

Kansoordelen door deskundigen: over ‘logisch’ rapporteren en wat daarbij mis kan gaan

Henry Prakken*

Kansoordelen door deskundigen worden door rechters vaak verkeerd begrepen. Dit artikel bespreekt hoe dat komt (rechters hebben vaak geen kennis of begrip van de Bayesiaanse kansrekening, terwijl deskundigen daar vaak onvoldoende rekening mee houden) en wat er aan is te doen (deskundigen moeten anders rapporteren, rechters moeten anders motiveren).

1 Inleiding

Op 13 november 2014 deed de Rechtbank Noord-Holland uitspraak in een moordzaak.¹ Een van de kwesties was of op de plaats delict gevonden autobandensporen gemaakt waren door de auto van de verdachte. Vastgesteld was dat het bandenprofiel van de auto van de verdachte overeenkwam met op de plaats delict aangetroffen bandensporen. Volgens het vonnis rapporteerde een bandensporendeskundige als volgt, in een volgens Broeders en Sjerps en Berger² ‘logisch correct’ formaat:

De kans dat een willekeurige Nederlandse auto (...) bandprofielen heeft die overeenkomen met de waarnemingen aan de sporen is naar schatting (...) 1 op 5000.

De rechtbank interpreteerde dit als volgt:

De kans dat de sporen door een willekeurige andere auto zijn veroorzaakt wordt (...) geschat op 1 op de 5000 ...

Waarna de rechtbank concludeerde ‘hetgeen een hoge mate van waarschijnlijkheid oplevert dat de auto van verdachte de bandensporen heeft gemaakt’.

* Prof.dr.mr. H. Prakken is hoogleraar rechtsinformatica en juridische argumentatie aan de Faculteit der Rechtsgeleerdheid, Rijksuniversiteit Groningen en universitair docent kunstmatige intelligentie aan de Faculteit Bètawetenschappen, Departement Informatica, Universiteit Utrecht. Dit artikel is mede mogelijk gemaakt door steun van het Isaac Newton Institute for Mathematical Sciences, Cambridge, Verenigd Koninkrijk, met name financiële steun en gastvrijheid in het kader van het programma Probability and Statistics in Forensic Science (gesteund door EPSRC grant no EP/K032208/1).

¹Rb Noord-Holland 13 november 2014, (ECLI:NL:RBNHO:2014:10689). Mijn bespreking van dit voorbeeld is ontleend aan H. Prakken & R. Meester, ‘Bayesiaanse analyses van complexe strafzaken door deskundigen. Betrouwbaar en zo ja: nuttig?’, *Expertise en Recht* 2017, afl. 5, p. 185-197.

² A.P.A. Broeders, “‘Begrijpt de rechter wat ik bedoel?’ Over het gebruik van “logisch correcte” conclusies in het forensisch identificatieonderzoek”, *Ars Aequi* 2010 afl. 7/8 (Bijzonder nummer *Bewijs*), p. 494-498. M. Sjerps & C.E.H. Berger, *Het Bayesiaanse model biedt een helder zicht op een complexe werkelijkheid* (rapport), Den Haag: Nederlands Forensisch Instituut 2011.

Op een cursus over kansrekening bij strafrechtelijk bewijzen legt een hoogleraar kansrekening uit dat de rechtbank zich schuldig maakt aan een drogreden en dat de conclusie van de rechtbank niet volgt uit de door de deskundige gerapporteerde kans.

Verderop zal ik uitleggen dat de hoogleraar gelijk heeft. Wat is hier dan misgegaan? Begrijpt de rechter in deze zaak de logica van de kansrekening niet of is het format waarin de deskundige rapporteerde toch niet zo logisch als wel beweerd wordt? Deze vragen komen in dit artikel aan de orde. Hierbij zal ik onder meer enkele kanttekeningen plaatsen bij een recente discussie tussen Van Koppen en Sjerps *et al.*³ Eerst zal ik uitleggen waarom de hoogleraar gelijk heeft met haar bewering dat de rechtbank zich in de bovenstaande zaak schuldig maakt aan een drogreden (paragraaf 2). Deze paragraaf geeft daarmee een basale inleiding in de zogenaamde Bayesiaanse kansrekening. Dan zal ik in paragraaf 3 de vraag bespreken: als de rechter inderdaad fout zit, hoe doen andere rechters het dan? Ik zal een klein en voorlopig onderzoekje presenteren waaruit blijkt dat het daarmee niet best gesteld is (wat eerder onderzoek van anderen bevestigt). In paragraaf 4 zal ik schetsen hoe de rechter het wel goed kan doen en in paragraaf 5 bespreek ik waarom een ‘logische’ manier van rapporteren van kansen zo vaak verkeerd geïnterpreteerd kan worden. In paragraaf 6 sluit ik af met enkele aanbevelingen voor deskundigen en rechters.

