

Bayesiaanse analyses van complexe strafzaken door deskundigen. Betrouwbaar en zo ja: nuttig?¹

Henry Prakken

Faculteit der Rechtsgeleerdheid, Rijksuniversiteit Groningen &
Faculteit Bètawetenschappen, Departement Informatica, Universiteit Utrecht

Ronald Meester

Afdeling Wiskunde
Vrije Universiteit Amsterdam

20 oktober 2017

Samenvatting

Forensische wetenschappers propageren tegenwoordig het gebruik van de Bayesiaanse kansrekening. Tot voor kort beperkte het gebruik hiervan in de rechtszaal zich tot de analyse van de bewijskracht van individuele bevindingen, maar recent is in een aantal strafzaken een Bayesiaanse analyse van de volledige zaak als deskundigenrapport ingebracht. Dat roept de vraag op of de Bayesiaanse kansrekening op een wetenschappelijk betrouwbare en praktisch werkbare wijze toegepast kan worden ter beantwoording van de schuldvraag in strafzaken.

In dit artikel formuleren we eerst een aantal noodzakelijke voorwaarden voor betrouwbaar gebruik van de Bayesiaanse kansrekening in strafgedingen. Vervolgens concluderen we dat hun vervulling in de rechtspraktijk zeer problematisch is en dat ze bovendien samen nog onvoldoende zijn voor een betrouwbare Bayesiaanse analyse van complexe strafzaken, omdat de wetenschap daarvoor nog onvoldoende specifieke methoden en richtlijnen heeft ontwikkeld.

1 Inleiding

De laatste tijd is er zowel in de wetenschap als de rechtspraktijk discussie over het gebruik van kansrekening, en over de zogenaamde regel van Bayes in het bijzonder, bij strafrechtelijk bewijzen. Naar aanleiding van enkele geruchtmakende rechterlijke dwalingen zoals de Schiedammer Parkmoordzaak en de zaak Lucia de Berk en (in Engeland) de zaak Sally Clark is de belangstelling in de rechtspraktijk voor wat wetenschappers te zeggen hebben over rationeel bewijzen toegenomen. Met name forensische wetenschappers propageren het gebruik van de regel van Bayes (Berger & Aben 2010a, 2010b, 2010c; Sjerps 2011; Sjerps & Berger 2011) en het Nederlands Forensisch Instituut (NFI) heeft sinds 2005 als beleid dat deskundigen van het NFI hun bevindingen rapporteren in termen van zogenaamde likelihood ratio's.

Er is overigens geen consensus over de toepasbaarheid van de regel van Bayes in strafzaken (Meester 2015). Anderen bestuderen benaderingen in termen van argumentatie (Bex et al. 2003; Prakken 2004; Anderson et al. 2005), van verhaals- of

¹ 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). Wij danken de redactie van *Expertise en Recht* voor het zeer nuttige commentaar op een eerdere versie van dit artikel.

scenarioconstructie (Crombag et al. 1994; Van Koppen 2011) of van combinaties daarvan (Bex et al. 2010; Verheij et al. 2016). Zie voor een bespreking van de verhouding tussen de verschillende benaderingen Prakken (2014). Ook onder de aanhangers van de Bayesiaanse benadering is er verschil van mening over de vraag of deze benadering alleen nuttig is op deelpunten (zoals de interpretatie van DNA-bewijs) of dat ze ook als denkmodel voor strafrechtelijk bewijzen gebruikt kan worden (Meester 2015).

Tot voor kort beperkte het gebruik van de regel van Bayes door deskundigen in de rechtszaal zich tot de analyse van individuele bevindingen, en wel op zogenaamd bronniveau, waarbij de vraag voorligt van welke bron een bepaald spoor (bijvoorbeeld DNA, voetsporen of vingerafdrukken) afkomstig is. Dit gebruik beperkt zich tot het vaststellen van de bewijskracht, ook wel likelihood ratio (LR) genoemd, van de bevindingen. Onder wetenschappers is dit gebruik van de regel van Bayes relatief oncontroversieel. Er is wel discussie, maar die spitst zich met name toe op de praktische vraag of juristen (rechters, vertegenwoordigers van het OM, advocaten) deze likelihood ratio's wel op de juiste manier (kunnen) interpreteren. Ze krijgen pas betekenis (en dus waarde voor juridisch bewijzen) als ze ingebed worden in een Bayesiaanse denkwijze, en de meeste juristen zijn met die denkwijze niet vertrouwd. Punt van discussie hierbij is onder meer of dit probleem verholpen kan worden door juristen en rechtenstudenten beter te trainen of op te leiden in de kansrekening of dat het probleem dieper ligt.

Recent is er een aantal pogingen geweest om een Bayesiaanse analyse van een volledige strafzaak als deskundigenrapport in te brengen. In de herzieningszaak *Zes van Breda* (Hof Den Haag 14 oktober 2015, ECLI:NL:GHDHA:2015:2860) bracht het OM een Bayesiaanse analyse van de gehele zaak in van de klimaat- en atmosferisch fysicus dr. Frans Alkemade, toentertijd werkzaam bij het Netherlands Institute for Space Research (SRON) in Utrecht. Aan een van ons (HP) was in deze zaak door het hof gevraagd om het rapport van dr. Alkemade te evalueren en beoordelen. In het hoger beroep van de zaak van de Oosterlandse brandstichtingen (ECLI:NL:GHSHE:2016:5165) bracht het OM opnieuw een Bayesiaanse analyse door Alkemade van de gehele zaak in. Het hof had een van ons (HP) gevraagd om in deze zaak een contraonderzoek te verrichten en daarin de betrouwbaarheid van de Bayesiaanse methode zoals door Alkemade in deze zaak toegepast te beoordelen. Tot slot heeft recent de verdediging in een zaak voor de rechtbank Zeeland-West-Brabant (ECLI:NL:RBZWB:2016:3060) een deskundigenrapport van prof. Ton Derksen, emeritus hoogleraar wetenschapsfilosofie aan de Radbouduniversiteit Nijmegen, ingebracht waarin de Bayesiaanse kansrekening werd gebruikt. Bij deze zaak waren wij niet betrokken.

In Prakken (2014) wordt betoogd dat de Bayesiaanse denkwijze berust op filosofische aannames omtrent de aard van oordelen en bewijzen. Een rechter die niet-Bayesiaans over bewijs denkt, komt dus niet noodzakelijkerwijs in strijd met de wiskunde en logica; ook andere denkwijzen, bijvoorbeeld argumentatief of in termen van scenario's, kunnen rationeel zijn. Aan de andere kant heeft de Bayesiaanse denkwijze als voordeel dat het een gedetailleerde wiskundige beschrijving geeft van een rationeel denkproces. Bovendien zijn de alternatieve benaderingen evenmin universeel aanvaard als enig juiste denkmodel over redeneren ingeval van

onzekerheid. De keuze voor de toepassing van een Bayesiaanse denkwijze op een juridische casus is daarom in sommige gevallen zeker verdedigbaar. In dat licht bezien is het de moeite waard om te onderzoeken in hoeverre ze praktisch bruikbaar is in de rechtszaal. Dat is wat we in dit artikel zullen doen. We zullen ons hierbij met name richten op het gebruik van de Bayesiaanse kansrekening voor het geven van een analyse van een complexe strafzaak in zijn geheel. De vraag die we willen bespreken is tweeledig:

- (1) Kan de regel van Bayes op een wetenschappelijk betrouwbare en nuttige wijze toegepast worden in de rechtszaal ter analyse van complexe strafzaken in hun geheel?
- (2) Is een dergelijk gebruik van de regel van Bayes in de rechtszaal praktisch gezien mogelijk en zo ja, onder welke voorwaarden?

Het is niet ons doel om gedetailleerde kritiek te leveren op de rapporten van Alkemade in deze zaken. Ons doel is veeleer om algemene conclusies te trekken; wel zullen we waar dat zinvol is onze observaties illustreren met aspecten uit de beide zaken.

We merken nog op dat naast de vraag of de regel van Bayes nuttig en betrouwbaar is bij de analyse van een complexe strafzaak als geheel, er ook nog verschillende andere problemen zijn die we hier en daar wel aanstippen maar niet expliciet bespreken.

Hieronder zullen we eerst in paragraaf 2 de regel van Bayes samenvatten voor zover van belang voor dit artikel. Meer uitgebreide inleidingen in een juridische context zijn bijvoorbeeld Berger & Aben (2010a, 2010b, 2010c), Van den Berg (2014) of de paragrafen 3-8 van Alkemade & Stikkelbroeck (2015a). Dan zullen we in paragraaf 3 de voorwaarden schetsen voor de bruikbaarheid van de regel van Bayes in de rechtspraak, waarbij we vooral in zullen gaan op de betrouwbaarheid van de methode. In het licht daarvan bespreken we in paragraaf 4 drie bovengenoemde rechtszaken waarin de Bayesiaanse methode als denkmodel voor de gehele zaak gebruikt werd. In paragraaf 5 geven we dan richtlijnen voor deugdelijk gebruik van de regel van Bayes door deskundigen in complexe strafzaken als methode om uitspraken te doen over de kans op schuld of onschuld van de verdachte. In paragraaf 6 sluiten we af met conclusies, onder meer over de praktische toepasbaarheid van Bayesiaans denken volgens deze richtlijnen.

2 De Bayesiaanse benadering

2.1 Inleiding

De Bayesiaanse benadering ziet juridisch bewijzen als denken over kansen. Een kans moet hierbij epistemisch worden geïnterpreteerd, dat wil zeggen dat de kans de mate van geloof uitdrukt die een persoon in (de juistheid van) een bepaalde hypothese of gebeurtenis heeft. Onder bepaalde aannames van rationaliteit kan men inzien dat deze interpretatie aan de gebruikelijke axioma's van de kansrekening moet voldoen, hoewel er ook redenen zijn om daar vraagtekens bij te zetten (Meester & Kerkvliet 2016). In dit artikel zullen we dat punt echter laten rusten en gaan we ervan uit dat

epistemische kansen aan de wetten van de kansrekening voldoen, en dus ook aan de regel van Bayes.

