

## Op jacht naar de causale verklaring met een statistische speurhond: een geschikte geurproef?

### Inleiding

We koesteren allemaal de diepgewortelde overtuiging dat iedere verandering of iedere gebeurtenis een oorzaak zal hebben en daarom veronderstellen we dat er steeds een antwoord te vinden moet zijn op de vraag wat de causale verklaring van zo'n verandering of gebeurtenis is – hoe moeilijk dat soms ook kan zijn. Maar als we zoeken naar oorzaak-gevolgrelaties, wanneer kunnen we dan met voldoende zekerheid iets als *de* oorzaak of oorzaken aanwijzen? Die juridisch relevante vraag is primair wetenschapsfilosofisch van aard en houdt denkers al millennia lang bezig, ook binnen het recht.<sup>1</sup> Maar wat is nu precies het begrip 'oorzaak'? Een definitief antwoord daarop hebben we nog steeds niet gevonden: 'Despite the attention, there is still very little agreement on the most central question: what is it?'.<sup>2</sup> En, gegeven deze worsteling, hoe zeker is het juridische bewijs van causaliteit?

Met name in het aansprakelijkheids- en het strafrecht is het vaststellen van het causale verband van cruciaal belang voor het juridische eindoordeel. Vanwege die praktische noodzaak echter zullen we in de rechtspraak ondanks alle kennistheoretische mitsen en maren bij het vaststellen van een oorzakelijke relatie wel moeten manoeuvreren binnen de juridische en praktische grenzen en zo zoeken naar de best mogelijke causale verklaring van een incident, een calamiteit of een misdaad.<sup>3</sup> Draagt de gedaagde of verdachte door diens gedragingen de verantwoordelijkheid voor het ontstaan van deze situatie of onheilsgebeurtenis?<sup>4</sup> Bij de speurtocht naar de juiste uitleg zullen we daarom een gefundeerde keuze dienen te maken voor de meest betrouwbare onderzoeksmethode en die kan hetzij kwalitatief of kwantitatief van aard zijn – of een combinatie daarvan. Binnen de juridische context heeft de beantwoording van de causaliteitsvraag niet alleen een juridische maar ook een wetenschappelijke dimensie. De juridische dimensie omvat vanzelfsprekend de materieel- en formeelrechtelijke regelgeving, ingebed in de doctrine. Alleen al juridisch gezien, is en blijft het een lastige kwestie.<sup>5</sup> Het zoeken naar de juiste verklaring vraagt een wetenschappelijk onderbouwde methodologie, maar dat is ook bepaaldelijk niet onproblematisch.<sup>6</sup> Praktisch wordt er vaak een beroep gedaan op een of meer deskundigen,

die het onderzoek naar de causale verklaring zullen uitvoeren en rapporteren. Een van de mogelijkheden die hun ter beschikking staan, is daarbij gebruik te maken van een kwantitatieve benadering waarbij met behulp van statistische onderzoeksmethoden gezocht wordt naar een causaal verband tussen het beweerde onrechtmatige gedrag en de afloop; het leveren van een statistisch bewijs. Overigens is het in dit verband essentieel erop te wijzen dat aan de term 'bewijs' uiteenlopende betekenissen kunnen worden toegekend. In het strafrecht wordt het begrip 'bewijs' als volgt gedefinieerd: 'Dat kan worden omschreven als aantonen dat in redelijkheid niet kan worden getwijfeld aan de juistheid van het verwijt dat aan de verdachte wordt gemaakt. Absolute zekerheid hoeft er niet te zijn. Wiskundig bewijs waar geen speld tussen is te krijgen, wordt niet vereist.'<sup>7</sup> Deze begripsomschrijving is tegelijk heel juridisch en praktisch georiënteerd, maar een epistemoloog – de wetenschapper op het gebied van kennis – zal deze zin met gefronste wenkbrauwen lezen.

