

Hoe meten weten wordt

INAUGURELE REDE DOOR PROF. DR. MARC H.M. THELEN

*in au
gurele
redo*

Radboud Universiteit



INAUGURELE REDE

PROF. DR. MARC H.M. THELEN



Medische laboratoriumdiagnostiek speelt een rol bij ruim 70 procent van alle medische beslissingen. Toch zijn de kwaliteitseisen die we aan die diagnostiek stellen nog maar zelden gebaseerd op de zekerheid waarmee artsen een bepaalde diagnose willen stellen, tot een bepaalde behandeling willen besluiten of het

succes daarvan willen beoordelen. Door onderzoek te doen naar het effect van de grootte van de meetfout op de zekerheid waarmee een bepaalde medische beslissing wordt genomen, wil Marc Thelen de kwaliteitsspecificaties voor medische laboratoria een onderbouwing geven op basis van hun consequentie. Onderzoek naar de relatie tussen de verschillen in werkwijze en methoden van laboratoria enerzijds en de kwaliteit van resultaten van die laboratoria anderzijds, kan inzicht geven in de wijze waarop die nieuwe specificaties door alle laboratoria kunnen worden gehaald. Academisch onderwijs kan volgende generaties medici en laboratoriumspecialisten opleiden om in onderlinge dialoog de kwaliteit van medische laboratoriumzorg te verbeteren.

Marc Thelen is bijzonder hoogleraar Kwaliteit in de medische laboratoriumzorg aan de Radboud Universiteit/het Radboudumc en werkt daarnaast als klinisch chemicus in STZ ziekenhuis Amphia Breda en als directeur van de Stichting Kwaliteitsbewaking Medische Laboratoriumdiagnostiek (SKML) in Nijmegen. Marc Thelen geniet internationale erkenning voor zijn bijdrage aan kennisontwikkeling van de kwaliteit van medische laboratoriumdiagnostiek en vertaling daarvan in richtlijnen en normen.

Radboud Universiteit



HOE METEN WETEN WORDT

Hoe meten weten wordt

*Rede uitgesproken bij de aanvaarding van het ambt van bijzonder hoogleraar
Kwaliteit in de medische laboratoriumzorg aan de Radboud Universiteit/het Radboudumc op
vrijdag 31 januari 2020*

door prof. dr. Marc H.M. Thelen

Opmaak en productie: Radboud Universiteit, Facilitair Bedrijf, Post & Print
Fotografie omslag: Bert Beelen

© Prof. dr. Marc H.M. Thelen, Nijmegen, 2020

Niets uit deze uitgave mag worden vermenigvuldigd en/of openbaar worden gemaakt middels druk, fotokopie, microfilm, geluidsband of op welke andere wijze dan ook, zonder voorafgaande schriftelijke toestemming van de copyrighthouder.

Mijnheer de rector magnificus, leden van het college van bestuur van de Radboud Universiteit, leden van de raad van bestuur van het Radboudumc, zeer gewaardeerde collega's, beste studenten, familieleden, vrienden en andere toehoorders, met veel plezier verwelkom ik u in mijn openbare les getiteld: Hoe meten weten wordt.

MEDISCHE LABORATORIUMDIAGNOSTIEK

In de medische laboratoriumdiagnostiek worden alleen al in Nederland dagelijks tienduizenden bloedmonsters geanalyseerd. In die monsters wordt de concentratie of activiteit gemeten van honderden verschillende biomarkers. Aan de hand van de uitkomst van die analyses trekken artsen conclusies over mogelijke diagnoses, beslissen ze over de noodzaak van een therapie of aanpassing daarvan en beoordelen ze het succes van die behandelingen. Samen hebben de uitkomsten van medische laboratoriumbepalingen invloed op ruim 70 procent van alle medische beslissingen¹. Vanwege deze enorme impact is het noodzakelijk om voor de juiste medische beslissingen ook te beschikken over de juiste resultaten van medische laboratoriumdiagnostiek. De resultaten van analyses dienen juist genoeg te zijn om de juiste diagnose te stellen of om tot een juiste beslissing te komen over het besluit tot een bepaald type behandeling en om het succes ervan vast te stellen. Als een laboratorium meer dan één apparaat of misschien zelfs meer dan één methode gebruikt voor dezelfde bepaling, dient vast te staan of resultaten van die apparaten of methoden voldoende gelijk zijn om als uitwisselbaar aan de aanvragend artsen te rapporteren. Zo dient ook te worden vastgesteld wanneer de uitkomsten van een ander laboratorium, bijvoorbeeld van een ander ziekenhuis, voldoende op die van het eigen laboratorium lijken om samen in één patiëntendossier naast en door elkaar te kunnen gebruiken.

