelijk maken m.b.v. 95% betrouwbaarheidsinterval
- systematische	selectie- en informatiebias Alleen te vermijden bij opzet van het onderzoek Niet corrigeerbaar in de analyse, alleen de richting kan bepaald worden
	confounding Te vermijden bij opzet van het onderzoek of (tot op zekere hoogte) kwantitatief corrigeerbaar in de analyse

SELECTIEBIAS EN INFORMATIEBIAS

Anders is het bij systematische fouten: hier gaat het erom of je mag verwachten dat de uitkomst (op toevalsfluctuaties na) de 'werkelijkheid' weergeeft. Om bij hetzelfde (deels fictieve) voorbeeld te blijven: is het verschil (afgezien van de toevalsfout) inderdaad ca. 0,4 mmol/L, of is het eigenlijk 0 mmol/L (geen verschil dus) of wellicht -0,4 mmol/L (een omgekeerd effect) of wellicht 1,5 mmol/L (een veel groter effect)?

Systematische fouten worden ook wel aangeduid met de Engelstalige term: bias, hetgeen vertekening betekent. Er zijn twee verschillende soorten 'bias' te onderscheiden (zie ook schema 1). Enerzijds vertekening die kan ontstaan bij de opzet of uitvoering van het onderzoek, bijvoorbeeld bij de selectie van de onderzoekspersonen (selectiebias) of bij het verzamelen van de informatie (informatiebias). De "fout" kan zowel gemaakt worden door de onderzoeker (hij of zij maakt een verkeerde keuze of een fout bij het informatie verkrijgen) als door de onderzochte (hij of zij besluit wel of niet aan een onderzoek mee te doen of geeft, al dan niet bewust, verkeerde informatie).

Eventuele fouten leiden pas echt tot bias als deze gerelateerd zijn aan de te meten variabele (afhankelijk van het soort onderzoek is dat de ziekte of de determinant) én als deze verschillend zijn voor de groepen die onderling vergeleken worden (zie schema 2). Als bijvoorbeeld in beide groepen een bepaald deel van de onderzoekspersonen geen zin heeft om mee te doen, maar deelname niets te maken heeft met het cholesterolgehalte, dan leidt deze non-respons niet tot selectiebias. Maar als in de niet-actieve groep met name degenen met een hoog cholesterolgehalte niet mee willen doen en in de actieve groep doet men mee, ongeacht de hoogte van zijn cholesterol, dan kan dit wel een vertekening geven (in werkelijkheid zou het cholesterolgehalte in de niet-actieve groep hoger zijn, en dus ook het verschil met de actieve groep). Kenmerkend voor deze twee vormen van bias (selectie- en informatiebias) is dat er in de analyse niet voor kan worden gecorrigeerd. Er kan uitsluitend beredeneerd worden in welke richting het verschil op zou schuiven als er een bepaalde vertekening zou hebben plaatsgevonden.



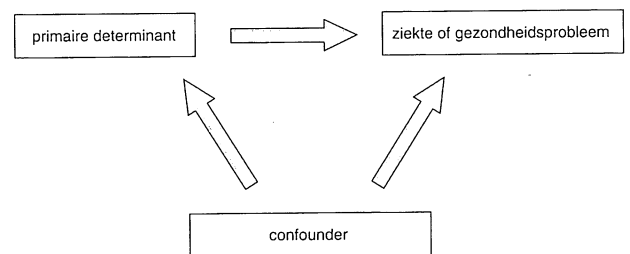
Schema 2 Schematische voorstelling van bias

CONFOUNDING

Dat is een groot verschil met die andere vorm van bias: de confounding (bias). Hiervoor kan wel in de analyse worden gecorrigeerd (mits informatie over confounders bij iedereen is verzameld). En, in tegenstelling tot selectie- en informatiebias, die afhankelijk zijn van de opzet en uitvoering van het onder-

zoek, is het gevaar van confounding in observationeel onderzoek altijd aanwezig, want het is inherent aan de vraagstelling. Dat komt omdat een confounder niets anders is dan een andere determinant van de te bestuderen ziekte of het gezondheidsprobleem. Omdat we bij etiologisch onderzoek doorgaans het 'zuivere' effect van één bepaalde determinant willen weten, is de mogelijkheid van confounding in principe altijd aanwezig. Letterlijk betekent confounding dan ook: verarring. Het effect van de ene determinant wordt verward met dat van een andere.