Als je iets in maat en getal weet uit te drukken, geeft dat al gauw het idee van objectiviteit. Maar juristen worden tijdens hun studie niet systematisch geschoold in onderzoeksmethodologie, laat staan in kwantitatieve denkmethoden en de meeste van hen zijn dus ongecijferd.<sup>8</sup> Maar als rechters niet geschoold zijn in dergelijke rekenmethodes hoe kunnen zij de (on)juistheid van een statistisch bewijs dan beoordelen?<sup>9</sup>

Elders in dit nummer treft u de bespreking aan van het boek *Math on trial* door Aitken. Niet voor het eerst wordt daarbij kritisch gekeken naar het gebruik van statistiek in de rechtszaal. In het boek worden gerechtelijke dwalingen beschreven, die hun oorzaak vonden in op kwantitatieve methoden gebaseerde getuigenissen of deskundigenrapporten, opgesteld door wél gecijferde geleerden. Maar de essentiële vraag die ook in deze bespreking toch nog onbeantwoord blijft is: is er überhaupt wel een plaats voor statistische methoden voor het beantwoorden van de kernvraag wie er verantwoordelijk gehouden kan worden voor de te beoordelen rampzalige gebeurtenis, het causale verband? Hoe betrouwbaar – en dus bruikbaar – is het statistisch bewijs inzake causaliteitskwesties?

\* Prof. dr. dr. R.W.M. Giard is hoogleraar methodologie en aansprakelijkheid aan de Erasmus School of Law, Erasmus Universiteit Rotterdam en redacteur van dit tijdschrift. Met dank aan Gert Haverkate, Ton Broeders en Diederik Aben voor hun commentaren en suggesties.

1. Bekend en veel geciteerd is het werk van H.L.A. Hart & T. Honoré, *Causation in the law*, Oxford: Clarendon Press 1959 (met een heruitgave in 2002).
2. H. Beebe, C. Hitchcock & P. Menzies (red.), *The Oxford handbook of causation*, Oxford: Oxford University Press 2009.
3. Zie A. Amaya, 'Inference to the best explanation', in: H. Kaptein et al. (red.) *Legal evidence and proof. Statistics, stories, logic*, Farnham, Ashgate Publishing Ltd 2009, p. 135-159.
4. M.S. Moore, *Causation and responsibility. An essay in law, morals, and metaphysics*, Oxford: Oxford University Press 2009.
5. Chr.H. van Dijk, 'Causale perikelen: het is moeilijk en zal moeilijk blijven', *TVP* 2013-3, p. 61-84.
6. Zie de bundel R. Goldberg (red.), *Perspectives on causation*, Oxford: Hart Publishing, 2011 en het in noot 2 genoemde boek *The Oxford handbook of causation*.
7. G.J.M. Corstens, *Het Nederlands strafprocesrecht*, Deventer: Kluwer 2008 (6e druk), p. 664.
8. De term werd geïntroduceerd door John Allen Paulos in zijn boek *Ongecijferdheid*, Amsterdam: Uitgeverij Bert Bakker 1989.
9. Zie ook W.H. van Boom & M.J. Borgers (red.), *De rekenende rechter. Van 'iudex non calculat' naar actieve cijferaar?*, Den Haag: Boom Juridische uitgevers 2004.

### De denkwereld van de statisticus

Wat omvat de term statistiek nu precies en wat zou een jurist van de basisprincipes van die denkwereld moeten weten? Deze hulp(!)wetenschap kan worden omschreven als een verzameling van verschillende empirische methoden om op systematische wijze uitspraken te kunnen doen over populaties. Dat 'leren' behelst hetzij het getalsmatig beschrijven van verschijnselen (inventariseren), hetzij het toetsen van veronderstellingen (hypothesen). Met de term 'populatie' wordt een nauwkeurig omschreven groep al dan niet levende objecten bedoeld die aan een onderzoek wordt onderworpen. Aan zo'n exploratie ligt steeds een concrete vraagstelling ten grondslag en die bepaalt welke data er gezocht gaan worden en welke statistische methode voor analyse daarop zal worden toegepast. Steevast zijn er, na het stellen van de onderzoeksvraag of -vragen, drie belangrijke kwesties:

1. De *steekproef*: hoe kom ik aan mijn onderzoeksobjecten? We worden praktisch gedwongen een keuze te maken: hoe heeft de selectie plaatsgevonden en wat kon er daarbij fout gaan?
2. De *opzet*: wat ga ik onderzoeken en waarom? Neem ik eenvoudig waar of experimenteer ik?
3. De *waarnemingen of metingen*: hoe kom ik aan mijn resultaten en hoe accuraat en reproduceerbaar zijn die methoden?

Het antwoord op die vragen naar relevantie, representativiteit, betrouwbaarheid en bruikbaarheid is van groot belang om te kunnen vaststellen of de uitkomsten van zo'n onderzoek in de rechtszaal gebruikt mogen worden voor het oplossen van de voorliggende normatieve juridische kwestie.

De uitkomsten van inventariserend statistisch onderzoek kunnen getalsmatige beschrijvingen zijn in de vorm van rechte tellingen ('hoeveel rechters zijn er in Nederland?') of percentages ('hoeveel procent van de rechters is vrouw?'). Dat is *descriptieve* statistiek: die beschrijft slechts. Daarnaast is er ook de *analytische* vorm: die probeert langs inductieve weg conclusies te trekken door bepaalde kenmerken van populaties nader te analyseren. De *inferentiële* (=analytische) statistiek kan dan weer hetzij observationeel hetzij experimenteel van aard zijn. Een tijdreeksanalyse kijkt naar veranderingen: 'neemt het aantal rechtszaken sinds 2010 toe of af?' of 'stijgt het percentage vrouwelijke rechters de laatste jaren?'. Een andere mogelijkheid is het zoeken naar verbanden, waarbij wordt gekeken naar relaties tussen verschillende kenmerken binnen een of meerdere groepen die worden onderzocht en vergeleken ('oordelen vrouwelijke rechters anders dan mannelijke?').

Een klassiek voorbeeld is natuurlijk de vraag of er een oorzakelijk verband bestaat tussen roken en longkanker. Bij zo'n onderzoek worden groepen rokers en niet-rokers langdurig geobserveerd en gebeurtenissen (eindpunten: kreeg wel of geen longkanker, gestorven) met elkaar vergeleken. Als blijkt dat de frequentie longkanker hoger is binnen de groep rokers worden de uitkomsten vervolgens nog op verschillende wijzen statistisch getoetst om te kunnen beoordelen of het resultaat eventueel aan het toeval kan worden toegeschreven. Maar als longkanker

statistisch significant vaker bij rokers voorkomt, dan wordt er op deze wijze vooralsnog slechts een associatie tussen roken en longkanker vastgesteld, maar nog geen causaal verband – hooguit een vermoeden daarvan.

Bij zo'n observationeel vergelijkend onderzoek is methodologisch gezien de *ceteris paribus*-doctrine van eminent belang: de te vergelijken groepen moeten, behoudens het onderzochte kenmerk, qua samenstelling en andere kenmerken volledig identiek zijn. Zie bijvoorbeeld de uitkomsten van het onderzoek naar het verband tussen rookgedrag, waarbij als uitkomstmaat de sterftekansen per 1000 personen per jaar werden berekend (tabel 1).

Tabel 1. Sterftekans en rookgedrag<sup>10</sup>

Groep	Ongecorrigeerd	Gecorrigeerd qua leeftijd
Niet-rokers	20,2	20,3
Sigarettenrokers	20,5	28,3
Sigaar- en pijprokers	35,3	21,2