Om u mee te kunnen nemen in de wereld van de kwaliteit van meten, zal ik eerst met u stil staan bij wat begrippen. Ten eerste het woord meten zelf. Het *Van Dale Groot woordenboek van de Nederlandse taal* zegt: me-ten (mat, heeft gemeten): de lengte, de oppervlakte of de inhoud van iets bepalen.

Daar kunnen we in medische laboratoriumdiagnostiek, maar ook in vele andere wetenschappelijke disciplines niet voldoende mee uit de voeten. Daarom is door het *Bureau International des Poids et Mesures* (BIPM) (in het Engels *international bureau of weights and measures*) - het instituut waar al eeuwen geleden de kilo en de meter werden vastgelegd - een uitleg over meten geformuleerd in bewoordingen die voor alle bètawetenschappen voldoet. Het BIPM heeft een gids gemaakt genaamd *the International Vocabulary of Metrology* (VIM), waarin vele begrippen uit de metrologie worden geduid². Metrologie is de kunst van het meten. VIM zegt over meten: het proces van het experimenteel verkrijgen van één of meer kwantitatieve waarden die redelijkerwijs aan een grootheid kunnen worden toegekend.

Het doen van een meting veronderstelt:

- een beoogd gebruik van het meetresultaat
- een beschreven meetprocedure
- een gekalibreerd meetsysteem volgens de beschreven meetprocedure, rekening houdend met de meetomstandigheden

Een meting kent een:

- ware waarde
- meetnauwkeurigheid
 - meetjuistheid
 - meetprecisie
- meeteenheid
- metrologische herleidbaarheid naar de meeteenheid
- meetonzekerheid

In het vervolg van mijn lezing zal ik de betekenis van deze begrippen bij het beoordelen van de kwaliteit van een meting de revue laten passeren.

METEN IS WETEN

Metten is weten is een veelgehoord adagium. Maar meestal willen we niet weten, maar zeker weten. Wat we dan vaak doen, is dat we iets vaker gaan meten of we gaan onze resultaten vergelijken met die van andermans metingen. Laten we als voorbeeld nemen dat we allemaal met een eigen thermometer de temperatuur gaan meten waarbij Noordzeewater bevriest en dat ook allemaal meer dan eens doen.

Wanneer onze eigen individuele resultaten niet identiek zijn, hebben we in ieder geval aan den lijve ondervonden dat bij herhaalde metingen met eenzelfde meetsysteem variatie bestaat, we noemen dat imprecisie. Wanneer we de uitkomsten van onze eigen metingen middelen en vervolgens dat gemiddelde met dat van anderen vergelijken, ervaren we dat tussen de gemiddelde resultaten van diverse meetsystemen ook verschillen bestaan, die we bias noemen. Wat we eigenlijk zouden willen weten is echter niet de bias ten opzichte van de gemiddelde waarde van alle deelnemers aan dit experiment, maar de bias ten opzichte van de ware waarde. Ik kom daar later op terug.