In de kansrekening wordt onzekerheid over de waarheid van een bewering uitgedrukt door een kans te geven dat de bewering waar is, waarbij het begrip kans zoals gezegd geïnterpreteerd dient te worden als mate van geloof in de juistheid van die bewering. Een kans wordt uitgedrukt in een getal tussen 0 en 1, of met andere woorden: tussen 0% en 100%. Een kans van 1 betekent dat de betrokkene een bewering *zeker waar* vindt, een kans op 0 dat hij de bewering *zeker onwaar* vindt, en elke andere kans drukt een mate van *onzekerheid* uit over de waarheid van de bewering. Een belangrijke eigenschap van kansen is dat de kans op waar plus de kans op onwaar 100% moet zijn. Dat lijkt logisch maar is dat niet altijd. Immers, de betrokkene kan bij gebrek aan informatie geen geloof hebben in zowel een bewering als in haar tegendeel; zie hiervoor Meester en Kerkvliet (2017).

De regel van Bayes is een eenvoudige formule die uit de axioma's van de kansrekening af te leiden is. Kort gezegd geeft deze formule een methode om 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*). De Bayesiaanse benadering is een *relatieve* benadering, in de zin dat het om *kansverhoudingen* tussen twee hypothesen gaat. Alleen als die twee hypothesen elkaars ontkenning zijn (dus elkaar uitsluiten en geen derde mogelijkheid toelaten) kunnen uit die verhouding ook de kansen op de hypothesen zelf worden uitgerekend, omdat hun kansen in dat geval zoals gezegd optellen tot 100%.

2.2 Statistische (on)afhankelijkheid

Een belangrijk begrip in de kansrekening is dat van *statistische (on)afhankelijkheid*. Twee beweringen zijn statistisch onafhankelijk van elkaar als informatie over het al dan niet waar zijn van de ene bewering irrelevant is voor het bepalen van de kans dat de andere bewering waar is. Als twee beweringen statistisch onafhankelijk van elkaar zijn, mogen we om de kans op de juistheid van beide beweringen te berekenen hun afzonderlijke kansen met elkaar vermenigvuldigen.

In strafzaken zijn bevindingen lang niet altijd statistisch onafhankelijk van elkaar. Sterker nog, bewijsmiddelen binnen dezelfde zaak zijn dat meestal niet. Zo wordt de kans dat een tweede beveiligingscamera de verdachte op een bepaalde locatie heeft opgenomen groter als een eerste beveiligingscamera de verdachte op dezelfde locatie heeft opgenomen (Fenton et al. 2013). Of stel dat DNA-sporen van de verdachte in een verkrachtingszaak op de blouse en op de rok van het slachtoffer zijn gevonden. De twee kansen op die sporen mogen niet met elkaar vermenigvuldigd worden, want zodra het spoor op de blouse is gevonden wordt de kans dat een spoor van dezelfde persoon op de rok wordt gevonden groter. Een beruchte zaak waarin dit mis ging was de Engelse strafzaak Sally Clark, waarin kort achter elkaar twee baby's uit hetzelfde gezin stierven terwijl ze alleen met hun moeder thuis waren. Een medisch expert

vermenigvuldigde de geschatte kans van 1 op de 8500 op een individueel sterfgeval met zichzelf om op een kans van 1 op de 73 miljoen uit te komen dat twee baby's uit hetzelfde gezin aan wiegendood sterven. Maar die vermenigvuldiging was volgens andere deskundigen in deze zaak onterecht, omdat de hiervoor benodigde aanname dat de twee sterfgevallen statistisch onafhankelijk van elkaar zijn ongegrond is. Na de eerste wiegendood is het heel goed mogelijk dat de kans op een tweede geval van wiegendood in hetzelfde gezin groter is geworden, want wiegendood zou bijvoorbeeld erfelijke of sociale oorzaken kunnen hebben (Dawid 2005, pp. 71-75).

2.3 Voorwaardelijke kansen

Een tweede belangrijk begrip in de kansrekening is dat van een voorwaardelijke kans. Een voorwaardelijke kans is de kans dat iets waar is of gebeurd is gegeven dat (oftewel aangenomen dat, of verondersteld dat) iets anders waar is of gebeurd is. Een onvoorwaardelijke en een voorwaardelijke kans op een gebeurtenis hoeven niet hetzelfde te zijn. Zo is de (onvoorwaardelijke) kans op het gooien van een zes met een zuivere dobbelsteen 1 op de 6, maar is de voorwaardelijke kans op het gooien van een 6 met een zuivere dobbelsteen, gegeven dat een even getal bovenkomt, 1 op de 3. We geven nu een paar juridisch relevante voorbeelden van voorwaardelijke kansen:

- De kans dat de verdachte op plaats delict was *gegeven* dat twee getuigen hem daar gezien hebben.
- De kans dat het op plaats delict gevonden DNA van de verdachte is *gegeven* dat het matcht met verdachtes DNA.
- De kans dat de verdachte het tenlastegelegde feit gepleegd heeft *gegeven* al het beschikbare bewijs.

In strafzaken zijn we uiteindelijk geïnteresseerd in de voorwaardelijke kans op schuld gegeven het beschikbare bewijs.

Ook de voorwaardelijke kansen dat een bewering waar, respectievelijk onwaar is gegeven een bepaalde andere bewering tellen op tot 1 (oftewel 100%).

2.4 Omkering van voorwaardelijke kansen: de 'prosecutor's fallacy'.

In de natuurlijke taal kunnen voorwaardelijke kansen op verschillende manieren uitgedrukt worden, en dat kan aanleiding geven tot interpretatieproblemen en misverstanden. Vaak gebeurt dit met zogenaamde 'random-match probabilities' bij sporenbewijs, zoals DNA-bewijs maar ook bewijs t.a.v. glassporen, bandensporen, enzovoorts. Forensische experts rapporteren hier de kans op een toevallige match, dat wil zeggen, de kans dat een gevonden spoor overeenkomt met dat een potentiële bron gegeven dat het spoor niet van die bron afkomstig is. Als die kans laag is (en doorgaans is dat zo), dan wordt ze vaak onterecht omgedraaid tot de lage kans dat het spoor niet van de potentiële bron afkomstig is gegeven dat het daarmee overeenkomt. Dit wordt wel de 'prosecutor's fallacy' genoemd (Thompson & Schumann 1987). Zie bijvoorbeeld een uitspraak van de rechtbank Noord-Holland van 13 november 2014 (ECLI:NL:RBNHO:2014:10689), waarin het bandenprofiel van de auto van de verdachte overeenkwam met op de plaats misdrijf aangetroffen bandensporen. Eerst

vermeldt de rechtbank een deskundigenschatting van de random-match probability van aangetroffen autobandsporen:

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

Dan maakt de rechtbank hier het volgende van:

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

Waarna de rechtbank concludeert “hetgeen een hoge mate van waarschijnlijkheid oplevert dat de auto van verdachte de bandensporen heeft gemaakt”.

Hier verwacht de rechtbank 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.

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. Meer in het algemeen maken mensen vaak de fout om als iets gebeurt waarop vooraf de kans erg klein was te denken: dat kan geen toeval meer zijn, er moet wel iets aan de hand zijn. Een bekend voorbeeld is de zaak Lucia de Berk, waarin zich het gevaar lijkt te hebben verwezenlijkt dat uit een schatting van een zeer kleine kans dat veertien zeer zieke kinderen in drie verschillende ziekenhuizen overlijden tijdens diensten van dezelfde verpleegster onterecht geconcludeerd wordt dat het daadwerkelijke overlijden van veertien zeer zieke kinderen tijdens diensten van Lucia de Berk geen toeval kon zijn en vervolgens dat Lucia de Berk er dus wel de hand in moest hebben gehad (Derksen & Meijnsing 2009, Meester et. al 2006). Maar deze redenering is verkeerd: als iets zeldzaams gebeurt is elke mogelijke oorzaak daarvan zeldzaam. Waar het om gaat is welke van de mogelijke zeldzame oorzaken het meest waarschijnlijk de werkelijke oorzaak is. Mensen die een grote loterij winnen worden ook niet vervolgd omdat vooraf de kans dat ze zouden winnen erg klein is.

In de zaak van de Oosterlandse brandstichtingen maakte Alkemade dezelfde fout. Op basis van statistieken gecombineerd met subjectieve aannames concludeerde hij ten aanzien van de 15 tenlastegelegde brandstichtingen dat de kans dat zich ergens in Nederland toevallig zo iets zou voordoen minder dan 1 op de miljoen is. Hieruit concludeerde hij onterecht dat de aanname dat deze brandstichtingen geen gemeenschappelijk verband zouden hebben onhoudbaar is.² Vervolgens concludeerde

² Overigens willen we hiermee niet zeggen dat de 15 brandstichtingen geen verband met elkaar houden, maar alleen dat Alkemade zijn conclusie dat ze wel verband met elkaar houden niet enkel kon

hij in één adem door dat die samenhang wel moest bestaan in het optreden van één of meer seriebrandstichters.

In de volgende paragraaf zullen we zien dat de regel van Bayes erg geschikt is om te laten zien dat dergelijke redeneringen incorrect zijn.