Uit de tweede kolom blijkt verrassend dat het sterfterisico het grootst is bij de sigaar- en pijprokers. Nader onderzoek van die drie groepen liet echter zien dat de rokersgroepen onderling qua gemiddelde leeftijden verschilden (resp. 54.9, 50.5 en 65.9) waarbij de sigaar- en pijprokers dus het oudst waren en alleen al daardoor een grotere sterftekans hadden. Correctie voor de leeftijd leverde een ander beeld op (derde kolom) waarbij nu het risico voor sigarettenrokers veel hoger bleek. De vergeleken groepen bleken door een selectiemechanisme op meer punten verschillend dan alleen het rookgedrag: hier was sprake van vertekening ('bias'). Daarnaast kan, als de vergelijkbaarheid wél klopt, er nog een andere vooralsnog onbekende en daardoor niet onderzochte factor (een 'confounder') in het spel zijn die een verschil verklaart.<sup>11</sup> Vandaar de mantra die ook iedere jurist beslist moet kunnen dromen: *association is not causation*. In civiele rechtszaken is de vraag naar voren gekomen of er een causale relatie kon worden vastgesteld tussen beroepsmatige asbestblootstelling en mesothelioom of tussen het slikken door vrouwen van DES tijdens de zwangerschap en het optreden van gynaecologische afwijkingen en tumoren bij hun dochters. Het is niet alleen onethisch om experimenten met mensen te gaan verrichten om die relatie te bewijzen, het is ook onpraktisch want het vergt vele jaren van onderzoek eer er resultaten zullen zijn. Daarom is men aangewezen op retrospectief observationeel onderzoek. Maar dan ontmoeten we weer de genoemde associatiemantra: wat dan? Zijn er argumenten denkbaar waarmee we een causale relatie voldoende aannemelijk kunnen maken?

Om bij associaties, geconstateerd bij retrospectieve observationele onderzoeken, de waarschijnlijkheid van een causale relatie aannemelijk te maken, werden met name in de geneeskunde criteria geformuleerd zoals de sterkte van de associatie, het constateren ervan bij meerdere onderzoeken (consistentie), de specificiteit voor die ene factor, de relatie in tijd, het bestaan van een biologische

10. Tabel ontleend aan G.W. Cobb & S. Gehlbach, 'Statistics in the courtroom: US v. Kristen Gilbert', <http://math.buffalostate.edu/~wilsondc/MED588/KristenGilbert.pdf>.

11. Zie H.I. Weisberg, *Bias and causation. Models and judgment for valid comparisons*, Hoboken, New Jersey: John Wiley & Sons 2010.

gradiënt (hoe sterker de expositie, des te groter de kans op ziekte), de biologische plausibiliteit van zo'n causale relatie en de coherentie ervan binnen de bestaande medische kennis.<sup>12</sup> Zowel voor DES als asbest kon de causale betekenis daardoor voldoende – maar niet met absolute zekerheid! – vastgesteld worden en daarmee als basis dienen voor het juridische oordeel.

Tot nu toe zijn dit allemaal voorbeelden van niet-experimenteel – dus observationeel en meestal retrospectief – onderzoek. Daarnaast bestaat er nog een experimentele tak van statistiek, waarbij prospectief gerandomiseerd vergelijkend onderzoek wordt gedaan met als grondplan twee door loting verkregen onderzoeksgroepen waarbij de een wel en de ander niet aan een te onderzoeken agens of interventie wordt blootgesteld. Mits de onderzoeksgroepen voldoende groot zijn, worden door de lotingsprocedure twee min of meer gelijkwaardige groepen verkregen omdat het lot ze qua leeftijd, leefgewoontes en risicofactoren 'homogeniseert', zodat de onderzoeksgroepen alleen maar zullen verschillen op het punt van de wel of niet uitgevoerde interventie – weer de genoemde ceteris paribus-eis. Dit is de best mogelijke methode voor het aantonen van een causaal effect. Deze onderzoeksvorm wordt vooral in de geneeskunde toegepast, maar kent ook praktische voorbeelden in bijvoorbeeld de criminologie.<sup>13</sup> De uitkomst is steeds een waarschijnlijkheidsoordeel.