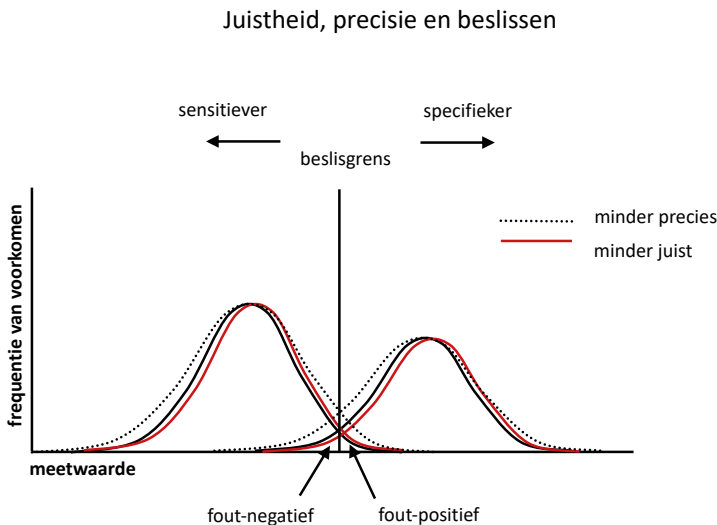
De mate van bias en imprecisie die we accepteren vormen samen de meetonzekerheid. Welke mate van onzekerheid we bereid zijn te accepteren hangt af van het beoogde gebruik van de meting.

BEOOGD GEBRUIK

De meting van de concentratie van eenzelfde stof kan verschillende doelen dienen en de eisen die we aan de kwaliteit van de meting stellen, zou idealerwijs moeten afhangen van dat doel. Zo wordt in de zorg om veel verschillende redenen de concentratie van glucose in bloed gemeten. Indien we met een eenmalige meting willen vaststellen of uitsluiten of iemand *diabetes mellitus* (suikerziekte) heeft, dan stellen we andere eisen aan die meting dan wanneer we in de ambulance willen vaststellen of uitsluiten of iemand bewusteloos is geraakt als gevolg van een *hypoglycemie* (een te lage bloedsuikerconcentratie). Om de noodzakelijke kwaliteit vast te stellen is niet de ernst van afwijking waarnaar onderzoek wordt gedaan bepalend, maar wel de relatie tussen een eventueel verkeerd resultaat en de behandelconsequentie daarvan. Zelfs als de meting in de ambulance een grote afwijking heeft, is de kans groot dat toch juist kan worden ingeschat of er sprake is van een *hypoglycemie* of niet.

DE RELATIE TUSSEN MEETONZEKERHEID EN DE ZEKERHEID VAN DE MEDISCHE BESLISSING

We hebben zojuist gezien dat wanneer de uitkomst van de meting ook met bias en imprecisie niet in de buurt komt van de beslisgrens er geen reden is om onzeker te worden over die beslissing. Er is dan geen overlap tussen de uitkomsten van metingen in de groep patiënten boven de beslisgrens en die eronder. Meestal is dat echter anders; en is er wel overlap (figuur 1).



Figuur 1. Door overlap van meetwaarden tussen patiënten met en zonder een aandoening is noch de sensitiviteit, noch de specificiteit ideaal. Indien juistheid en/of precisie slechter worden, heeft dat effect op de sensitiviteit en de specificiteit.

Door die overlap kunnen metingen in de groep van lage meetresultaten toch boven de beslisgrens uitkomen; dat noemen we fout-positief. Anderzijds kunnen metingen uit de groep van hoge meetresultaten toch onder de beslisgrens uitkomen; dat noemen we fout-negatief. Door de beslisgrens te verleggen, kunnen we test sensitiever of specifiekere maken, maar nooit beide tegelijk.

Als we de imprecisie van alle metingen vergroten zal dat ertoe leiden dat de curves door de extra variatie breder worden. Daardoor neemt zowel het aantal fout-positieve als het aantal fout-negatieve uitslagen toe.