2.5 De regel van Bayes

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. Met een random-match probability van 1 op de 5000 zal dus naar verwachting van 1000 van die auto's het bandenprofiel overeenkomen met dat van de sporen op de plaats delict, en de auto van de daders (die het spoor gemaakt heeft) is daar maar één van. 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 regel van Bayes geeft een meer algemene manier om de kansverhouding tussen twee elkaar uitsluitende hypothesen (bijvoorbeeld de schuldhypothese versus de onschuldhypothese in een strafzaak) gegeven het aangetroffen bewijs (hieronder 'bevindingen' genoemd) te bepalen. Stel we hebben twee hypothesen en we hebben een bevinding. Om de a-posteriori kansverhouding van de hypothesen gegeven de bevinding te bepalen hebben we drie dingen nodig:

- De a-priori kansverhouding van de hypothesen
- De kans op de bevinding gegeven dat de eerste hypothese waar is
- De kans op de bevinding gegeven dat tweede hypothese waar is.

In ons autobandenvoorbeeld is de bevinding de overeenkomst tussen de gevonden bandensporen en de auto van de verdachte en is de eerste hypothese dat de sporen door de auto van de verdachte gemaakt zijn, terwijl de tweede hypothese stelt dat dat niet zo is en de sporen door een andere, onbekende auto zijn gemaakt. De a-priorikans op de eerste hypothese is 1 op de 5 miljoen, want we hebben aangenomen dat er 5 miljoen auto's zijn die de sporen gemaakt kunnen hebben. De a-priorikans op de tweede hypothese is bijna 1. De a-priori kansverhouding is dus ongeveer 1 gedeeld door 5 miljoen. De kans op de overeenkomst gegeven dat de auto van de verdachte de sporen gemaakt heeft is 1 (oftewel 100%), en de kans op de overeenkomst gegeven dat de auto van de verdachte de sporen niet gemaakt heeft is 1 op de 5000. Deze twee voorwaardelijke kansen bepalen samen de *bewijskracht* (hierboven ook de *likelihood ratio* genoemd) van de bevinding met betrekking tot de twee hypothesen: hoeveel waarschijnlijker of onwaarschijnlijker is de bevinding gegeven dat de eerste

baseren op de kleine kans op de brandstichtingen aangenomen dat er géén verband is tussen de brandstichtingen (kort gezegd: gegeven de hypothese van 'toeval').

hypothese waar is dan gegeven dat de tweede hypothese waar is? Deze bewijskracht vinden we door de twee voorwaardelijke kansen door elkaar te delen:

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

In ons voorbeeld is de kans boven de streep 1 (oftewel 100%), want aangenomen dat de auto van de verdachte de sporen gemaakt heeft, zal het bandenprofiel van die auto zeker overeenkomen met de sporen. De kans onder de streep is de random-match probability en die was door de deskundige geschat op 1 op de 5000. De bewijskracht van de overeenkomst is dus 1 gedeeld door $1/5000 = 5000$: de overeenkomst is 5000 maal zo waarschijnlijk wanneer de sporen door de auto van de verdachte gemaakt zijn dan wanneer ze door een andere, onbekende auto gemaakt zijn.

Maar dat is niet alles: de regel van Bayes zegt dat de bewijskracht vermenigvuldigd moet worden met de a-priori kansverhouding ten aanzien van de hypothesen, dat wil zeggen:

$$\frac{\text{De kans dat de eerste hypothese waar is}}{\text{De kans dat de tweede hypothese waar is}}$$

En omdat de a-priorikans op de overeenkomst 1 op de 5 miljoen is, is de a-priorikans dat de sporen niet met het profiel van de auto van de verdachte overeenkomen vrijwel 100%. Dus de a-priori kansverhouding is vrijwel 1 op de 5 miljoen.

De regel van Bayes zegt nu dat we om de a-posteriori kansverhouding op de hypothesen gegeven de bevindingen te berekenen, de bewijskracht moeten vermenigvuldigen met de a-priorikansverhouding.

$$\frac{\text{De kans dat de eerste hypothese waar is gegeven de bevinding}}{\text{De kans dat de tweede hypothese waar is gegeven de bevinding}}$$

=

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

x

$$\frac{\text{De kans dat de eerste hypothese waar is}}{\text{De kans dat de tweede hypothese waar is}}$$

Oftewel:

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

In ons voorbeeld is de a-posteriorikansverhouding dus $5000 \times 1 \text{ op de } 5 \text{ miljoen} = 1 \text{ op de } 1000$. 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 onterecht de a-priori kansverhouding tussen de twee hypothesen buiten beschouwing gelaten.

Dit is meestal nog niet alles, want vaak zal er meer bewijs zijn ten aanzien van dezelfde hypothesen. De regel van Bayes zegt dan dat we de a-priori kansverhouding t.a.v. de twee hypothesen achtereenvolgens moeten vermenigvuldigen met de bewijskracht van elke bevinding, als tenminste aan bepaalde onafhankelijkheidsaannames is voldaan. Op die manier wordt de a-posteriori kansverhouding gegeven de eerste bevinding die beschouwd is, de a-priori kansverhouding in de berekening met de volgende bevinding die beschouwd wordt, enzovoorts. Stel dat in onze voorbeeldzaak ander bewijs beschikbaar is met een bewijskracht van 500. Vermenigvuldiging van de a-priorikansverhouding van 1 op de 5 miljoen met 500 levert dan een a-posteriorikansverhouding gegeven dat andere bewijs op van 1 op de 10.000, en die moet vervolgens vermenigvuldigd worden met de bewijskracht van 5000 van het bandensporenbewijs, wat een nieuwe a-posteriorikans van iets meer dan 33% oplevert dat de auto van de verdachte de sporen gemaakt heeft, gegeven zowel het bandensporenbewijs en het andere bewijs. Dat is ook als volgt in te zien. Het andere bewijs heeft het aantal potentiële bronnen van de sporen teruggebracht tot 10.000. Met de random-match probability van 1 op de 5000 zal dan naar verwachting van twee van die 10.000 auto's het bandenprofiel toevallig overeenkomen met dat van de sporen op de PD. Verder zal ook weer het bandenprofiel van de auto van de daders met de sporen overeenkomen. Dus in totaal zal van drie van de 10.000 auto's het bandenprofiel overeenkomen met dat van de gevonden sporen, terwijl maar één van die auto's heeft de sporen gemaakt. De auto van de verdachte is één van die drie auto's dus de kans dat de auto van de verdachte de sporen gemaakt heeft gegeven de overeenkomst tussen de sporen en het andere bewijs is 1 op de 3, oftewel iets meer dan 33%. Ook hier hoeft het niet te stoppen, want verder bewijs kan deze kans weer veranderen: belastend bewijs (d.w.z. bewijs met een bewijskracht groter dan 1) zal de kans verhogen en ontlastend bewijs (d.w.z. bewijs met een bewijskracht lager dan 1) zal de kans verlagen.

Dit alles is precies de door Alkemade & Stikkelbroeck (2015a, 2015b) bepleite en de door Alkemade in de zaken Zes van Breda en de Oosterlandse brandstichtingen toegepaste denkwijze. Overigens is het hierbij van belang te beseffen dat een hoge a posteriori kans op schuld na het beschouwen van een aantal bevindingen geen enkele garantie is dat dit ook zo blijft bij meerdere bevindingen. Zelfs bij een zeer hoge a posteriori kans op schuld relatief aan bepaalde bevindingen kan elke nieuwe ontlastende bevinding de a posteriorikans sterk doen dalen. Elke bepaling van de a posteriori kans op schuld is dus relatief aan de beschouwde bevindingen. Het is daarom essentieel om goed onderzoek te doen naar alle mogelijke belastende en ontlastende bevindingen. Als dat in het misdaadonderzoek niet gebeurd is, kan dat in de rechtszaal een reden zijn om zelfs bij een hoge a posteriori-kans op schuld niet schuldig te verklaren, om de eenvoudige reden dat mogelijk ontlastend materiaal niet in aanmerking is genomen. Om die reden is het te kort door de bocht om, zoals Van den Berg (2014) doet, de *toereikendheid* van bewijs voor een veroordeling gelijk te stellen met een voldoende hoge a posteriori kans op de schuldhypothese gegeven het beschikbare bewijs.

Er speelt hier overigens ook nog iets anders mee. Zoals in elke methode om tot schuld of onschuld te redeneren staat het ieder vrij om naar believen belastende of ontlastende bewijsmiddelen aan te voeren. Dit geeft in een Bayesiaanse analyse ruimte voor manipulatie. Om een voorbeeld te noemen: als ontlastend materiaal zou de verdediging kunnen aanvoeren dat verdachte geen dagboek aantekening heeft gemaakt over het gebeurde. De bewijskracht van deze verklaring hangt af van de omstandigheden en zal meestal gering zijn, maar als je er maar genoeg van dergelijke “bewijsmiddelen” opneemt kan dat best accumuleren tot een serieus te nemen getal.

3 Voorwaarden voor de bruikbaarheid van Bayesiaans denken in de rechtspraktijk

De bovenstaande bespreking was bedoeld om zowel de kracht als de beperkingen van Bayesiaans denken te kunnen laten zien. De kracht van Bayesiaans denken zit vooral in de therapeutische en pedagogische waarde: met kleine voorbeeldjes kan betrekkelijk eenvoudig getoond worden dat bepaalde verleidelijke denkpatronen inzake kansen incorrect zijn. Kennis van de basisprincipes van Bayesiaans denken is dus voor juristen heel belangrijk om denkfouten te voorkomen. Maar de therapeutische en pedagogische waarde van Bayesiaans denken impliceert nog geen geschiktheid als algemeen denkmodel voor strafrechtelijk bewijzen in de praktijk. Over die geschiktheid gaat het in deze paragraaf.