De kwesties die in mijn artikel aan de orde komen betreffen toepassingen van de Bayesiaanse kansrekening. In recente rechtszaken is deze toepassing controversieel gebleken. Zo hebben de gerechtshoven in Den Haag en ‘s-Hertogenbosch rapportages van de klimaatfysicus Frans Alkemade terzijde gesteld en hebben de Rechtbank Zeeland-West-Brabant en het Gerechtshof ‘s-Hertogenbosch datzelfde gedaan met rapportages van de wetenschapsfilosoof Ton Derksen.⁴ Deze zaken betreffen echter toepassingen van de Bayesiaanse kansrekening als geheel analysemodel van complexe strafzaken en daarvoor is ook naar mijn mening de Bayesiaanse kansrekening nog lang niet geschikt.⁵ In het huidige artikel heb ik het over een veel beperkter gebruik van de Bayesiaanse kansrekening door deskundigen in de rechtszaal, namelijk voor de analyse van individuele bevindingen, en wel meestal op zogenaamd bronniveau (maar soms ook op zogenaamd activiteitsniveau, bijvoorbeeld bij de vraag of bepaalde verwondingen veroorzaakt zijn door geweld of door een ongeluk). Hierbij ligt de vraag voor van welke bron een bepaald spoor (bijvoorbeeld DNA, bandensporen of vingerafdrukken) afkomstig is. Deskundigen beperken zich hier tot het vaststellen van de bewijskracht, ook wel *likelihood ratio* genoemd, van de vastgestelde overeenkomst tussen de sporen en een mogelijke bron.⁶ Onder wetenschappers is dit gebruik van de regel van Bayes oncontroversieel. Er is wel discussie, maar die spitst zich met name toe op de praktische vraag

³ P.J. van Koppen, ‘De tengelhamer en het schedeldak. Over rapportages door het Nederlands Forensisch Instituut’, *Nederlands Juristenblad* 2017/1121. M. Sjerps, A. Kloosterman & C.E.H. Berger, ‘Over de rapportage van het NFI: een weerwoord’, *Nederlands Juristenblad* 2017/2140. P.J. van Koppen, ‘Hoe onlogisch werken toch logisch kan worden gevonden’, *Nederlands Juristenblad* 2017/2141.

⁴ Zie respectievelijk Gerechtshof Den Haag 14 oktober 2015, ECLI:NL:GHDHA:2015:2860, Gerechtshof Den Bosch 22 november 2016, ECLI:NL:GHSHE:2016:5165, Rechtbank Zeeland-West-Brabant 23 mei 2016, ECLI:NL:RBZWB:2016:3060 en Gerechtshof Den Bosch 5 februari 2018, ECLI:NL:GHSHE:2018:421.

⁵ Zie Prakken & Meester 2017. Mijn mening wordt gedeeld door NFI-medewerkers; zie H. van den Heuvel & J.A. de Koeijer, naschrift bij J.A. Coster van Voorhout, ‘het Interdisciplinair Forensisch Onderzoek: meer dan de som der delen’, *Trema* 2017, afl. 8.

⁶ Zie voor dit gebruik van de Bayesiaanse kansrekening onder meer Broeders (2010) en C.E.H. Berger, ‘Criminalistiek is terugredeneren’, *Nederlands Juristenblad* 2010/646.

of juristen deze *likelihood ratios* wel op de juiste manier (kunnen) interpreteren.⁷ Ze worden namelijk pas relevant voor juridisch bewijzen als ze ingebed worden in een Bayesiaanse denkwijze, en de meeste juristen zijn met die denkwijze niet vertrouwd. Mijn artikel gaat over dit laatste probleem.

2 Basale inleiding in de Bayesiaanse kansrekening

Laten we het bandensporenvoorbeeld nader bekijken.⁸ De hoogleraar kansrekening kan als volgt uitleggen dat de rechter zich schuldig maakt aan een drogreden. Om te beginnen: de door de deskundige gerapporteerde kans wordt wel de *random-match probability* genoemd. Dat wil zeggen: de kans dat gevonden sporen overeenkomen met een potentiële bron van de sporen terwijl die potentiële bron toch niet de echte bron is. Met ons bandensporenvoorbeeld kan uitgelegd worden dat er naast de *random-match probability* nog meer informatie nodig is om de kans dat de gevonden sporen door de auto van de verdachte gemaakt zijn te bepalen. Stel dat er in deze zaak initieel niet meer bekend is dan dat het om een in Nederland rondrijdende auto gaat en stel dat er in Nederland 5 miljoen auto's rondrijden (inclusief auto's van bezoekers uit het buitenland). Als we alleen van deze informatie uitgaan,⁹ zijn er dus 5 miljoen potentiële bronnen van de gevonden bandensporen en zijn ze allemaal even waarschijnlijk als de bron van de sporen. Met een *random-match probability* van 1 op de 5000 zal naar verwachting van 1000 van die 5 miljoen auto's het bandenprofiel overeenkomen met dat van de sporen op de plaats delict, terwijl maar één van die auto's de bron is van de sporen, namelijk de auto van de dader(s). De auto van de verdachte kan elk van die 1000 auto's zijn, dus de kans dat de auto van de verdachte de sporen gemaakt heeft, gegeven de overeenkomst tussen de sporen, is 1 op de 1000. Dit rechtvaardigt niet de conclusie van de rechtbank dat de *random-match probability* van 1 op de 5000 'een hoge mate van waarschijnlijkheid oplevert dat de auto van verdachte de bandensporen heeft gemaakt'. De fout van de rechtbank wordt wel de *prosecutor's fallacy* genoemd of ook wel de drogreden van de omkering van de voorwaardelijke kans. De rechtbank verwacht namelijk de volgende twee voorwaardelijke kansen met elkaar:

De kans dat een Nederlandse auto bandprofielen heeft die overeenkomen met de waarnemingen aan de sporen *gegeven dat* de auto de sporen niet gemaakt heeft.

De kans dat een Nederlandse auto de bandensporen niet gemaakt heeft *gegeven dat* de auto bandprofielen heeft die overeenkomen met de waarnemingen aan de sporen.

⁷ Zie bijvoorbeeld Van Koppen 2017, Sjerps et al. 2017 en J.W. de Keijzer, H. Elffers, R.M. Kok & M.J. Sjerps, *Bijkans begrepen? – feitelijk en vermeend begrip van forensische deskundigenrapportages onder rechters, advocaten en deskundigen*, Den Haag: Boom Juridische Uitgevers 2009.