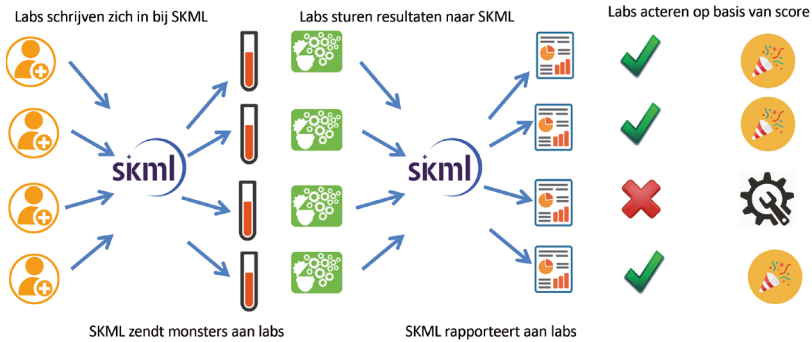
Als we ook nog bias introduceren, zal afhankelijk van de richting van de afwijking, ofwel het aantal fout-positieven, of het aantal fout-negatieven toenemen. U zult begrijpen dat het daarom van belang is om zowel bias als imprecisie te beheersen binnen grenzen die een goede sensitiviteit en specificiteit toelaten. U weet inmiddels ook dat dit gemakkelijker gaat naarmate de resultaten die we willen onderscheiden verder weg liggen van de beslisgrens.

DE ROL VAN DE STICHTING KWALITEITSBEWAKING MEDISCHE LABORATORIUMDIAGNOSTIEK (SKML)

Tot zover de problemen. Zijn er ook oplossingen? Vertrouwen is goed, controle is beter. De SKML heeft er haar missie van gemaakt om de juistheid en daarmee de metrologische traceerbaarheid van meetresultaten van zoveel mogelijk bepalingen in de laboratoriumgeneeskunde te controleren. De laboratoriumgeneeskunde strekt zich uit over verschillende wetenschappelijke disciplines zoals klinische chemie, medische microbiologie, pathologie, medische immunologie en klinische farmacie. Daarbinnen bestaan vaak nog subdisciplines zoals stolling en endocrinologie binnen de klinische chemie of virologie en parasitologie binnen de medische microbiologie. De SKML is een netwerkorganisatie waarbinnen elk van deze subdisciplines een eigen sectie heeft met vertegenwoordigers uit de daarbij betrokken wetenschappelijke verenigingen. De SKML bestaat momenteel uit veertien secties die de kwaliteit van de laboratoria in het betreffende veld toetsen door het organiseren van zogenaamde rondzendingen. We noemen dat een extern kwaliteitsbewakingsprogramma, omdat het buiten de laboratoria wordt georganiseerd.

Daarbij wordt een grote hoeveelheid van een rondzendmateriaal gemaakt dat aan deelnemende laboratoria wordt toegezonden met de instructie om specifieke bepalingen te verrichten en soms ook om op basis van de bevindingen conclusies te trekken en deze naar de SKML te sturen via SKML software (figuur 2).

Wat doet de SKML?



Figuur 2. Schematische weergave van proces van deelname aan externe kwaliteitsbewaking.

De SKML past daar vervolgens statistische analyses op toe en rapporteert aan de deelnemers. Elke deelnemer ziet in zijn eigen rapport hoe de eigen resultaten zich verhouden tot die van andere deelnemers. Daarbij worden statistische analyses gepresenteerd over de verschillen tussen de eigen resultaten en die van anderen, zoals gemiddelden en variatie rond die gemiddelden. Resultaten worden ook gegroepeerd per methode. Zo kunnen deelnemers herkennen in welke mate ze afwijken van andere deelnemers met dezelfde methode. De laboratoria gebruiken de rapportages om te beoordelen of de kwaliteit van de betreffende test voldoende is of dat er correctieve acties nodig zijn.

Om voor dat doel geschikt te zijn moeten de rapportages meer doen dan laboratoria en methoden met elkaar vergelijken. Anders weten de deelnemers immers net zoveel als na ons experiment met temperatuurmetingen van bevroren Noorzeewater: er zijn verschillen en ieder zou gelijk kunnen hebben. De meeste internationale collega-organisaties van de SKML in de externe kwaliteitsbewaking gaan ervan uit dat de gemiddelde waarde na verwijdering van de uitbijters de juiste waarde is. De resultaten van de deelnemers worden dan vergeleken met deze zogenaamde consensuswaarde.