Of een potentiële confounder in de onderhavige data ook een 'echte' confounder blijkt te zijn, hangt ervan af of deze variabele ook verschillend verdeeld is over de primaire determinant (zie schema 3). Laten we opnieuw even kijken naar de relatie tussen lichamelijke activiteit en het cholesterolgehalte bij volwassenen. Indien niet-actieve personen gemiddeld ouder zijn dan actieve, en ouderen hebben een hoger cholesterolgehalte, dan zou leeftijd (mede) het verschil tussen actieven en niet-actieven kunnen verklaren. Echter als een andere variabele, zoals bijvoorbeeld de vetinname in beide groepen precies gelijk verdeeld is, dan zal deze (hoewel het een sterke determinant van cholesterol is) toch de relatie tussen lichamelijke activiteit en het cholesterolgehalte niet verstoren, en dus geen confounder blijken te zijn. Overigens is het wel of niet confounder-zijn niet iets wat je met een statistische toets kunt bepalen. Ook een variabele die maar een klein beetje verstorend werkt komt toch in aanmerking om bij de analyse betrokken te worden.



Schema 3 Schematische voorstelling van confounding

VERMIJDEN VAN OF CORRIGEREN VOOR CONFOUNDING

Gelukkig is confounding goed vermijdbaar dan wel behandelbaar. Essentieel daarvoor is dat de onderzoeker veel weet van de te bestuderen ziekte. Immers, andere determinanten (potentiële confounders) moeten bekend zijn wil men ervoor kunnen zorgen dat de schattingen (hier: van het cholesterolgehalte) in beide determinantgroepen vergelijkbaar zijn dan wel worden.

Zo kan confounding in de *onderzoekopzet* worden vermeden door restrictie toe te passen (alleen een bepaalde leeftijdsgroep of alleen niet-rokers mogen aan het onderzoek meedoen) of te matchen, hetgeen met name in patiënt-controle onderzoek gebeurt. De controlegroep wordt dan gekozen uit personen die qua verdeling van een aantal potentiële confoun-

ders (zoals leeftijd en geslacht) gelijk zijn aan de 'zieken'. Het randomiseren (het toeval laten bepalen wie in welke determinantgroep komt) bij experimenten (trials) leidt er eveneens toe dat de twee determinantgroepen, naar we mogen hopen, vergelijkbaar zijn ten aanzien van zowel bekende als ook nog onbekende confounders.

Bij de analyse kan voor confounding gecorrigeerd worden, mits informatie over de primaire determinant, de ziekte en de confounders voor iedere onderzochte persoon op individueel niveau beschikbaar is. Er kan altijd vooraf beredeneerd worden welke richting correctie voor confounding opgaat. Kijken we nog eens naar het voorbeeld van lichamelijke activiteit en het cholesterolgehalte. Stel we vinden dat niet-actieve personen (rechts in de figuur) een hoger cholesterol hebben dan actieve personen (zie figuur 1). Maar tegelijkertijd blijkt een groter deel van de niet-actieve personen ouder te zijn (gearceerd gedeelte). Correctie betekent niets anders dan dat we beide determinantgroepen 'gelijk maken' ten aanzien van iets dat nu ongelijk is, namelijk de verdeling van de confounder, hier: leeftijd. (In dit voorbeeld onderscheiden we voor het gemak slechts twee categorieën: jong en oud). We weten dat het cholesterolgehalte hoger is bij ouderen. De niet-actieve groep heeft meer ouderen. Als we ons voorstellen wat de consequentie is voor het cholesterolgehalte wanneer het aandeel ouderen in de niet-actieve groep een minder grote bijdrage (gewicht) levert aan het totaal, dan is duidelijk dat het gemiddelde cholesterolgehalte in deze groep zal dalen. Omgekeerd wanneer de bijdrage van de ouderen in de actieve groep iets meer gewicht krijgt, om het aandeel van de ouderen even groot te maken als in de niet-actieve groep, dan zal dit gepaard gaan met een stijging van de gewogen schatting van het cholesterolgehalte in deze groep. Het resultaat is dat het verschil in cholesterol kleiner zal worden, mogelijk geheel verdwijnt, of zelfs omkeert. Als het verschil geheel verdwijnt dan kan gesteld worden dat