⁸ De onderstaande uitleg is niet bedoeld om origineel te zijn. Er zijn vele leerzame teksten die statistische drogredenen als deze en andere uitstekend uitleggen. Zie bijvoorbeeld Broeders 2010, Ph. Dawid, *Probability and proof. Online appendix to 'Analysis of Evidence' by T.J. Anderson, D.A. Schum and W.L. Twining*. www.statslab.cam.ac.uk/~apd/, C.E.H. Berger & D.J.C. Aben, 'Bewijs en overtuiging: Een helder zicht op valkuilen', *Expertise en Recht* 2010, afl. 5/6, p. 159-165 en T. Derksen, *De Ware Toedracht. Praktische Wetenschapsfilosofie voor Waarheidszoekers*, Diemen: Uitgeverij Veen Magazines 2010. Mijn uitleg is grotendeels ontleend aan Prakken & Meester 2017.

⁹ Dat hoeft niet, maar als er meer informatie bij betrokken wordt, blijft het verhaal in essentie hetzelfde. Zie hiervoor het einde van deze paragraaf.

Alleen uit de tweede voorwaardelijke kans volgt de conclusie van de rechtbank, maar de deskundige had niet de tweede maar de eerste voorwaardelijke kans gerapporteerd.

De fout van de rechtbank kan in meer algemene termen uitgelegd worden met de zogenaamde *regel van Bayes*, een wiskundige formule die uit de axioma's van de kansrekening af te leiden is. Kort gezegd geeft deze formule een methode om de verhouding van de kansen op de juistheid van twee hypothesen gegeven het beschikbare bewijs (de *a-posteriori kansverhouding*) te berekenen uit (1) de verhouding van de kansen op de juistheid van de hypothesen voordat het bewijs bekend is (de *a-priori kansverhouding*) en (2) de verhouding van de voorwaardelijke kansen op het bewijs gegeven de eerste respectievelijk de tweede hypothese (de *likelihood ratio* of *bewijskracht*). Schematisch:

$$\frac{\text{De kans dat hypothese 1 waar is gegeven de bevinding}}{\text{De kans dat hypothese 2 waar is gegeven de bevinding}}$$

=

$$\frac{\text{De kans op de bevinding gegeven dat hypothese 1 waar is}}{\text{De kans op de bevinding gegeven dat hypothese 2 waar is}}$$

x

$$\frac{\text{De kans dat hypothese 1 waar is}}{\text{De kans dat de hypothese 2 waar is}}$$

Oftewel:

$$\text{a-posteriori kansverhouding} = \text{bewijskracht} \times \text{a-priori kansverhouding}$$

In ons bandensporenvoorbeeld wordt dit:

$$\frac{\text{De kans dat de auto van de verdachte de bron is gegeven de overeenkomst}}{\text{De kans dat een andere auto de bron is gegeven de overeenkomst}}$$

=

$$\frac{\text{De kans op de overeenkomst gegeven dat de auto van de verdachte de bron is}}{\text{De kans op de overeenkomst gegeven dat een andere auto de bron is}}$$

x

$$\frac{\text{De kans dat de auto van de verdachte de bron is}}{\text{De kans dat een andere auto de bron is}}$$

In ons voorbeeld is de bewijskracht van de overeenkomst tussen de bandensporen 5000, want de kans op de overeenkomst gegeven dat de auto van de verdachte *niet* de bron is van de bandensporen is 1 op de 5000 (de random-match probability) terwijl de kans op de overeenkomst gegeven dat de auto van de verdachte *wel* de bron is van de bandensporen natuurlijk 1, oftewel 100 procent is. Verder is de a-priorikansverhouding tussen de hypothesen 1 op de 5 miljoen, want er zijn (als er niets meer bekend is) 5 miljoen potentiële bronnen van

de bandensporen. Dus is in ons voorbeeld de a-posteriorikansverhouding dus 5000×1 op de 5 miljoen = 1 op de 1000.¹⁰ Met andere woorden: gegeven de overeenkomst tussen de sporen is het 1000 maal zo waarschijnlijk dat een andere auto de sporen gemaakt heeft dan dat de auto van de verdachte ze gemaakt heeft. We zien nu op een andere manier welke fout de rechtbank gemaakt heeft. De rechtbank heeft de a-priorikansverhouding tussen de twee hypothesen buiten beschouwing gelaten.

Het is leerzaam te zien dat zelfs bij een zeer hoge bewijskracht de a posteriorikansverhouding niet hoog hoeft te zijn. Als bijvoorbeeld de *random-match probability* in ons voorbeeld niet 1 op de 5000 is maar 1 op de 5 miljoen, dan zijn bij 5 miljoen auto's die de bandensporen gemaakt kunnen hebben beide hypothesen even waarschijnlijk (1 op de 5 miljoen maal 5 miljoen is 1), dus de a posteriorikans dat de auto van de verdachte de bandensporen gemaakt heeft is dan slechts 50 procent.

Bij de vraag of iets of iemand de bron van gevonden sporen is, is de bewijskracht het omgekeerde van de *random-match probability*, want de kans op de sporen gegeven dat de potentiële bron ook de echte bron is, is 100 procent. Dus als de *random-match probability* 1 op de miljoen is, is de bijbehorende *likelihood ratio* (bewijskracht) 1 miljoen. In het algemeen hoeft dat niet zo te zijn. Zo kan een deskundige vinden dat de kans op bepaalde verwondingen 80 procent is als de verwondingen het resultaat van geweld zijn en slechts 10 procent als de verwondingen het resultaat van een ongeluk zijn. De bewijskracht is dan 0,8 gedeeld door 0,1 is 8.