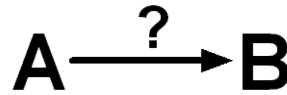
Belangrijk is dat we ons realiseren dat statistiek een discipline is die als antwoord op de gestelde onderzoeksvraag of -vragen de (on)zekerheden daarvan beschrijft en kwantificeert. Maar dat onderzoek wordt op grond van praktische motieven steeds uitgevoerd op qua aantallen beperkte en geselecteerde groepen en niet door alle in deze wereld beschikbare elementen of individuen te onderzoeken die daarvoor in aanmerking zouden kunnen komen. De uitkomsten van onderzoek spreken nooit vanzelf, maar moeten steeds worden geïnterpreteerd. Inherent aan de gehanteerde methoden en door het gebruik van steekproeven, bestaat er kans dat de uitkomsten worden beïnvloed door systematische fouten en toevalsfactoren en daarom is grote voorzichtigheid geboden met het doen van algemeen geldende uitspraken. Uitkomsten van descriptieve en analytische statistiek kunnen behulpzaam zijn bij de opzet en uitvoering van het proces van waarheidsvinding, bijvoorbeeld welke factoren belangrijk zijn en dus onderzocht moeten worden, hoe betrouwbaar informatie is etc. Tot zover globaal de denkwereld van de statistiek maar we zullen ons nu verder moeten concentreren op de vraag of, en zo ja hoe causale relaties met statistische methoden kunnen worden aangetoond.

#### Statistisch bewijs bij het vaststellen van causaal verband

Waar de vraag 'is A verantwoordelijk voor B?' luidt, zal de rechter het causaliteitsbegrip deterministisch hantieren, terwijl de statisticus causaliteit als een kwantificeer-

bare maar onzekere relatie tussen A en B zal zien, die aan de hand van berekende waarschijnlijkheden aansluitend moet worden gekwalificeerd, bijvoorbeeld 'vrijwel zeker', 'zeer waarschijnlijk', 'mogelijk' of 'onwaarschijnlijk'.

Stel, de rechter wordt geconfronteerd met de vraag of A de oorzaak is van B, zoals weergegeven in onderstaand schema:



Dan zijn er vanuit het perspectief van richting van de tijd twee essentieel verschillende typen causale vragen denkbaar.<sup>14</sup>

- *Prospectief*: zal A vervolgens B veroorzaken? Dan hebben we het over effecten van oorzaken (=EvO), een vooruitblikkende vraag.
- *Retrospectief*: werd B veroorzaakt door A? Dan gaat het om oorzaken van effecten (=OvE), een terugblikkende vraag.

Het EvO-type vraag is algemeen van aard en treft men in een wetenschappelijke context aan. Dan is het onder meer mogelijk om aan de hand van experimenteel onderzoek een causale relatie aan te tonen, zoals het hierboven gegeven voorbeeld van de gerandomiseerde studie. Het OvE-type vraag daarentegen handelt over kwesties rond individuele veroorzaking zoals we die in de rechtspraak aantreffen: werd B vermoord door A of was de schade van B het gevolg van onrechtmatig handelen van A? Gezien het retrospectieve karakter kan alleen van uitkomsten van observationeel onderzoek gebruikgemaakt worden. Kan bij het beantwoorden van het OvE-type vraag in een juridische context statistiek een bruikbaar én betrouwbaar antwoord leveren?

De volgende casus is illustratief. Een verpleegkundige, werkend op een afdeling intensieve zorg, werd door haar collega's ervan verdacht regelmatig patiënten om het leven te brengen. Bij het strafvorderlijke onderzoek werd de hulp van een statisticus gezocht, die een analyse maakte van de sterfgevallen op de afdeling waar zij werkte met de vraag of daartussen een verband bestond.<sup>15</sup> De uitkomsten zijn samengevat in tabel 2:

Tabel 2. Sterftetekans bij de aan- of afwezigheid van verpleegster G.<sup>16</sup>

Zuster G. aanwezig	Overleden	Niet overleden	Totaal
<b>Ja</b>	40 (16%)	217 (84%)	257 (100%)
<b>Nee</b>	34 (2%)	1350 (98%)	1384 (100%)
<b>Totaal</b>	74 (4%)	1567 (96%)	1641 (100%)

Tijdens een dienst van zuster G. heeft een patiënt inderdaad een grotere kans om te overlijden (16%), dan bij

12. Zie bijvoorbeeld Austin Bradford Hill, 'The environment and disease: association or causation?', *Proceedings of the Royal Society of Medicine*, 58, 1965, p. 295-300.