3.1 Onbruikbaarheid van naïeve Bayes

De vraag is nu: werkt de Bayesiaanse denkwijze bij complexe strafzaken echt in de praktijk? Dat is nog maar zeer de vraag. Op kleine voorbeeldjes zoals hierboven werkt de Bayesiaanse denkwijze zoals die hierboven geschetst is prima, maar echte zaken zijn vaak veel ingewikkelder en dan blijft van de elegantie en helderheid van de Bayesiaanse denkwijze vaak weinig over. Om deze reden wordt bovenstaande kijk op Bayes wel *naïeve Bayes* of soms zelfs *idiot Bayes* genoemd. Er zijn (ten minste) twee problemen: statistische afhankelijkheden en stap-voor-stap bewijsconstructies.

Een eerste beperking van naïeve Bayes is dat ze alleen correct is als, gegeven de hypothesen, alle bewijsmiddelen statistisch onafhankelijk van elkaar zijn. Praktisch gesproken zegt deze aanname dat om de bewijskracht van een bevinding t.a.v. twee vergeleken hypothesen te schatten al het andere bewijs genegeerd kan worden. Alleen de volgende kansen hoeven op elkaar gedeeld te worden:

$$\frac{\text{De kans op de bevinding gegeven de ene hypothese}}{\text{De kans op de bevinding gegeven de andere hypothese}}$$

Zonder deze aanname wordt Bayesiaans updaten veel complexer. Dan moeten om de bewijskracht van een bevinding t.a.v. de twee vergeleken hypothesen te schatten de volgende voorwaardelijke kansen geschat worden:

$$\frac{\text{De kans op de bevinding gegeven de ene hypothese en al het andere bewijs}}{\text{De kans op de bevinding gegeven de andere hypothese en al het andere bewijs}}$$

En dat is in het algemeen veel moeilijker. Een voorbeeld. In de zaak van de Oosterlandse brandstichtingen schatte Alkemade zijn a-priori kansen met onder meer de bevindingen dat de verdachte door de politie verdacht werd van eerdere brandstichtingen en dat zijn medeverdachte veroordeeld was voor latere brandstichtingen. Verder gaf Alkemade de bevinding dat de verdachte en medeverdachte vrienden waren een zeer sterke bewijskracht: hij vond de vriendschap 200 maal zo waarschijnlijk onder de hypothese dat beiden actief waren als seriebrandstichter dan onder de hypothese dat ze geen van beiden actief waren. Beschouw nu de volgende voorwaardelijke kansen:

- (1) De kans dat de verdachte en medeverdachte vrienden waren gegeven dat geen van beiden actief was als seriebrandstichter
- (2) De kans dat verdachte en medeverdachte vrienden waren gegeven dat geen van beiden actief was als seriebrandstichter *en dat de medeverdachte door de politie verdacht werd van eerdere brandstichtingen en dat de verdachte veroordeeld is voor latere brandstichtingen.*

Ieder kan voor zichzelf bepalen of hij of zij vindt dat de tweede kans gelijk is aan de eerste. Ons gezond verstand zegt ons dat de tweede kans beduidend hoger is dan de eerste kans. Zodra bekend is dat de medeverdachte verdacht en de verdachte veroordeeld is voor andere brandstichtingen, hebben ze kennelijk iets gemeen (in karakter, achtergrond, omstandigheden, interesses), waardoor de kans op vriendschap ook onder de onschuldhypothese hoger wordt en daarmee de bewijskracht van de vriendschap lager.

Helaas zijn, zoals hierboven in paragraaf 2.2 opgemerkt, in de praktijk bewijsmiddelen vaak niet (conditioneel) onafhankelijk van elkaar, en het ontbreken van deze afhankelijkheden kan tot grote fouten leiden. Om dan toch Bayesiaans te kunnen blijven denken zijn zogenaamde Bayesiaanse netwerken nodig of in ieder geval bijzonder handig, waarin statistische afhankelijkheden op een grafische manier aangegeven kunnen worden en de methode van Bayesiaans updaten in gecompliceerder vorm toch toegepast kan worden.

Het tweede probleem is dat de relatie tussen bevinding enerzijds en schuld- of onschuldhypothese anderzijds vaak niet rechtstreeks is maar met tussenstappen gaat. Bezie weer ons voorbeeld. De eigenlijke schuldhypothese is niet dat de auto van de verdachte de bandensporen gemaakt heeft maar dat de verdachte betrokken was bij het misdrijf in kwestie. De stap van de overeenkomst tussen de sporen naar de betrokkenheid bij het misdrijf heeft op zijn minst twee tussenstappen: van 'de bandensporen op de PD en het bandenprofiel van de auto van de verdachte komen overeen' via 'de auto van de verdachte heeft de sporen gemaakt' naar 'de verdachte was op de PD' naar 'de verdachte was betrokken bij het misdrijf'. Naïeve Bayes kan zulke meerstapsverbanden tussen bevindingen en (on)schuldhypothese niet expliciet maken, maar bij complexe zaken moet dat wel gebeuren, om aan te geven *waarom* er bepaalde statistische invloeden tussen hypothesen en bewijsmiddelen bestaan. Ook hiervoor zijn Bayesiaanse netwerken heel handig.

We kunnen dus concluderen dat als we Bayesiaans denken willen toepassen in complexe strafzaken (voor het analyseren van de gehele zaak), dit niet met naïeve

Bayes kan maar op een veel complexere manier moet, waarbij Bayesiaanse netwerken een goed hulpmiddel zijn. Zijn Bayesiaanse netwerken dan de oplossing voor de twee problemen? Die vraag zullen we aan het eind van dit artikel bespreken.

3.2 Betrouwbaarheid van Bayesiaanse analyses van complexe strafzaken

Een van de vragen die in de zaak van de Oosterlandse brandstichtingen beantwoord moest worden is of de Bayesiaanse methode als methode ter beantwoording van de schuldvraag, zoals toegepast in Alkemades rapport in deze zaak, betrouwbaar is. Betrouwbaarheid betekent in dit geval of verschillende personen die deze methode op dezelfde zaak toepassen tot dezelfde uitkomst zouden komen. De vraag is in hoeverre Bayesiaanse netwerkmodellen van complexe strafzaken betrouwbaar kunnen zijn in deze zin.

De vraag naar de betrouwbaarheid van Bayesiaanse analyses wordt het vaakst gesteld ten aanzien van het schatten van kansen maar het is belangrijk te beseffen dat de betrouwbaarheid van zulke analyses van meer afhangt dan alleen het schatten van kansen. Bij een Bayesiaanse modellering moeten op zijn minst de volgende keuzes gemaakt worden:

1. Het opstellen van de hypotheses,
2. De selectie van mogelijk relevante bevindingen,
3. De bepaling van statistische afhankelijkheidsrelaties tussen de verschillende bevindingen gegeven de hypotheses,
4. Het schatten of poneren van de relevante kansen,
5. Bepalen of verdere potentiële bevindingen relevant zouden kunnen zijn.

De kwaliteit van al deze beslissingen samen bepaalt de betrouwbaarheid van een Bayesiaanse analyse. Maar laten we toch eerst de betrouwbaarheid van kansschattingen bespreken.

Als leken met toepassingen van de kansrekening geconfronteerd worden, dreigt het gevaar dat ze geïmponeerd raken door de wiskundige vorm van de berekeningen en conclusies en niet beseffen dat de uitkomst van een kansberekening net zo goed of slecht is als zijn uitgangspunten. Om betrouwbare conclusies uit een Bayesiaanse analyse te kunnen trekken is meer nodig dan alleen wiskundige berekeningen. Het zegt dan niets dat ze wiskundig uit de gegeven kansen volgen. Hier geldt: *garbage in, garbage out*. De kansrekening kan nog zo mooi in elkaar zitten, maar als men er ongefundeerde kansen instopt, komen er ongefundeerde kansen uit. Overigens bestaat dit probleem bij elke redeneermethode, maar bij kansrekening of statistiek is er nog het extra gevaar dat het gebruik van wiskunde de ontorechte schijn kan wekken van objectiviteit en wetenschappelijkheid.

Veel voorbeelden in leerboeken gaan over kansspelen en dat is niet zomaar: juist op kansspelen is de kansrekening zeer goed toepasbaar. Een kansspel is een kleine, kunstmatige en volledig gedefinieerde wereld, waarin de 'inputkansen' volgen uit de specificatie van het spel. In rechtszaken is dit heel anders: die gaan over de echte wereld, die groot, complex, veranderlijk en niet volledig kenbaar is. Dan zijn er in essentie drie manieren om kansen te schatten: *statistieken*, schatting door *experts* en

een beroep op het *gezond verstand* (of in juridische termen: de algemene ervaringsregels).

Bijkomende complicatie is dat deze kansen vaak verschillende interpretaties hebben, maar dat ze wel in dezelfde berekening gebruikt moeten worden. Een kans op basis van statistieken is heel wat anders dan een kans op basis van een expertopinie.

Statistieken als bron van kansen

Betrouwbare *statistieken* kunnen een bron zijn van (voorwaardelijke) kansen. De 'random match probabilities' bij DNA-bewijs worden bijvoorbeeld met statistische methoden bepaald op basis van frequenties van het voorkomen van bepaalde DNA-kenmerken in de bevolking. In de rechtspsychologie is empirisch onderzoek gedaan naar de betrouwbaarheid van herkenningen in Osloconfrontaties, waarbij een slachtoffer of ooggetuige van een misdrijf de dader moet aanwijzen uit een rij van een aantal op elkaar gelijkende personen. Zo geven Wagenaar & Van der Schrier (1996) de bewijskracht van herkenningen van verschillende afstanden en onder verschillende lichtsterktes. Ook statistieken over het 'gewone' leven kunnen een bron van kansen zijn, zoals in ons voorbeeld statistieken over de frequentie van bepaalde typen autobanden.