Tot slot: hoe kan de rechter een a-priorikansverhouding bepalen? Dat is verre van triviaal en in ons voorbeeld kan een andere redenering (bijvoorbeeld één die ook kijkt naar de geografische verdeling van auto's over Nederland), heel goed tot een andere a priorikans komen. Maar voor mijn betoog is alleen relevant dat een rechter die logisch correcte conclusies wil trekken uit *random-match probabilities* of *likelihood ratios*, daarbij altijd een oordeel moet geven over de a priorikans op de onderzochte hypothesen, en dat altijd moet doen op grond van al het andere relevante bewijs. Dat kunnen statistieken zijn, zoals in ons voorbeeld over het aantal auto's in Nederland, maar het kan ook specifiek bewijs zijn, zoals verklaringen van getuigen die een bepaald type auto op of rond de plaats delict gezien hebben.

3 Hoe doen rechters het in de praktijk?

Nu we hebben gezien dat de rechtbank in onze zaak inderdaad fout zit, rijst de vraag hoe andere rechters het doen. Hier heb ik een klein en voorlopig onderzoek naar gedaan. Op rechtspraak.nl heb ik over de jaren 2014-2017 (niet uitputtend) gezocht naar uitspraken met deskundigenrapportages in termen van *likelihood ratios*. Ik heb dit gedaan met zoektermen als 'kans', 'waarschijnlijker', 'NFI / Nederlands Forensisch Instituut'. Daarbij heb ik 30 relevante zaken gevonden. Soms geeft de deskundige daarin een expliciete *random-match probability*, zoals in ons voorbeeld uit paragraaf 1 en 2. In andere zaken formuleert de deskundige de *likelihood ratio* als vergelijkend oordeel, en dat meestal zonder getallen, als in 'de bevindingen zijn (zeer) (veel) waarschijnlijker als hypothese 1 waar is dan als hypothese 2 waar is'. Dat bij de laatste formulering geen getallen worden gebruikt, maakt niets uit voor de toepasbaarheid van de regel van Bayes, zoals ik in paragraaf 4 zal uitleggen.

¹⁰ Om de berekening iets begrijpelijker maken negeer ik hier een klein benaderingsfoutje: strikt genomen is de a-priorikansverhouding $1/4.999.999$.

Van de door mij onderzochte 30 zaken is naar mijn mening in 22 zaken door de rechter de *prosecutor fallacy* begaan.¹¹ Soms draait de rechter hierbij een voorwaardelijke kans(verhouding) expliciet om, zoals in de volgende uitspraak van de Rechtbank Limburg:

‘... Er is een match gevonden met het DNA-profiel van verdachte waarbij de matchkans kleiner is dan 1 op 1 miljard. Dat betekent dat de kans dat het celmateriaal op die uiteindes van een willekeurig gekozen persoon is, kleiner is dan 1 op 1 miljard.’¹²

In andere uitspraken wordt een *random-match probability* of *likelihood ratio* niet expliciet omgedraaid maar wordt zonder nadere motivering een conclusie getrokken die uit zo’n omdraaiing volgt. Dit heb ik ook als voorbeelden van de *prosecutor fallacy* beschouwd. Een voorbeeld is:

‘DNA-onderzoek heeft bevestigd dat verdachte zeer waarschijnlijk (de kans op de verkregen DNA-profielen is meer dan één miljard keer groter is wanneer verdachte de biologische moeder van deze baby is dan wanneer een andere vrouw, niet verwant aan verdachte, de biologische moeder van deze baby is) de biologische moeder is van baby [baby 1]’¹³

Een ander voorbeeld van dit type is een uitspraak van de Rechtbank Rotterdam,¹⁴ waarin een NFI-deskundige bij een DNA-match een *random-match probability* rapporteerde van kleiner dan één op één miljard en daarbij aangaf dat dit de hoogste vorm van waarschijnlijkheid is die het NFI in dergelijke onderzoeken afgeeft. De rechtbank concludeerde hieruit:

‘De rechtbank acht de uitkomst van het DNA-onderzoek betrouwbaar, op grond waarvan met aan zekerheid grenzende waarschijnlijkheid kan worden vastgesteld dat het verdachte is geweest die op de badkamermuur een bloedspoor heeft achtergelaten.’

Ook dit is de *prosecutor fallacy*, want zoals we hierboven gezien hebben, volgt deze conclusie slechts als bovendien de a-priorikansverhouding niet te laag is. Als deze bijvoorbeeld ook minder dan één op één miljard is, kan het heel goed zijn dat de kans dat de verdachte de bron van de bloedsporen is minder dan 50 procent is, of in ieder geval beduidend lager dan bijna 100 procent. Maar kan hier niet gezegd worden dat een zo lage a-priorikansverhouding onrealistisch is en de conclusie van de rechtbank toch gerechtvaardigd is? Bijvoorbeeld: zelfs als de gehele Nederlandse bevolking (in gelijke mate) als potentiële bron beschouwd wordt, is, gegeven de match, de kans dat de verdachte de bron is op zijn minst 58 maal zo groot als de kans dat iemand anders de bron is, dus op zijn minst 98 procent. Bovendien kan het heel goed zijn dat op basis van het andere bewijs er niet zoveel potentiële bronnen van de bloedsporen over zijn en dan volgt de conclusie van de rechtbank inderdaad uit de *random-match probability* in combinatie met het overige bewijs. Maar dat is niet het punt: het punt is dat de rechtbank niet naar overig bewijs verwijst en dus geen besef toont dat een *random-*

¹¹ De appendix bevat de ECLI's van de door mij onderzochte zaken, met daarbij aangegeven in welke zaken naar mijn mening sprake is van een *prosecutor fallacy*.