13. Zie bijv. A.A. Braga & B.J. Bond, 'Policing crime and disorder hot spots: a randomized controlled trial', *Criminology* 2008, vol. 36, nr. 3, p. 577 e.v.

14. Zie A.P. Dawid, 'The role of scientific and statistical evidence in assessing causality', in: R. Goldberg (red.), *Perspectives on causation*, Oxford: Hart Publishing 2011, p. 133-147.

15. Ik gebruik hier het voorbeeld van een Amerikaanse rechtszaak, United States Court of Appeals, First Circuit. 21 juni 1998, No. 98-1563 (*US/Gilbert*).

16. Ontleend aan G.W. Cobb & S. Gehlbach, zie noot 10.

haar afwezigheid (2%) en er is dus duidelijk sprake van een statistisch beduidende associatie daartussen. Maar we stuiten hier weer op de onverbidelijke mantra: uit de associatie valt dus nog geen causaal verband af te leiden! Er zijn meerdere verklaringen denkbaar voor deze associatie, bijvoorbeeld deze verpleegster heeft altijd late dienst, juist de periode van de dag waarop de meeste mensen overlijden. Forensisch bewijs, waaruit blijkt dat de patiënten daadwerkelijk werden omgebracht en niet 'natuurlijk' overleden, is hier het doorslaggevende argument.

Een hypothese bij het onderzoeken van het vermoeden van een door één persoon gepleegde serie levensdelicten is de stelling 'dit kan geen toeval zijn'. Bij de zaak Lucia de B.<sup>17</sup> werd door de forensisch statisticus uitgerekend hoe groot de kans is op het aantreffen van een dergelijk incidentenpatroon indien er geen verband zou bestaan tussen Lucia de B. en de overlijdensgevallen: dat was minder dan 1 op 342 000 000. De medisch-deskundige bij de zaak Sally Clark becijferde op soortgelijke wijze de kans op het natuurlijk overlijden van twee kinderen in een gezin van 1 op 73 000 000. Maar hoe kom je betrouwbaar tot zo'n conclusie? Welk rekenmodel hanteert je daarvoor? En heel belangrijk: waar haal je de getallen vandaan en hoe betrouwbaar zijn die?<sup>18</sup> Zowel bij Lucia de B. als bij Sally Clark bleek er statistisch-technisch achteraf niet alleen sprake van een onjuiste kansschatting maar – ernstiger nog – van een toepassingsfout ten aanzien van de uitkomsten, de zogeheten 'prosecutor's fallacy'.<sup>19</sup> Er werd een extreem kleine kans berekend op een gebeurtenis of een eigenschap als 'normale' situatie, maar bij zowel Lucia de B. als bij Sally Clark werd dat door procesdeelnemers geïnterpreteerd als kans op onschuld gegeven de gebeurtenis. En door dat extreem kleine getal in aanmerking te nemen, volgde daarop de interpretatie dat de gebeurtenissen géén onschuldig toeval konden zijn. We willen echter de kans op daderschap gegeven de feitelijke situatie weten en die moet anders worden berekend. Met nadruk moet bij deze misslagen gesteld worden dat het statistisch onderbouwde dit-is-geen-toevalargument primair door de gecijferde deskundigen werd aangedragen! Maar kunnen de verdediging en de rechter dergelijke fouten vaststellen?