Overigens vereist een correct gebruik van statistieken niet alleen statistische maar ook inhoudelijke expertise. Ten eerste vereist de selectie en interpretatie van relevante en valide statistieken vakinhoudelijke kennis op het gebied waarover uitspraken in de vorm van kansschattingen worden gedaan. In de zaak van de Oosterlandse brandstichtingen gebruikte Alkemade onder meer statistieken omtrent brandstichtingen uit Japan en Groot-Brittannië. Om te bepalen of dit de juiste statistieken zijn, of het aangehaalde onderzoek deugdelijk is, of er misschien geen recenter relevant onderzoek is enzovoorts, is vakinhoudelijke kennis over brandstichtingen nodig. Hetzelfde geldt voor de vraag of die buitenlandse statistieken ook op Nederland van toepassing zijn en zo ja, of ze dan wel zonder meer van toepassing zijn op de situatie in Oosterland. Ook dat zijn vakinhoudelijke vragen, die alleen betrouwbaar beantwoord kan worden door een expert in brandstichtingen.

Een belangrijk punt van zorg omtrent statistische kansen is de vraag of statistische gegevens gebruikt kunnen worden in een rechtszaak tegen een individu. Als (stel) 60% van de vermoorde getrouwde vrouwen in Nederland door hun echtgenoot zijn omgebracht, mag dat gegeven dan gebruikt worden in een specifieke rechtszaak?

Deskundigen als bron van kansen

Als kansen niet op statistieken kunnen worden gebaseerd (en in de zaken van de Zes van Breda en de Oosterlandse brandstichtingen kon dat meestal niet), dan kunnen ze soms toch nog enigszins objectief vastgesteld worden door ze door *experts* te laten schatten. Zo'n beroep op een expert als bron van kansen is overigens net als een beroep op statistieken weerlegbaar, want experts mogen dan wel doorgaans gelijk hebben, daar zijn uitzonderingen op. Maar toch zouden rechtspsychologen bijvoorbeeld kansen kunnen schatten omtrent de betrouwbaarheid van getuigenverklaringen of bekentenissen. Of een medisch expert zou de kans kunnen

schatten dat een combinatie van alcohol en een bepaald medicijn iemand gewelddadig maakt. Het is hierbij essentieel dat het om deskundigen gaat op het gebied waarop de kansen geschat moeten worden.

Inzake de Zes van Breda en de Oosterlandse brandstichtingen constateren we dat Alkemade als klimaatfysicus geen enkele bewezen expertise bezit op de vele gebieden waarop hij kansen schat (en de andere modelleerbeslissingen neemt). In beide zaken was de betrouwbaarheid van verklaringen van getuigen en verdachten in het geding. In de zaak Zes van Breda was er verder onder meer sporenbewijs (DNA, bloed, haar), een verklaring van een anonieme informant, het feit dat verdachten naast het overvallen restaurant werkten, bewijs omtrent een mogelijke eerdere veroordeling, en het feit dat de verdachten elkaar kenden. En in de zaak van de Oosterlandse brandstichtingen ging het onder meer om de verschillende aspecten van brandstichtingen (de fysische en chemische maar ook de psychologische, sociologische en criminologische aspecten), de kans op vriendschappen in kleine gemeenschappen, en de bewijswaarde van recidive of verdenkingen van soortgelijke feiten. Op geen van deze gebieden bezit Alkemade als klimaatfysicus enige bijzondere expertise. Toch schat hij op al deze gebieden kansen. Dat is in strijd met het basale uitgangspunt in de forensische wetenschappen dat deskundigen alleen verklaren over het gebied van hun expertise.

We concluderen daarom dat Alkemades kansschattingen voor het grootste deel berusten op zijn eigen puur persoonlijke schattingen en aannames: niet op zijn expertise in de Bayesiaanse kansrekening, want die omvat slechts het rekenen met gegeven kansen en bevat geen enkel handvat voor het betrouwbaar schatten van de kansen waarmee wordt gerekend. En ook niet op zijn expertise als atmosferisch en klimaatfysicus, want die gebieden spelen in de onderhavige zaak geen rol. Om deze redenen hebben de kansschattingen van Alkemade niet meer waarde dan die van ieder willekeurig persoon die zich over de zaak zou kunnen buigen. Bovendien geldt deze conclusie niet alleen voor Alkemades kansschattingen, maar ook voor zijn beslissingen op alle punten die daaraan voorafgaan, zoals het opstellen van de relevante hypothesen, de selectie van de relevante bevindingen en de bepaling van hun statistische afhankelijkheden. Kortom, deze conclusie geldt voor zijn Bayesiaanse analyses van de zaken als geheel.

In andere gebieden waarin de Bayesiaanse kansrekening toegepast wordt, worden deze inzichten doorgaans gerespecteerd. Zo zal een expert in Bayesiaanse netwerken die een beslissingsondersteunend systeem voor medische diagnose wil bouwen (zoals collega's dat in Utrecht doen) niet zelf de kansen schatten maar dit overlaten aan een medisch expert. Alleen dan kunnen de kansen betrouwbaar geschat worden.

Ten aanzien van Alkemade bestaat bovendien specifieke twijfel over het niveau van zijn deskundigheid in de kansrekening (het ontbreken van peer-reviewed publicaties op relevante gebieden, het vasthouden aan een methode die algemeen in de literatuur als onbruikbaar wordt beschouwd, het begaan van de prosecutor's fallacy in de zaak van de Oosterlandse brandstichtingen). Bij erkende experts in de kansrekening (zoals vele statistici) geldt deze twijfel vanzelfsprekend niet, maar het is belangrijk te zien dat de bovenstaande conclusies ook voor hen gelden. Ook voor statistici geldt dat voor zover ze geen inhoudelijke deskundigheid beschikken op de gebieden waarop ze

kansen schatten, hun schattingen (en andere modelleerbeslissingen) niet op inhoudelijke expertise berusten en daarom wellicht niet betrouwbaar zijn. Veel statistici erkennen dat overigens en benadrukken dat het daarom belangrijk is dat een statisticus zijn of haar subjectieve aannames expliciet maakt zodat ze door de rechter gecontroleerd kunnen worden. Dat is inderdaad een goed gebruik, maar dit is alleen zinvol als de kansrekening op deelpunten gebruikt wordt en in combinatie met betrouwbare statistieken. Bij gebruik van de Bayesiaanse kansrekening om complexe zaken in hun geheel te analyseren (en daar hebben we het in dit artikel over) is de kans groot dat vrijwel de gehele analyse één grote subjectieve aanname is, ook als die door een gekwalificeerde statisticus gemaakt wordt.

Het gezond verstand als bron van kansen

Tot slot het *gezond verstand*. Veel bewijs in strafzaken gaat niet over kwesties waarover statistieken bestaan of waarover experts kunnen verklaren, maar over het leven van alledag. Stel bijvoorbeeld dat iemand wegrent van de plaats delict als de politie daar arriveert. Iedereen voelt aan dat dit relevant is, maar betrouwbare statistieken of deskundigen omtrent de kans dat dit gebeurt bij schuld dan wel onschuld van degene die wegrent zullen er niet zijn. Twee voorbeelden uit de Zes van Breda: daarin schatte Alkemade kansen omtrent het feit dat de verdachten elkaar kenden en omtrent het feit dat twee verdachten naast het overvallen restaurant werkten. Ook hierover bestaan geen statistieken of bijzondere deskundigheid. Een rechter die volgens de kansrekening wil denken zal dan zijn of haar gezond verstand (in juridische termen deel van de algemene ervaringsregels) moeten gebruiken om kansen te schatten, als dat überhaupt mogelijk is.

Het is duidelijk dat lang niet in alle gevallen een betrouwbare of zelfs maar betekenisvolle inschatting van kansen kan worden gemaakt. Het zal dan ook vaak zo zijn dat de regel van Bayes vooral kwalitatief gebruikt wordt. Met deze regel wapent men zich tegen redeneerfouten, zoals bijvoorbeeld in het eerder genoemde geval van Sally Clark. Als men destijds de regel van Bayes had gebruikt, dan had men kunnen inzien dat men de a-priorikans op dubbele moord helemaal niet beschouwd had, terwijl dat wel had moeten.

3.3 Vermeend nut van onbetrouwbare Bayesiaanse analyses

Welke waarde kan een deskundigenrapport hebben als de analyse niet op inhoudelijke deskundigheid gebaseerd is? Alkemade claimt in zijn rapporten dat de waarde van zijn analyse elders ligt, namelijk dat het laat zien welke vragen beantwoord moeten worden. Maar ook daarvoor kan hij aan de wetenschap geen beproefde methode ontlenuen. Bovendien rijst de vraag hoe iemand de juiste vragen kan stellen op een gebied waarop hij geen deskundigheid bezit. Om een analogie te gebruiken, zou een medisch expert die een ernstig zieke patiënt moet behandelen aan een klimaatfysicus advies vragen over welke vragen hij moet stellen en welke onderzoeken hij moet doen? Natuurlijk niet: de medisch specialist zou heel goed beseffen dat alleen iemand met inhoudelijke medische kennis op deze punten betrouwbaar advies kan geven.

Meer in het algemeen stellen Bayesianen dat ook wanneer er geen volledig betrouwbare maatstaven zijn voor het schatten van kansen, er in ieder geval

‘sensitivity analysis’ mogelijk is, die de invloed van verschillende schattingen op de a posteriori-kans berekent. Dat kan zeker nuttig zijn, vooral op deelpunten, en veel software voor Bayesiaanse netwerken ondersteunt sensitivity analysis. Maar voordat kansen kunnen worden bepaald, moeten zoals gezegd eerst andere keuzes worden gemaakt, namelijk van de te onderzoeken hypothesen, de mogelijk relevante bevindingen en hun statistische afhankelijkheidsrelaties. Al die elementen samen bepalen welke kansen moeten worden bepaald en als de keuzes op die punten niet op inhoudelijke deskundigheid gebaseerd zijn, dan heeft sensitivity analysis vervolgens weinig zin. Dat heeft alleen zin als in ieder geval de ‘lay-out’ van een Bayesiaans netwerk betrouwbaar te bepalen is, en hoe complexer de zaak, hoe minder snel dat het geval zal zijn. Bovendien is er meestal niet zoiets als “de” lay-out, en is de kans groot dat verschillende deskundigen tot verschillende structuren komen.