¹² Rb Limburg 25 februari 2015, ECLI:NL:RBLIM:2015:1588.

¹³ Rb Noord-Holland 12 mei 2016, ECLI:NL:RBNHO:2016:3947.

¹⁴ Rb Rotterdam 14 juni 2017, ECLI:NL:RBROT:2017:4573.

match probability altijd gecombineerd moet worden met een a-priorikansverhouding. Het kan bijvoorbeeld ook zijn dat het overige bewijs juist naar andere bronnen van de bloedsporen wijst en daarmee de a priori- en dus ook de a-posteriorikansverhouding niet verhoogt maar juist verlaagt, bijvoorbeeld tot beduidend lager dan 98 procent.

In een aantal andere zaken verwijst de rechter wel naar ander bewijs, maar niet op een systematische manier. Een uitspraak vermeldt dan eerst één of meer *random-match probabilities* of *likelihood ratios*, gevolgd door een opsomming van overig bewijs en dan een algemene conclusie als ‘op grond van het bovenstaande acht de rechtbank bewezen dat ...’, zonder expliciet op de gerapporteerde kansen of kansverhoudingen terug te komen. In deze gevallen is het onduidelijk of de rechters zich er van bewust waren dat de door de deskundigen gerapporteerde kansen gecombineerd moeten worden met een a-priorikansverhouding. Deze zaken heb ik niet geteld als voorbeelden van de *prosecutor fallacy*, maar ze geven wel te denken. Een voorbeeld is te vinden in een vonnis van de Rechtbank Midden-Nederland.¹⁵

In slechts één van de 30 door mij onderzochte uitspraken toont de rechter besef van het feit dat een *likelihood ratio* gecombineerd moet worden een a-priorikansverhouding. Een fragment hieruit:

‘Voorts geldt naar het oordeel van de rechtbank dat, hoewel de bewijswaarde van de door het TMFI berekende frequentie van 1:5000 zonder context bezien niet zo groot is, de kans dat verdachte de donor is van het DNA-materiaal op de huls aanzienlijk groter is indien daarbij al het andere bewijs wordt betrokken, met name de aanwezigheid van het andere DNA-spoor op de plaats delict, met een frequentie van minder dan 1 op 1 miljard.’¹⁶

Samengevat vond ik in ruim tweederde van de onderzochte zaken voorkomens van de *prosecutor fallacy* en vond ik slechts in één zaak expliciet besef dat *random-match probabilities* en *likelihood ratios* met ander bewijs gecombineerd moeten worden voor conclusies over de bewijsbaarheid van de onderzochte hypothesen getrokken kunnen worden. Hierbij geldt het voorbehoud dat dit slechts een klein en voorlopig onderzoek betreft, zonder claim van formele validiteit. Maar mijn bevindingen zijn wel in lijn met eerder, meer systematisch onderzoek. Zo vonden De Keijzer et al.¹⁷ in een experiment met rechters, advocaten en NFI-deskundigen dat 80 procent een *likelihood ratio* omdraaide tot een a posteriorikansverhouding. Ook veel algemeen psychologisch onderzoek laat zien dat mensen kansinformatie vaak onjuist interpreteren.¹⁸

Tot slot: hoe erg zijn deze bevindingen in praktische zin? Bevatten de door mij onderzochte zaken gerechtelijke dwalingen? Dat kan niet geconcludeerd worden, want het kan in al deze zaken heel goed zijn dat het overige bewijs de a-priorikansverhouding voldoende hoog maakt

¹⁵ Rb. Midden-Nederland 15 december 2015, ECLI:NL:RBMNE:2015:8972.

¹⁶ Rb. Oost-Brabant 21 november 2014, ECLI:NL:RBOBR:2014:7096.

¹⁷ J.W. de Keijzer, H. Elffers, R.M. Kok & M.J. Sjerps, *Bijkans begrepen? – feitelijk en vermeend begrip van forensische deskundigenrapportages onder rechters, advocaten en deskundigen*, Den Haag: Boom Juridische Uitgevers 2009.

¹⁸ Zie bijvoorbeeld D. Kahneman, *Thinking, Fast and Slow*, Londen: Penguin 2011.

om de belastende conclusies te trekken. Vervolgonderzoek moet uitwijzen hoe vaak de gemaakte fouten of bestaande onduidelijkheden relevant zijn voor de uitkomst van een zaak. Maar verontrustend is dat vrijwel geen enkele uitspraak expliciet begrip toont van hoe door deskundigen gerapporteerde kansen gebruikt kunnen worden om valide conclusies te trekken. In combinatie met het zojuist genoemde andere empirische onderzoek geeft dit weinig reden om te geloven dat dit begrip wel bij rechters bestaat en alleen maar impliciet gelaten wordt.