#### **De statistische speurtocht: hebben juristen een goede neus voor de juiste getallen?**

Het boek van Daniel Kahneman *Ons feilbare denken*<sup>20</sup> is voor een belangrijk deel gewijd aan problemen rond het hanteren en interpreteren van kwantitatieve methoden en dus die van statistisch onderzoek. Aan de hand van diverse voorbeelden toont hij dat we ten onrechte een buitensporig vertrouwen koesteren in wat we denken te weten. Kahnemans overzicht presenteert ook ons kennelijke onvermogen om de volle omvang van onze onwe-

tendheid en onzekerheid over kwesties waarover we moeten oordelen te erkennen. We zijn gewaarschuwd! Als we ons op statistisch onderzoek richten, doemen twee kwesties op. Ten eerste: is een kwantitatieve benadering van de causaliteitsvraag binnen een procedure wel gerechtvaardigd? Als we die vraag bevestigend beantwoorden, volgt het probleem van de interpretatie van uitkomsten: uit psychologisch onderzoek blijkt 'praktisch' eerder 'intuïtief' dan 'rationeel' te zijn en dat maakt de juridische beslisser kwetsbaar.<sup>21</sup> Empirisch onderzoek laat derhalve zien dat ook bij gecijferden hun intuïtieve modus domineert over de rationele; we hebben vaak geen goede neus voor de juiste beoordeling van statistiek.

Wat betreft de eerste vraag is het belangrijk dat we ons realiseren dat die gaat over het aantonen van het causale verband: we kunnen dan kiezen hetzij voor een getalsmatige hetzij voor een verklarende benadering. We hebben gezien dat de benadering met de 'dit kan geen toeval zijn'-hypothese fundamenteel onjuist is. Een juiste kansrekening om te becijferen hoe waarschijnlijk het is dat iemand het misdrijf of de daad gepleegd heeft gegeven het bewijs, strandt praktisch op het gebrek aan de juiste getallen.<sup>22</sup> Ook de benadering van tabel 2 in het geval van de van moorden verdachte verpleegkundige maakt duidelijk dat nader onderzoek geboden is maar brengt ons niet verder dan een associatie.

Daarmee komen we weer terug bij de fundamentele vraag wanneer we van een causaal verband mogen spreken. Een oorzakelijk verband beschrijft een relatie tussen gebeurtenissen en is niet alleen een kwestie van dat A steeds voorafgaat aan B en dat zonder A er geen B volgt (c.s.q.n.!), maar ook dat er een genererend mechanisme moet zijn waardoor B kan plaatsvinden.<sup>23</sup> Als we de keuze hebben tussen een rekenkundig (= probabilistisch) of een verklarend model, kiezen we daarom voor de laatste. Maar is dat praktisch altijd haalbaar in de procespraktijk? Hoe verhoudt zich dat tot bijvoorbeeld art. 338 Sv ('Het bewijs dat de verdachte het telastegelegde feit heeft begaan, kan door den rechter slechts worden aangenomen, indien hij daarvan uit het onderzoek op de terechtzitting door den inhoud van wettige bewijsmiddelen de overtuiging heeft bekomen') of tot de civielrechtelijke bewijsregels van art. 149-155 Rv waarbij vooral partijen verantwoordelijk zijn voor het stellen en weer spreken van de feiten? Er ligt een spanningsveld tussen epistemologisch verantwoorde causale verklaringen en argumentatie door de rechter, want de rechter moet steeds beslissen; dit is daardoor ook de geboorteplaats van gerechtelijke dwalingen.

17. Een vergelijkbare redenering werd ook toegepast door de medisch deskundige betrokken bij de zaak Sally Clark; zie de boekbespreking.

18. Zie ook R. Meester, 'Statistiek en kansrekening in het strafrecht', AA 2007, 9, p. 675 e.v.

19. Zie ook R.W.M. Giard, 'Waarheidsvinding. Hoe de juridische werkelijkheid het methodologische ideaal kan naderen', *Trema* (36) 2013-3, p. 89-96 en G. Gigerenzer, *Calculated risks. How to know when numbers deceive you*, New York: Simon Schuster 2002, p. 151 e.v.