3.4 Conclusies

De eerste vraag die we in dit artikel wilden beantwoorden is of Bayesiaans denken op een wetenschappelijk betrouwbare en nuttige wijze toegepast kan worden in de rechtszaal ter analyse van complexe strafzaken in hun geheel. We beantwoorden deze vraag nu in de vorm van enkele conclusies uit het voorgaande.

1. De wetenschap biedt geen beproefde en algemeen aanvaarde methode voor de toepassing van Bayesiaans denken in complexe strafzaken.
2. Er is consensus in de wetenschap dat de specifieke versie van de Bayesiaanse methode die door Alkemade in de zaken van de Zes van Breda en de Oosterlandse brandstichtingen is toegepast, namelijk met een globale statistische onafhankelijkheidsaannname (ook wel *naïeve Bayes* genaamd), niet toepasbaar is op complexe strafzaken. Als Bayesiaans denken op complexe zaken toepasbaar zal zijn, dan zal dat zijn met behulp van Bayesiaanse netwerken.
3. De betrouwbaarheid van een Bayesiaanse analyse hangt cruciaal af van de inhoudelijke expertise van de analist op het gebied van de zaak. In de zaken van de Zes van Breda en de Oosterlandse brandstichtingen miste Alkemade (zijnde klimaatfysicus) die inhoudelijke expertise.
4. Ook het nut van een niet-betrouwbare Bayesiaanse analyse om te laten zien welke vragen beantwoord moeten worden hangt cruciaal af van de inhoudelijke expertise van de analist op het gebied van de zaak.
5. Sensitivity analysis om de invloed van variaties in kansschattingen te onderzoeken heeft alleen zin als een zaak niet te complex is.

4 Huidig gebruik van de Bayesiaanse kansrekening in de rechtspraktijk

In de rechtspraktijk wordt de Bayesiaanse kansrekening afgezien van de optredens van Alkemade en Derksen tot nu toe vrijwel uitsluitend gebruikt op deelaspecten en op zogenaamd bronniveau, waarbij het de vraag is van welke bron een bepaald spoor (bijvoorbeeld DNA, voetsporen of vingerafdrukken) afkomstig is. Dit gebruik beperkt zich tot het vaststellen van de bewijskracht van sporenbewijs.

In drie recente rechtszaken is de toelaatbaarheid van Bayesiaanse analyses van complexe strafzaken aan de orde gekomen: de bovengenoemde twee zaken waarin Alkemade en een van ons optraden en een derde zaak.

Eerst de Zes van Breda. Het gerechtshof Den Haag heeft in zijn arrest van 14 oktober 2015 (ECLI:NL:GHDHA:2015:2860) eerst geoordeeld dat Alkemade “voor de onderhavige procedure als deskundige kan worden aangemerkt”.

“Naar het oordeel van het hof dient, indien sprake is van de introductie in het strafrecht van niet eerder bekende wetenschapsgebieden en -methoden, bij de beoordeling of een voorgedragen deskundige als zodanig kan worden beschouwd, een enigszins ruimer beoordelingskader te worden gehanteerd dan bij de beoordeling van deskundigen op binnen de rechtspraak reeds langer bekende wetenschapsgebieden. Een al te strikte toetsing (bijvoorbeeld op het punt van ervaring) zou immers het risico met zich kunnen brengen dat op bepaalde terreinen aanvankelijk niemand voldoende deskundig zou kunnen worden geacht, waardoor ook de introductie van nieuwe wetenschappelijke inzichten in het recht zou kunnen stagneren.

Het hof stelt vast dat [deskundige 1] een opleiding als fysicus heeft genoten, dat hij is gepromoveerd en ervaring heeft met de toepassing van Bayesiaanse analyse, zij het op andere wetenschaps- en onderzoeksgebieden dan het recht. Voorts geeft hij op het opleidingsinstituut voor de rechterlijke macht (SSR-) cursussen op het gebied van de Bayesiaanse statistiek en analyse. Blijkens zijn verklaring ter terechtzitting is hij eenmaal eerder zij het niet als deskundige maar als adviseur betrokken geweest bij het Bayesiaans benaderen van een strafzaak. Het hof constateert voorts dat [deskundige 1] in zijn rapport en ter terechtzitting in herziening uitvoerig heeft aangegeven volgens welke methode hij zijn onderzoek heeft uitgevoerd, op welke wijze hij deze methode heeft uitgevoerd en waarom deze methode zijns inziens betrouwbaar kan worden geacht. Gezien voorgaande feiten en omstandigheden, en in aanmerking nemende het hiervoor geschetste enigszins ruimere toetsingskader, is het hof van oordeel dat [deskundige 1] voor de onderhavige procedure als deskundige kan worden aangemerkt.”

Of voor Bayesiaanse deskundigheid minder strikte eisen moeten gelden dan in het algemeen voor deskundigheid is een rechtspolitieke kwestie, waar wij geen oordeel over zullen geven. Wel valt op dat het hof niet expliciet aangeeft op welk gebied Alkemade als deskundige kan worden aangemerkt. De overwegingen van het hof doen vermoeden dat het hier om deskundigheid op het gebied van de Bayesiaanse kansrekening gaat, omdat het hof niets overweegt ten aanzien van Alkemades deskundigheid op de vele gebieden waarop hij kansen schat (waaronder sporenbewijs, getuigenbewijs, en alibi's en bekentenissen van de verdachten). Het hof gaat hiermee voorbij aan onze argumentatie dat kennis van de Bayesiaanse kansrekening geen inhoudelijke kennis op deze gebieden impliceert, en dat Alkemade als klimaatfysicus op geen van deze gebieden gebleken inhoudelijke deskundigheid bezit. Waarom het hof hieraan voorbijgaat is uit het arrest niet op te maken.

Vervolgens heeft het hof over de Bayesiaanse methode geoordeeld

“... dat de betrouwbaarheid van voormelde methode en daarmee van de uit het gebruik van de methode voortvloeiende uitkomsten, naar de huidige stand van de wetenschap te onzeker is voor daadwerkelijke toepassing in een complexe strafzaak als de onderhavige. Om deze reden heeft het hof de conclusie(s) uit voormelde rapportage van dr. Alkemade bij de beoordeling van het bewijs dan ook geheel buiten beschouwing gelaten.”

Dit oordeel komt overeen met onze conclusies.

In de zaak van de Oosterlandse brandstichtingen oordeelde het hof niet over de vraag of Alkemade voor de procedure als deskundige kon worden beschouwd. Het hof beperkte zich in zijn arrest tot de volgende korte overweging:

“Het hof heeft het rapport van dr. Alkemade, gelet op de tegenwerpingen van prof.dr.mr. H. Prakken in het tegenonderzoek, niet laten bijdragen aan het oordeel van het hof omtrent de bewezenverklaring.”

De rechtbank Zeeland-West-Brabant merkte in haar vonnis in de zaak ECLI:NL:RBZWB:2016:3060 het volgende op over een deskundigenrapport van prof. Ton Derksen waarin de Bayesiaanse kansrekening werd gebruikt:

“De berekening van Derksen waaruit zou blijken dat verdachte – kort gezegd – waarschijnlijk onschuldig is, is gemaakt volgens de theorema van Bayes. Daarmee is een – naar het oordeel van de rechtbank – niet onomstreden regel uit de kansrekening gebruikt voor de strafrechtelijke waarheidsvinding waarvan de uitkomst in hoge mate af hangt van de selectie en waardering van de bewijsmiddelen die door Derksen is gemaakt.

De rechtbank sluit niet uit dat deze methode, mits juist toegepast, een hulpmiddel kan zijn bij de waarheidsvinding, maar de uiteindelijke beantwoording van de vraag of een tenlastegelegd feit wettig en overtuigend kan worden bewezen, is geen vraag die zich leent voor kansrekening. De beantwoording van die vraag is en blijft voorbehouden aan de rechter. Daarbij komt dat de rechtbank in deze zaak een andere selectie en waardering van de bewijsmiddelen maakt dan Derksen.

De rechtbank laat om bovengenoemde redenen de rapporten van Derksen buiten beschouwing bij de beantwoording van de vragen van de artikelen 348 en 350 Sv ...”

De nadruk die de rechtbank legt op de afhankelijkheid van de uitkomst van een Bayesiaanse analyse van de selectie en waardering van de bewijsmiddelen doet vermoeden dat de rechtbank goed begrepen heeft dat de criteria voor de betrouwbaarheid van kansschattingen buiten de Bayesiaanse kansrekening liggen. We merken ten slotte op dat de regel van Bayes zelf natuurlijk helemaal niet omstreden is, zoals de rechtbank beweert, want ze is wiskundig afleidbaar uit de axioma's van de kansrekening. Wat wel omstreden is, is de toepassing ervan bij bewijs in strafzaken, en dat is ongetwijfeld wat de rechtbank bedoelt.

5 Richtlijnen voor Bayesiaanse analyses van complexe strafzaken door deskundigen in een rechtsgeding

Op grond van het bovenstaande kunnen we nu bespreken hoe een Bayesiaanse analyse van een complexe strafzaak door deskundigen uitgevoerd zou moeten worden. Vervolgens kunnen we zien of zo'n analyse uitvoerbaar is en zo ja, nuttig kan zijn.