4 Hoe kunnen rechters het goed doen?

Dan nu hoe rechters het goed kunnen doen. Moet dit volledig in termen van de kansrekening? Dat hoeft niet per se. Er is al veel gewonnen als de rechter expliciet beargumenteert dat de a-priorikansverhouding de getrokken conclusies uit de kansoordelen van de deskundigen toelaat, en dat kan ook op andere manieren, bijvoorbeeld met argumentatief of scenariogebaseerd denken.¹⁹ Hierbij hoeft de rechter niet per se getallen te gebruiken. Een motivering als:

‘de match is zeer veel waarschijnlijker bij hypothese 1 dan bij hypothese 2 en gegeven de overige bevindingen is de a-priorikansverhouding niet erg laag (en wel hierom), dus gegeven alle bevindingen is hypothese 1 veel waarschijnlijker dan hypothese 2’

is consistent met de kansrekening omdat getallen te bedenken zijn die passen bij de relatieve oordelen en de conclusie rechtvaardigen. Of die getallen zelf te rechtvaardigen zijn, is een andere vraag, die buiten het bestek van dit artikel valt. Belangrijk is dat uit een dergelijke motivering blijkt dat de rechter beseft welke kansen bepaald moeten worden en hoe die gecombineerd moeten worden. Dat is voldoende voor een logisch gezien deugdelijke motivering.

5 Hoe kan ‘logisch’ rapporteren zo vaak verkeerd geïnterpreteerd worden?

De vraag rijst: als het format waarin de deskundigen rapporteren zo logisch is, hoe kan het dan zo vaak fout gaan? Het korte antwoord is: omdat te veel deskundigen zich beperken tot dit format en geen of onvoldoende verdere uitleg geven, en zo onterecht impliciet aannemen dat rechters wel begrijpen hoe ze *likelihood ratios* moeten interpreteren en combineren met ander bewijs.²⁰ Maar er is meer over te zeggen.

Ten eerste, forensisch deskundigen spreken nogal eens van een logisch correct raamwerk of logisch correcte formaten/conclusies, maar wat betekent ‘logisch correct’ in dit verband? Belangrijk is dat niet conclusies of formuleringen maar *redeneringen* logisch correct of incorrect kunnen zijn. Een redenering bestaat uit één of meer premissen en een conclusie. Ze is logisch correct als iemand die de premissen als waar aanvaardt, ook de conclusie als waar moet aanvaarden, anders heeft die persoon de betekenis van de premissen en conclusie niet

¹⁹ Zie voor een vergelijking tussen de argumentatieve, scenariogebaseerde en Bayesiaanse denkwijzen H. Prakken, ‘Strafrechtelijk bewijzen: met Bayes of met verhalen? Of is er een derde weg?’, *Expertise en Recht* 2014, afl. 1, p. 4-19.

²⁰ Tenminste, dat is de indruk die oprijst uit de door mij geanalyseerde 30 uitspraken. Omdat ik geen rechtstreekse toegang had tot de deskundigenrapporten, moet nader onderzoek uitwijzen in hoeverre dit echt zo is.

goed begrepen. De waarheid van de premissen is hierbij hypothetisch: een geldige redenering kan heel goed feitelijk onware premissen hebben, zoals in ‘Assen is de hoofdstad van Nederland, Assen ligt in Friesland, dus de hoofdstad van Nederland ligt in Friesland’. In de context van de kansrekening betekent dit dat als een deskundige een kansuitspraak doet en onderbouwt met andere kansuitspraken, zijn redenering logisch correct is mits de gedane kansuitspraak volgens de axioma's van de kansrekening wiskundig volgt uit de ter onderbouwing aangedragen kansuitspraken. De frase ‘logisch correcte conclusie’ moet dan gelezen worden als ‘de conclusie volgt logisch (volgens de wetten van de kansrekening) uit de aangedragen onderbouwing’. Iets anders is of de premissen van de kansredenering op de expertise van de deskundige gebaseerd kunnen worden. Dat is de vraag of de premissen waar zijn, en die vraag ligt buiten de logica en dus ook buiten de kansrekening.

Laten we nu de zaak bekijken die in de recente discussie tussen Sjerps et al. en van Koppen aan de orde kwam.²¹ In deze zaak was iemand dood aangetroffen met verwondingen aan de schedel. Bij de verdachte was onder meer een tengelhamer gevonden (‘de’ tengelhamer in de volgende fragmenten). De rechter-commissaris stelde de volgende vraag aan een NFI-deskundige:

‘Zijn de beschadigingen in het gesepareerde schedeldak veroorzaakt door de tengelhamer?’

De deskundige beantwoordde deze vraag als volgt:

‘Naar aanleiding van de vraag: “Zijn de beschadigingen in het gesepareerde schedeldak veroorzaakt door de tengelhamer” zijn de volgende hypothesen gesteld:

Hypothese 1: De aangetroffen beschadigingen in het gesepareerde schedeldak zijn veroorzaakt door *de tengelhamer*;

Hypothese 2: De aangetroffen beschadigingen in het gesepareerde schedeldak zijn veroorzaakt door *een andere klauwhamer*.

Hypothese 3: De aangetroffen beschadigingen in het gesepareerde schedeldak zijn veroorzaakt door *een andere soortgelijke tengel hamer*.

De bevindingen van het onderzoek zijn zeer veel waarschijnlijker wanneer hypothese 1 juist is, als wanneer hypothese 2 juist is.

De bevindingen van het onderzoek zijn waarschijnlijker wanneer hypothese 1 juist is, als wanneer hypothese 3 juist is’.

