



Ontbrekende waarnemingen

PRAKTIJKPROBLEEM

Lang geleden, toen er nog geen NHG-Standaard Cardiovasculair risicomangement was, waren veel huisartsen gewend om de rookgewoontes van de patiënt alleen in het dossier te vermelden als de patiënt ook daadwerkelijk rookte. Voortschrijdend inzicht wees uit dat dit niet zo handig is. Het ontbreken van een rooknotitie kan twee dingen betekenen: deze patiënt rookt niet, of het is hem nooit gevraagd en we weten het niet.

Ook bij wetenschappelijk onderzoek kan onduidelijkheid bestaan over de interpretatie van op het eerste gezicht duidelijke informatie. Waar moet je op letten bij het lezen van onderzoeksartikelen als het gaat om ontbrekende gegevens?

ACHTERGROND

Bij wetenschappelijk onderzoek hebben we vaak te maken met ontbrekende waarnemingen, en die zijn er in vele soorten. Het komt voor dat deelnemers aan een vragenlijstonderzoek niet alle vragen beantwoorden. In cohortonderzoeken waarbij deelnemers op gezette tijden metingen ondergaan, zijn er mensen bij wie sommige metingen niet gedaan (blijken te) zijn. In een randomised controlled trial ontbreken van sommige deelnemers bijvoorbeeld de laboratoriumbepalingen omdat er iets mis is gegaan met de bloedafname.

UITWERKING

Vaak is het ontbreken van een deel van de gegevens niet zo erg: bij de mislukte bloedafname zal toeval een grote rol spelen en het is dan niet waarschijnlijk dat de bloedwaarden van deze persoon erg afwijken van die van de andere deelnemers.

Anders is het wanneer mensen uitvallen vanwege aan het onderzoek gerelateerde redenen. Neem het voorbeeld van een gerandomiseerde placebocontroleerde trial waarbij patiënten in de interventiegroep de interventie zo belastend vinden in verhouding tot het ervaren effect, dat ze er na enige tijd geen zin meer in hebben. Als de uitvallers zich onttrekken aan de vervolgmetingen, zijn de overblijvers geen goede afspiegeling van de hele groep en volgen er te rooskleurige conclusies.

Soms is uitval inherent aan het onderzoek. De Leiden 85-plus Studie, waarin een grote groep mensen van 85 jaar en ouder werd gevolgd om inzicht te krijgen in factoren die samenhangen met ziekte en sterfte op hoge leeftijd, had begrijpelijkerwijs bij ieder volgend meetmoment te maken met een steeds kleiner wordende onderzoekspopulatie.

Bij het beoordelen van onderzoeken, bijvoorbeeld bij het schrijven van Cochrane-reviews, zijn ontbrekende metingen een belangrijk criterium om de kans op vertekende resultaten in te schatten. Het follow-uppercentage dat nog door de beugel kan is een arbitraire keuze. Bij Cochrane-reviews wordt de grens gelegd bij 80%. Is in een trial het percentage deelnemers van wie vervolgmetingen zijn verkregen lager dan 80%, dan krijgt het onderzoek op dit onderdeel als oordeel 'high risk of bias'.

Er zijn oplossingen voor het probleem van de ontbrekende waarnemingen, ook als de dataverzameling al is afgesloten. Een simpele werkwijze is om, wanneer de bloeddruk van enkele deelnemers niet gemeten is, hiervoor de laatst bekende meting te gebruiken, of bij de personen zonder meting het gemiddelde van alle metingen in te vullen. De laatste decennia zijn er statistische technieken ontwikkeld om ontbrekende waarnemingen op een meer verantwoorde manier 'op te vullen'. We noemen dit imputeren. Daarbij wordt voor een ontbrekende waarde in de gegevens gebruikgemaakt van de waarden van vergelijkbare deelnemers, maar bovendien een snufje toeval (*random error*) toegevoegd. Bij zogenaamde multipele imputatie wordt op deze wijze eenzelfde dataset op verschillende manieren aangevuld en vindt de uiteindelijke analyse plaats op meerdere datasets. Deze techniek is tegenwoordig de *state of the art*.

BETEKENIS

In onderzoeken zijn ontbrekende waarnemingen vaak niet te voorkomen. Het is immers altijd mogelijk dat een meting (bijvoorbeeld bloeddrukmeting) mist omdat de patiënt niet kwam opdagen of omdat de onderzoeker de meting was vergeten uit te voeren. Let bij het lezen van onderzoeksartikelen op de omvang van de uitval, de beschrijving van de follow-up en eventuele verklaringen voor de ontbrekende waarnemingen. Een hoge uitval vermindert de betrouwbaarheid van de resultaten, vooral als de uitval niet willekeurig was. Er zijn statistische technieken beschikbaar om ontbrekende gegevens aan te vullen. ■

LITERATUUR

- 1 Donders AR, Van der Heijden GJ, Stijnen T, Moons KG. Review: a gentle introduction to imputation of missing values. *J Clin Epidemiol* 2006;59:1087-91.
- 2 Bell ML, Fiero M, Horton NJ, Hsu CH. Handling missing data in RCTs; a review of the top medical journals. *BMC Med Res Methodol* 2014;14:118.

De serie Praktische epidemiologie laat zien dat er een wetenschappelijke onderbouwing bestaat voor veel handelingen die de huisarts in de dagelijkse praktijk intuïtief uitvoert. Aan de hand van een herkenbaar praktisch gegeven in de praktijk geven we kort aan hoe de wetenschap achter dit praktijkprobleem in elkaar zit.
Correspondentie: j.eekhof@nhg.org

EMGO+ VUmc, afdeling Huisartsgeneeskunde & Ouderengeneeskunde, De Boelelaan 1117, 1081 HV Amsterdam; dr. J.C. van der Wouden, universitair hoofddocent. Radboudumc, IQ healthcare; dr. M.K. Dees, senior onderzoeker • Correspondentie: j.vanderwouden@vumc.nl