Vooraf is het goed om de verschillende aannames waaronder we deze richtlijnen geven op een rijtje te zetten. Ten eerste gaat het om toepassing van Bayes in de context van een rechtsgeding (zowel de opsporings- en vervolgingsfase als het onderzoek ter rechtszitting). De richtlijnen gelden niet voor wetenschappelijke case studies. Overigens zijn wij van zulke case studies een groot voorstander.

Ten tweede stellen we hier de theoretische grondslagen van de Bayesiaanse denkwijze als rationeel model van juridisch bewijzen niet ter discussie. Dat kan wel degelijk en het gebeurt ook, zoals in Prakken (2014) en Meester & Kerkvliet (2016, 2017). Maar in ieder geval is Bayesiaans denken, hoewel niet onomstreden, een theoretisch plausibele en door veel wetenschappers aanbevolen wetenschappelijk methode voor redeneren in bewijskwesities en dat is voldoende reden om te bespreken of en zo ja hoe die methode in de rechtspraktijk toegepast kan worden.

Tot slot zijn de richtlijnen niet bedoeld voor gebruik van de Bayesiaanse kansrekening op deelpunten, zoals het gebruik door forensische wetenschappers van likelihood ratio's of random match probabilities bij individuele bewijsmiddelen. Ze zijn bedoeld voor globaal gebruik van Bayesiaans denken zoals door Alkemade in de zaken van de Zes van Breda en de Oosterlandse brandstichtingen, namelijk het op basis van de totaliteit van de bevindingen uitspraken doen over de kans op schuld van de verdachte.

Richtlijn 1: Er moet samenwerking zijn tussen gebiedsexperts en Bayesiaanse experts.

In paragraaf 3.2 hebben we gezien dat expertise in de kansrekening geen inhoudelijke expertise in de aspecten van een zaak impliceert en dat de verschillende modelleerbeslissingen (bepalen van de relevante bevindingen, tussen- en eindhypothesen, bepalen van statistische afhankelijkheden en schatten van kansen) door op het relevante gebied inhoudelijk deskundige experts genomen moeten worden. Maar gebiedsexperts zijn doorgaans geen experts in de Bayesiaanse denkwijze, zeker niet als die inhoudt dat een Bayesiaans netwerk geconstrueerd moet worden. Daarom moet een Bayesiaanse analyse van een complexe strafzaak in samenwerking tussen gebieds- en Bayesiaanse deskundigen opgesteld worden.

Richtlijn 2: verschillende soorten bevindingen vereisen verschillende soorten expertise dus (doorgaans) is meer dan één gebiedsdeskundige nodig.

Complexe strafzaken zullen een variëteit aan categorieën van bevindingen hebben. Daarom is één gebiedsdeskundige niet genoeg; voor elk type bevinding moet er een bijbehorende expert zijn. Zo zou er in de Zes van Breda op zijn minst een

psychologisch expert, een bloedexpert en een DNA-expert moeten zijn. En in de zaak van de Oosterlandse brandstichtingen zou er op zijn minst een psychologisch expert op het gebied van verklaringen, één of meer experts in de verschillende aspecten van brandstichtingen (psychologisch, sociologisch, technisch) en een sociologisch expert omtrent groepsgedrag in kleine gemeenschappen moeten zijn.

Teamwerk voorkomt ook het gevaar dat een enkele deskundige op een gegeven moment overtuigd raakt van de schuld dan wel onschuld van de verdachte en zich in de verdere analyse ongemerkt door die overtuiging laat leiden. Dit zou gebeurd kunnen zijn in de zaak van de Oosterlandse brandstichtingen. In het begin van zijn analyse concludeerde Alkemade met een instantie van de prosecutor's fallacy (gegeven de toevalshypothese zijn deze 15 brandstichtingen zeer onwaarschijnlijk, dus ze kunnen geen toeval zijn) dat er wel een verband moest zijn tussen de 15 tenlastegelegde brandstichtingen. Vervolgens concludeerde hij in één adem door dat die samenhang wel moest bestaan in het optreden van één of meer seriebrandstichters. Pas in een later deel van zijn rapport schatte Alkemade de a-priorikansen van zijn hypothesen en de bewijskracht van de bevindingen (we weten uit paragraaf 2.5 dat deze schattingen noodzakelijk zijn om überhaupt te kunnen oordelen of de 15 brandstichtingen berusten op toeval). Het is denkbaar dat Alkemades ongefundeerde initiële oordeel dat de 15 brandstichtingen wel moesten samenhangen zijn hele denken over de zaak ten ongunste van de verdachte heeft beïnvloed. Dit gevaar is nog groter als (zoals in Alkemades rapporten) de kansschattingen niet gebaseerd zijn op inhoudelijke deskundigheid maar op puur subjectieve oordelen.

Richtlijn 3: Subjectieve aannames en schattingen op basis van gezond verstand moeten expliciet als zodanig aangeduid worden.

In paragraaf 3.2 hebben we ook gezien dat veel aspecten van een zaak niet onder de expertise van een deskundige valt maar een kwestie van gezond verstand of zelfs puur subjectieve aannames zijn. Om die reden moet in de analyse precies en expliciet worden aangegeven welke schattingen en andere beslissingen gebaseerd zijn op gebiedsdeskundigheid en welke slechts een kwestie van gezond verstand of puur subjectieve aannames zijn.

Richtlijn 4: Er zijn op zijn minst twee alternatieve analyses nodig van twee verschillende groepen experts.

Tot slot hebben we in paragraaf 3.2 gezien dat in een Bayesiaanse analyse van een complexe zaak niet alleen kansen moeten worden geschat maar ook de relevante bevindingen en tussenhypothesen moeten worden geselecteerd en statistische afhankelijkheden moeten worden bepaald. Omdat voor het maken van al deze beslissingen nog geen duidelijke wetenschappelijk gefundeerde methodes bestaan, kan dit in een complexe zaak al gauw maken dat verschillende teams van experts tot significant verschillende Bayesiaanse modelleringen komen, zelfs als alle experts hun schattingen zoveel mogelijk op hun expertise baseren. Dit geldt temeer vanwege het hierboven besproken probleem dat Bayesiaanse analyses onvermijdelijk subjectieve aannames zullen bevatten. Om die redenen moet een zaak idealiter door op zijn minst twee teams van deskundigen geanalyseerd worden.

6 Tot slot

Dan is er nog het probleem van de begrijpelijkheid van een Bayesiaanse analyse voor de bij een zaak betrokken juristen. De meeste juristen zijn niet gewend om met wiskundige formules om te gaan en bovendien is Bayesiaans denken voor veel mensen tegenintuïtief. Er is veel empirisch bewijs dat correct Bayesiaans denken voor niet-statistici (en zelfs soms voor statistici) te hoog gegrepen is (bijv. Tversky & Kahneman 1974, Kahneman 2011). Het bewijs is zelfs zo overvloedig dat het nog maar zeer de vraag is of de oplossing ligt in beter onderwijs en betere training van juristen (Sjerps & Berger, 2011, p. 2 suggereren dat dit de juiste aanpak is). Het zou daarom kunnen dat een Bayesiaanse denktrant bij bewijs in complexe strafzaken de zaak niet zal verbeteren maar eerder zal verslechteren (zie ook Lempert, 1986).

Veel forensische wetenschappers en statistici vestigen hun hoop op Bayesiaanse netwerken. Ten eerste zijn er efficiënte computerprogramma's die de constructie van en het rekenen met Bayesiaanse netwerken ondersteunen. Verder zijn Bayesiaanse netwerken volgens bijvoorbeeld Fenton & Neil (2011) geschikt om Bayesiaanse analyses aan statische leken (zoals rechters en jury's) uit te leggen zonder in wiskundige details te hoeven treden. De grafische structuur van een netwerk zou zonder wiskundige kennis begrepen kunnen worden en dan zouden rechters alleen nog maar de a-priori- en voorwaardelijke kansen die in zo'n netwerk ingevoerd worden hoeven te controleren (bijvoorbeeld door te checken of ze op deskundigenverklaringen gebaseerd kunnen worden of door ze aan hun eigen gezond verstand te toetsen). De rechters hoeven dan niet de door de computer gemaakte berekening van de a posteriori-kansen te begrijpen: net zoals mensen de uitkomsten van rekenmachientjes accepteren zonder de details van de berekening te hoeven zien, zouden rechters ook de berekening van het Bayesiaans netwerk kunnen accepteren zonder de details van de berekening te hoeven zien. In beide gevallen zou het voldoende zijn dat de rechter er op kan vertrouwen dat de juiste wiskundige theorie ingeprogrammeerd is.

Toch is dat te simpel gesteld. Ten eerste wordt de grafische structuur van Bayesiaanse netwerken al gauw heel complex, zeker bij complexe strafzaken. Er is wel steeds meer onderzoek naar het op een begrijpelijk manier opstellen van Bayesiaanse netwerken (Hepler et al. 2007; Fenton et al. 2013; De Zoete et al. 2014; Timmer et al. 2017; Vlek et al. 2016), maar dat is nog in volle gang en er is nog geen enkel empirisch onderzoek naar de vraag of de structuur van zulke netwerken voor juristen inderdaad begrijpelijk is. Ten tweede is het niet voldoende dat rechters de inputkansen controleren. Zoals we in paragraaf 3.2 hebben gezien, moeten voor er kansen geschat kunnen worden, eerst de relevante bevindingen, tussen- en eindhypothesen en statistische afhankelijkheden bepaald worden. Met andere woorden: eerst moet de structuur van het Bayesiaans netwerk vastliggen; en de rechter moet ook kunnen controleren of die structuur deugdelijk bepaald is. En omdat die structuur in de praktijk nooit helemaal zonder subjectieve keuzes en aannames bepaald kan worden, kan een rechter hier nooit volledig op de experts vertrouwen, maar moet hij of zij die subjectieve keuzes zelf toetsen. Dat vereist op zijn minst enig inzicht in de theorie van Bayesiaanse netwerken.