Als we de vraag van de rechter-commissaris interpreteren als ‘Is het gegeven de beschadigingen in het gesepareerde schedeldak (zeer) waarschijnlijk dat ze veroorzaakt zijn door de tengelhamer?’ dan zien we dat dit een a-posteriorikans is terwijl de deskundige deze vraag beantwoordt door het geven van twee *likelihood ratios*. Hierboven hebben we gezien dat een *likelihood ratio* op zich nog geen conclusie toelaat over een a-posteriorikans of kansverhouding, maar de deskundige heeft de rechter-commissaris daar kennelijk niet voor gewaarschuwd. Het lijkt erop dat hij de *likelihood ratios* over de schutting van de rechtbank

²¹ Deze zaak is geanalyseerd in het project Gereede Twijfel van de Vrije Universiteit. Zie M.L. van Dobben, F. Verkuijlen, L. Jovanović, M.R. Stokhof, L.L.M. Borst, Y. van Schaik, M.A.H. van Thiel, M. Willemsen, J.J. van der Kemp & P. van Koppen, *De Dansende Dader*. Den Haag: Boom Juridische Uitgevers 2018. De onderstaande fragmenten zijn geciteerd uit dit boek, met een enkele inkorting voor de leesbaarheid.

heeft gegooid en heeft gedacht ‘zie maar wat je er mee doet’. Naar mijn mening werkt rapporteren in deze stijl het begaan van de *prosecutor fallacy* sterk in de hand. De logica van de kansrekening impliceert dat een *likelihood ratio* nooit een antwoord kan zijn op een vraag als die van de rechter-commissaris, maar het is bekend dat juristen een *likelihood ratio* vaak wel als zodanig opvatten. Daarom had de deskundige dit antwoord niet zo mogen geven.

Sjerps et al. vinden dat een deskundige geen a-posteriorikansverhoudingen kan geven, omdat de deskundige daarvoor een oordeel moet geven over de a-priorikansverhouding, wat doorgaans buiten diens expertisegebied zal liggen. Dit is op zich een sterk punt²² maar het impliceert niet dat deskundigen kunnen volstaan met het geven van *likelihood ratios* als antwoord op een vraag hoe bepaalde bevindingen veroorzaakt zijn. Eerder zouden ze moeten zeggen: op deze vraag kan ik geen antwoord geven. Een deskundige die zo’n vraag wel wil beantwoorden, ontkomt niet aan een oordeel over *posteriors* en dus ook niet aan een oordeel over *priors*. Zo’n oordeel is (als de deskundige de axioma’s van de kansrekening juist toepast) op geen enkele manier in strijd met de logica van de Bayesiaanse kansrekening. Of een deskundige zo’n oordeel op diens expertise kan baseren, is een andere vraag.

Verder claimen Sjerps et al. dat de bevindingen van de deskundige de hypothese dat de verwondingen zijn veroorzaakt door de tengelhamer van de verdachte “steunen”. Dit is om twee redenen problematisch. Ten eerste is de term ‘steunen’ ambigu. Het kan betekenen: ‘gegeven het bewijs is Hypothese 1 waarschijnlijk waar’. Dit is niet wat Sjerps et al. bedoelen, maar het kan heel goed zijn dat veel juristen het wel zo interpreteren. Een andere interpretatie van ‘steunen’ is dat het bewijs de kans vergroot dat Hypothese 1 waar is (meer precies: de voorwaardelijke kans op Hypothese 1 gegeven het bewijs is groter dan de onvoorwaardelijke (a priori) kans op Hypothese 1). Maar de logica van de kansrekening impliceert dat dit in zijn algemeenheid niet opgaat. Het gaat alleen op als de door de deskundige vergeleken hypothesen elkaar niet alleen uitsluiten maar ook samen alle mogelijkheden afdekken. Als dat niet zo is, dan is het logisch mogelijk dat de bevindingen een niet-beschouwde Hypothese 4 nog sterker ondersteunen dan Hypothese 1 (dat wil zeggen, de bevindingen zijn waarschijnlijker gegeven Hypothese 4 dan gegeven Hypothese 1), zodat de a-posteriorikans op Hypothese 1 lager kan zijn dan de a-priorikans op Hypothese 1 en de bevindingen dus niet belastend maar ontlastend zijn.²³ In casu zijn in het Gerede-Twijfelboek zulke niet door de deskundige beschouwde hypothesen geopperd, namelijk dat de verwondingen toegebracht zijn door de naast het slachtoffer gevonden wandelstok van het slachtoffer, of veroorzaakt zijn door een ongelukkige val van de trap. Dus volgens hun eigen Bayesiaanse logica hadden Sjerps et al. de conclusie dat de bevindingen van de deskundige de hypothese van het OM steunen niet mogen trekken zonder de expliciete aanname dat er geen andere mogelijke oorzaken van het schedelletsel zijn.

²² Ook hier is trouwens discussie over mogelijk. Zo vindt Van Koppen in zijn weerwoord op Sjerps et al. (2017) dat deskundigen alles moeten aandragen dat relevant is voor de vraag die aan hen is voorgelegd, ook als dat niet strikt binnen hun expertisegebied valt. Deze kwestie, hoe belangrijk ook, valt buiten het bestek van dit artikel, omdat het niet over de logische interpretatie van kansen gaat. Wel wil ik opmerken dat Van Koppens visie niet de activiteiten van de klimaatfysicus Frans Alkemade (zie paragraaf 1) rechtvaardigt, die zonder enige inhoudelijke expertise allerlei kansen in complexe strafzaken schat. Het vertrekpunt dient altijd inhoudelijke deskundigheid ten aanzien van een bepaald aspect van de zaak te zijn. Zie hiervoor ook Prakken & Meester 2017.

²³ Zie hiervoor N.E. Fenton, D. Berger, D. Lagnado, M. Neil & A. Hsu, ‘When “neutral” evidence still has probative value (with implications from the Barry George Case)’, *Science and Justice* (54) 2014, afl. 4, p. 274-287.