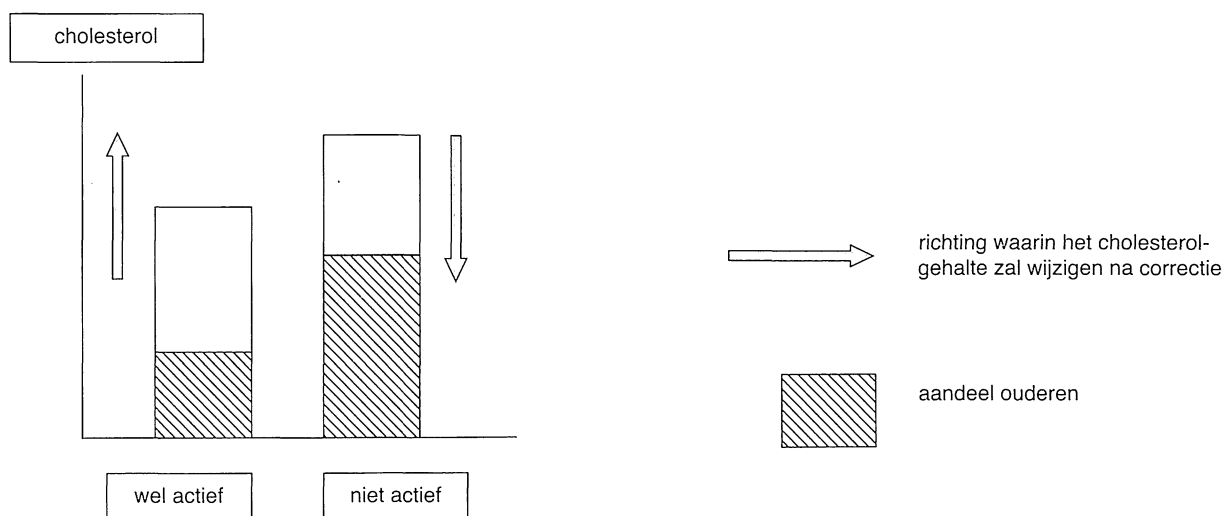
niet het verschil in lichamelijke activiteit maar het verschil in leeftijd verantwoordelijk was voor het gevonden verschil in cholesterolgehalte. Een mooi voorbeeld van verwarring.

Onder bepaalde aannamen is het mogelijk exact uit te rekenen hoe groot het verschil tussen de cholesterolgehalten na correctie voor confounding is. Ook is het mogelijk om voor meerdere confounders tegelijkertijd te corrigeren, al wordt dat niet alleen erg bewerkelijk, maar zelfs onmogelijk als de aantallen per subgroep niet groot genoeg zijn. Dan moeten we onze toevlucht wel nemen tot modellen, al zitten daar dan ook wel weer haken en ogen aan. De volgende keer zullen we ingaan op twee statistische modellen en hoe de coëfficiënten in dit kader dan geïnterpreteerd kunnen worden.

Voor de lezer hopen we het begrip 'confounding' iets toegankelijker te hebben gemaakt. Met wellicht de geruststellende gedachte dat het effect van correctie in ieder geval steeds beredeneerd kan worden, mits men weet hoe de confounder verdeeld is over categorieën van de primaire determinant en hoe de relatie tussen de confounder en de ziekte is. Daarom roepen we auteurs op om in artikelen aan te geven op welke relevante factoren (potentiële confounders) de groepen in welke richting verschillen en welke relatie die confounders hebben met de uitkomstvariabele (positief of negatief effect). Dat maakt het voor lezers mogelijk meer inzicht te krijgen in de statistische analyse.

CORRESPONDENTIEADRES

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



Figuur 1 Schematische voorstelling van de gemiddelde waarden van de cholesterolgehalten en het aandeel ouderen bij wel- en niet-actieven