Overigens kunnen onze conclusies voor een flink deel gegeneraliseerd worden tot andere benaderingen van rationeel juridisch bewijzen, zoals de argumentatie- en de verhaalsbenadering of combinaties daarvan (zie voor meer details Prakken 2014, 2016). De richtlijnen van paragraaf 5 gelden grotendeels ook voor die benaderingen, want ook daarin moeten relevante bevindingen en hypothesen bepaald worden en moet de aard en sterkte van hun onderlinge verbanden bepaald worden (zij het nu zonder getallen) en ook daarbij rijst de vraag naar de betrouwbaarheid van keuzes op die punten. Bovendien kan bij een complexe zaak ook een analyse in termen van argumentatie of scenario's complex worden. Aan de andere kant claimen proponenten van deze denkwijzen dat ze intuïtiever en daardoor begrijpelijker zijn dan Bayesiaans denken en dat ze niet het probleem oproepen van de schijn van precisie gewekt door de wiskundige vorm van Bayesiaans denken en door het feit dat Bayesiaans denken vooral door exacte wetenschappers toegepast wordt. Bovendien heeft voor zover bekend nog nooit iemand geprobeerd om als deskundige in een rechtszaak een complexe zaak in zijn geheel met argumentatie- of scenario-denken te analyseren. Bij deze benaderingen komt kennelijk minder snel het onjuiste idee op dat expertise in deze denkwijzen expertise in de aspecten van een strafzaak impliceert.

We beantwoorden nu de tweede vraag van dit artikel, namelijk of wetenschappelijk betrouwbaar en nuttig gebruik van Bayesiaans denken in de rechtszaal praktisch gezien mogelijk is en zo ja, onder welke voorwaarden. De voorwaarden hebben we geschetst in de vorm van de vijf richtlijnen in Paragraaf 5, en in deze paragraaf hebben we geconcludeerd dat hun vervulling in de praktijk zeer problematisch is. Voorlopig lijkt toepassing van de regel van Bayes op deelvragen (zoals sporenbewijs) door forensisch deskundigen het hoogst haalbare, plus 'therapeutisch' gebruik van Bayesiaans denken om veel voorkomende statistische drogredenen tegen te gaan, bijvoorbeeld zoals een van ons (RM) in het hoger beroep van de zaak Lucia de Berk geprobeerd heeft om de AG en raadsheren in de zaak drogredenen te doen vermijden; zie Meester et al. (2006) en ook Grünwald (2011). Dat laatste zou wel eens het meest tastbare nut van Bayesiaans denken kunnen zijn. Om die reden is onderwijs en training van Bayesiaans denken aan rechtenstudenten en praktijkjuristen erg belangrijk. Maar systematische toepassing in de rechtszaak is een heel andere zaak.

Literatuur

Alkemade, F. & Stikkelbroeck, H. (2015a) In onderling verband en samenhang bezien (een praktische uitleg van de Bayesiaanse aanpak in het strafrecht). Deel 1: Het vaststellen van de bewijskracht van bevindingen en de regel van Bayes. *Trema* November 2015, 285-297.

Alkemade, F. & Stikkelbroeck, H. (2015b) In onderling verband en samenhang bezien (een praktische uitleg van de Bayesiaanse aanpak in het strafrecht). Deel 2: Het bepalen van de Bayesiaanse 'prior odds' in strafzaken. *Trema* December 2015, 340-349.

Anderson, T.J., Schum, D.A. & Twining, W.L. (2005). *Analysis of Evidence*, second edition. Cambridge: Cambridge University Press.

Berg, B. van den (2014), Een onoverbrugbare kloof. Een vergelijking van de strafrechtelijke en de forensische zienswijze bij de waardering van feiten. *Nederlands Juristenblad* 2014/1057, afl. 21, p. 1420-1426.

- Berger, C.E.H. & Aben, D.J.C. (2010a), Bewijs en overtuiging: Rationeel redeneren sinds Aristoteles. *Expertise en Recht* 2010-2, 52-56.
- Berger, C.E.H. & Aben, D.J.C. (2010b), Bewijs en overtuiging: Redeneren in de rechtszaal. *Expertise en Recht* 2010-3, 86-90.
- Berger, C.E.H. & Aben, D.J.C. (2010c), Bewijs en overtuiging: Een helder zicht op valkuilen. *Expertise en Recht* 2010-5/6, 159-165.
- Bex, F.J., Prakken, H., Reed, C. & Walton, D.N. (2003). Towards a formal account of reasoning about evidence: argumentation schemes and generalisations. *Artificial Intelligence and Law* 11: 125–165.
- Bex, F.J., Koppen, P.J. van, Prakken, H. & Verheij, B. (2010). A Hybrid formal theory of arguments, stories and criminal evidence. *Artificial Intelligence and Law* 18: 123–152.
- Crombag, H.F.M., Koppen, P.J. van & Wagenaar, W.A. (1994) *Dubieuze Zaken: de Psychologie van Strafrechtelijk Bewijs* 2nd edition, Contact, Amsterdam.
- Dawid, Ph. (2005). Probability and proof. Online appendix to “Analysis of Evidence” by T.J. Anderson, D.A. Schum and W.L. Twining.
<http://www.statslab.cam.ac.uk/~apd/>
- Derksen, T. & Meijsing, M. (2009). The fabrication of facts: the lure of the incredible coincidence. In H. Kaptein, H. Prakken & B. Verheij, (eds.): *Legal Evidence and Proof: Statistics, Stories, Logic (Applied Legal Philosophy Series)*. Farnham: Ashgate, 2009, pp. 39-70.
- Fenton, N.E., Lagnado, D. & Neil, M. (2013). A general structure for legal arguments using Bayesian networks. *Cognitive Science* 2013-37, p. 61-102.
- Fenton, N.E. & Neil, M. (2011). Avoiding legal fallacies in practice using Bayesian networks', *Australian Journal of Legal Philosophy* 36, 114-151, 2011.
- Grünwald, P. (2011). Over het bedrijven van statistiek in kansloze situaties. Voordracht Zwolle, 18 mei 2011. <http://www.luciadeb.nl/metta/Grunwalds-Metta-analyse.pdf>.
- Hepler, A.B., Dawid, A.P. & Leucari, V. (2007). Object-oriented graphical representations of complex patterns of evidence. *Law, Probability & Risk* 6 (1–4): 275–293.
- Kahneman, D. (2011). *Thinking, Fast and Slow*. Penguin, London.
- Koppen, P.J. van (2011), *Overtuigend Bewijs. Indammen van Rechterlijke Dwalingen*. Amsterdam: Nieuw Amsterdam 2011.
- Lempert, R. (1986) The new evidence scholarship: analyzing the process of proof. *Boston University Law Review* 1986-66: p. 439-477.
- Meester, R. (2015), When you have a hammer, everything looks like a nail. Een reactie op ‘In onderling verband en in samenhang bezien’. *Trema* December 2015, 355-359.
- Meester, R. & Kerkvliet, T. (2016) Assessing forensic evidence by computing belief functions. *Law, Probability and Risk* doi 10.1093/lpr/mgw002 (2016).

- Meester, R. & Kerkvliet, T. (2017) A behavioural interpretation of belief functions, *Journal of Theoretical Probability*, DOI 10.1007/s10959-017-0776-y.
- Meester, R., Collins, M., Gill, R. & Lambalgen, M. van (2006). On the (ab)use of statistics in the legal case against the nurse Lucia de B. *Law, Probability & Risk* 5: 233–250.
- Prakken, H. (2004). Analysing reasoning about evidence with formal models of argumentation. *Law, Probability & Risk* 3:33–50.
- Prakken, H. (2014) Strafrechtelijk bewijzen: met Bayes of met verhalen? Of is er een derde weg? *Expertise en Recht* 2014-1, p. 4-19
- H. Prakken (2016), Een reactie op H.W.J. de Groot: 'Moord op de A73'. *Expertise en Recht* 2016-2, 63-70.
- Sjerps, M. (2011), *Bewijskracht 10, Volle Vaart Recht Vooruit*. Oratie Universiteit van Amsterdam, 4 november 2011.
- Sjerps, M. & Berger, C.E.H. (2011), Het Bayesiaanse model biedt een helder zicht op een complexe werkelijkheid. Den Haag: Nederlands Forensisch Instituut.
- W.C. Thompson & E.L. Schumann (1987), Interpretation of statistical evidence in criminal trials. The prosecutor's fallacy and the defense attorney's fallacy, *Law and Human Behavior* 11: 167-187.
- Timmer, S., Meyer, J.-J.Ch., Prakken, H., Renooij, S. & Verheij, B. (2017). A two-phase method for extracting explanatory arguments from Bayesian networks. *International Journal of Approximate Reasoning* 80 (2017): 475-494.
- Tversky, A. & Kahneman, D. (1974). Judgement under uncertainty: Heuristics and biases. *Science* 185: 1124–1131.
- Verheij, B., Bex, F.J., Vlek, Ch., Timmer, S., Meyer, J.-J.Ch., Renooij, S. & Prakken, H. (2016). Arguments, scenarios and probabilities: connections between three normative frameworks for evidential reasoning. *Law, Probability & Risk* 15: 35-70.
- Ch. Vlek, Prakken, H., Renooij, S. & Verheij, B. (2016), A method for explaining Bayesian networks for legal evidence with scenarios. *Artificial Intelligence and Law* 24: 285-324.
- W.A. Wagenaar & J.H. van der Schrier (1996), Face recognition as a function of distance and illumination: A practical tool for use in the courtroom. *Psychology, Crime, and Law* 2: 321-332.
- Zoete, J.C. de, Sjerps, M.J., Lagnado, D.A. & Fenton, N.E. (2014). Modelling crime linkage with Bayesian networks. *Science & Justice* 56: 209-217.