6 Conclusies en aanbevelingen

Dit artikel bevestigt aanwijzingen uit eerder onderzoek dat juristen de door forensisch deskundigen gerapporteerde *likelihood ratios* en *random-match probabilities* vaak verkeerd interpreteren. Dit komt door een combinatie van twee factoren: veel juristen zijn onbekend met of niet vaardig in de Bayesiaanse denkwijze die nodig is om dergelijke kansoordelen correct te interpreteren, en veel deskundigen houden daar bij de presentatie van hun kansoordelen te weinig rekening mee.

Deze conclusies leiden tot de volgende aanbevelingen.

Ten eerste moeten forensisch deskundigen in hun rapportages volstrekt duidelijk maken dat hun *likelihood ratios* en *random-match probabilities* zonder een oordeel over de a-priorikansverhouding niet voor conclusies omtrent bewijsbaarheid van de hypothesen gebruikt kunnen worden. En ze mogen vragen hoe bepaalde bevindingen veroorzaakt zijn niet beantwoorden met slechts *likelihood ratios*, zonder een nadere toelichting dat die geen antwoord op de vraag zijn.

Volgens Marjan Sjerps en Charles Berger (persoonlijke communicatie) moeten NFI-rapporten tegenwoordig een voetnoot bevatten met daarin onder meer de volgende tekst: “De conclusie verwoordt de bewijskracht van de resultaten ten aanzien van de hypothesen. De conclusie geeft niet de kans weer dat een bepaalde hypothese waar is. Die kans hangt namelijk ook af van overig bewijs en informatie buiten het forensische expertiseterrein en valt daardoor buiten de reikwijdte van dit rapport. Zie voor meer informatie de vakbijlage 'De reeks waarschijnlijkheidstermen van het NFI'”. De vraag rijst of deze voetnoot, die generiek is en niet naar specifieke aspecten van de zaak verwijst, voldoende is, want in 26 van de door mij onderzochte uitspraken werden de kansoordelen gegeven door NFI-deskundigen en in 19 daarvan werd de *prosecutor fallacy* gemaakt. Overigens wordt uit geen van deze 26 uitspraken duidelijk of de NFI-deskundigen de voetnoot in hun rapporten opgenomen hebben.

Ten tweede moeten rechters in hun motiveringen, als ze conclusies trekken uit door deskundigen gegeven kansoordelen, expliciet de vraag beantwoorden of de a-priorikansverhouding met betrekking tot de onderzochte hypothesen toestaat om in combinatie met de kansoordelen van de deskundigen conclusies te trekken omtrent de bewijsbaarheid van hypothesen. Overigens hoeft de rechter hierbij, zoals gezegd in paragraaf 4, niet per se Bayesiaans te denken; het kan vaak ook met argumentatief of scenariogebaseerd denken. En de rechter hoeft ook niet per se getallen te geven: kwalitatieve oordelen als ‘(zeer) veel waarschijnlijker’ passen ook in de logica van de kansrekening.

Tot slot zou een degelijke inleiding in de kansrekening een verplicht onderdeel van elke juridische opleiding moeten worden, temeer daar statistisch bewijs niet alleen in het strafrecht maar ook in civiele zaken steeds vaker een rol speelt (bijvoorbeeld DNA-bewijs bij vaderschapskwesities).²⁴

Deze drie aanbevelingen zijn naar zijn mening noodzakelijke voorwaarden voor een rationele interpretatie van door deskundigen gegeven kansoordelen door rechters. Of ze ook voldoende

²⁴ In Groningen geef ik sinds 2015 zo'n inleiding in het kader van een vak Juridische Argumentatie, samen met inleidingen in de argumentatietheorie en de formele logica.

zijn is nog maar de vraag, want uit allerlei onderzoek blijkt dat zelfs in de kansrekening getrainde mensen nog vaak moeite hebben om kansinformatie juist te interpreteren.²⁵

Appendix: Lijst van onderzochte zaken

Zaken met naar mijn mening voorkomens van de *prosecutor fallacy*:

ECLI:NL:RBZWB:2014:4249
ECLI:NL:RBROT:2014:9661
ECLI:NL:RBAMS:2014:3655
ECLI:NL:RBLIM:2015:1588
ECLI:NL:RBDHA:2016:4166
ECLI:NL:RBROT:2017:4573
ECLI:NL:RBROT:2017:9969
ECLI:NL:RBNNE:2017:4349
ECLI:NL:RBMNE:2015:1422
ECLI:NL:RBOBR:2016:6438
ECLI:NL:RBOVE:2014:3905
ECLI:NL:RBLIM:2014:4453
ECLI:NL:RBDHA:2017:4664
ECLI:NL:RBNHO:2016:3947
ECLI:NL:RBZWB:2015:6889
ECLI:NL:RBLIM:2015:4772
ECLI:NL:RBNHO:2016:220
ECLI:NL:RBOBR:2016:3594
ECLI:NL:RBLIM:2017:12401
ECLI:NL:RBDHA:2017:13834
ECLI:NL:RBGEL:2015:6785
ECLI:NL:RBOVE:2015:2815

Overige onderzochte zaken

ECLI:NL:RBDHA:2015:6688
ECLI:NL:RBNNE:2016:3096
ECLI:NL:GHSHE:2017:3109
ECLI:NL:RBNNE:2016:5823
ECLI:NL:RBZWB:2016:3137
ECLI:NL:RBNNE:2014:1086
ECLI:NL:RBMNE:2015:8972
ECLI:NL:RBOBR:2014:7096
ECLI:NL:RBGEL:2017:1023

²⁵ Aldus bijvoorbeeld S. Lindsey, R. Hertwig & G. Gigerenzer, 'Communicating statistical DNA evidence' *Jurimetrics* (43) 2003, p. 147-163.