20. Amsterdam: Uitgeverij Business Contact 2011.

21. Zie J.J. Rachlinski et al., 'Probable Cause, Probability, and Hindsight', *Journal of Empirical Legal Studies* 2011, vol. 8 (S1), p. 72-98 en ook diens artikel 'Judicial Psychology', *Rechtstreeks* 2012-2, p. 15 e.v.

22. Met het theorema van Bayes: maar dan moeten betrouwbaarheidskenmerken van de (forensische) test en de à priori kans op daderschap met voldoende zekerheid bekend zijn.

23. S. Sloman, *Causal models. How people think about the world and its alternatives*, Oxford: Oxford University Press 2009, p. 24 en L.A. Paul & N. Hall, *Causation. A user's guide*, Oxford: Oxford University Press 2013, hoofdstuk 1.

### Conclusie

'Essentially all models are wrong, but some are useful', aldus statisticus George Box.<sup>24</sup> Hoe bruikbaar zijn statistische modellen in de rechtszaal? Kan daarmee ondubbelzinnig de schuld van gedaagde of verdachte worden vastgesteld door het causale verband getalsmatig te bewijzen? Dat is het concrete doel, maar is het numerieke middel wel geschikt daarvoor? Mijn antwoord daarop is, zoals hierboven beargumenteerd, eenvoudigweg: nee, want dat ligt gewoon buiten het mandaat van de analytische observationele statistiek. Bovendien dienen we onze kaarten niet alleen op de statistiek te zetten. We dienen oog te hebben voor het totaal van de informatie en niet in de valkuil van tunnelvisie te belanden.

Om de waarheid met waarschijnlijkheidsrekening zo dicht als mogelijk te benaderen, dient de modellering daarvan precies te zijn en de data die in dit rekenmodel gestopt moeten worden onberispelijk. Maar hoe modelleer je wiskundig bijvoorbeeld het grillige sterftepatroon in een ziekenhuis en hoe relateer je dat aan de aanwezigheid van één persoon als oorzaak van overlijden? En zelfs als dat volmaakt is uit te voeren, is dat dan, indachtig de mantra associatie  $\neq$  causaliteit, een bewijs van causaliteit? Hier is deugdelijk forensisch bewijs nodig. De in dit artikel genoemde voorbeelden van dwalingen waren *niet* het gevolg van de ongecijferdheid van de rechter. Dat neemt niet weg dat er in de academische en postacademische scholing van juristen meer aandacht voor de mogelijkheden en ook de beperkingen van deze kwantitatieve benaderingswijzen zou mogen zijn. Illustratief hiervoor is dit citaat van de statisticus die genoemde berekening van de kans van 1 op 342 000 000 bekritiseerde:

"Toen ik tijdens de zitting in hoger beroep in de zaak Lucia de B. de rechter duidelijk trachtte te maken hoe statistiek werkt, en ik vertelde over het maken van berekeningen onder bepaalde hypothesen, werd ik onderbroken: "Mijnheer Meester, hier in de rechtszaal houden we ons niet bezig met hypothetische situaties, maar met de realiteit".<sup>25</sup>

Er is behoefte aan meer descriptieve en analytische statistiek ten behoeve van het proces van waarheidsvinding. Een statistische speurhond kan helpen bij het zoeken naar relevante informatie en naar een mogelijke dader, maar helpt *niet* bij het aantonen van een causaal verband, want daarbij staat juist het causale mechanisme centraal. Immers, de juiste benadering bij Lucia de B., bij de intensivecarezuster G. en bij Sally Clark was beantwoording van de vraag *waarom* al dat overlijden had plaatsgevonden, niet hoe (on)gewoon dat was.

24. G.E.P. Box & N.R. Draper, *Empirical model-building and response surfaces*, Oxford, England: John Wiley & Sons 1987, p. 424.

25. Zie R. Meester, noot 18.