De SKML heeft dat altijd onbevredigend gevonden en heeft er daarom voor gekozen om de waardetoekenning van haar rondzendmonsters zoveel mogelijk met behulp van referentiemethoden in referentielaboratoria te doen. Dat zijn methodes die metrologisch herleidbaar zijn naar de ware waarde. Door dat in monsters te doen die zogenaamd commuteerbaar zijn, zorgt de SKML er bovendien voor dat de gevonden afwijkingen representatief zijn voor patiëntmonsters.

Is daarmee dan alles geregeld? Nee, wat nog niet is gedefinieerd is welke afwijkingen toelaatbaar zijn. Tussen meerdere metingen zullen altijd verschillen bestaan. Het is daarom van belang om vast te stellen welke verschillen (klinisch) relevant of acceptabel zijn. Dat noemen we de analytische prestatiespecificaties.

ANALYTISCHE PRESTATIESPECIFICATIES

In de afgelopen twintig jaar is in de wetenschap van de laboratoriumgeneeskunde consensus ontstaan over de hiërarchie van mechanismen waarmee analytische prestatiespecificaties aan bepalingen worden toegekend. Bovenaan de hiërarchie staat 'medische uitkomsten'. Daarmee bedoelen we dat wanneer er een aangetoonde relatie is tussen de kwaliteit van een meetresultaat en de kwaliteit van een medisch besluit dat op basis van dit meetresultaat wordt genomen, dit bepalend is voor de eisen aan de analytische prestatiespecificaties. Op deze manier zijn de analytische prestatiespecificaties voor *cardiaal troponine*, een test die wordt gebruikt bij de diagnose van een myocardinfarct (hartinfarct), bepaald op basis van studies waarin is onderzocht welke kwaliteit nodig is om geen myocardinfarcten te missen. Omdat dit soort informatie maar zelden beschikbaar is, wordt meestal snel overgestapt op de tweede bron voor analytische prestatiespecificaties: biologische variatie. De gedachte daarbij is dat de uitkomst van een herhaalde meting op een later moment in een ander monster van dezelfde patiënt niet alleen wordt bepaald door verschillen in de meting (bias en imprecisie), maar ook door verschillen die binnen de periode tussen de metingen in de patiënt optreden. We noemen dat de binnenpersoonsvariatie. En dus moeten we bij het vaststellen van de maximaal toelaatbare variatie tussen twee meetmomenten (de imprecisie) rekening houden met deze binnenpersoonsvariatie.

Bij de beoordeling of een meetresultaat afwijkt van een bepaalde beslisgrens moet daarnaast rekening worden gehouden met de tussenpersoonsvariatie. Voor gevallen dat informatie over biologische variatie niet beschikbaar is, geldt zogenaamde *state of the art* als laatste in de hiërarchie voor analytische prestatiespecificaties. Met *state of the art* wordt bedoeld wat met de huidige beschikbare methoden haalbaar is. Deze hiërarchie voor analytische prestatiespecificaties staat bekend als de Milaan-consensus voor analytische prestatiespecificaties en stamt uit 2014³.

De SKML baseert haar analytische prestatiespecificaties ook op deze Milaan-consensus. Op dit moment heeft van alle kwantitatieve bepalingen er slechts één analytische prestatiespecificaties op basis van medisch noodzakelijke kwaliteit (Milaan 1). Dat is de eerder genoemde cardiaal troponine, een bepaling waarvan het resultaat wordt gebruikt voor de diagnostiek van het myocardinfarct. Alle andere bepalingen hebben analytische prestatiespecificaties op basis van biologische variatie of *state of the art*.

OP NAAR MEER MILAAN 1

Wanneer een exacte relatie tussen de meetkwaliteit en medische uitkomsten bekend is, gebruiken we dat als basis voor de analytische prestatiespecificaties, maar zulke kennis is slechts zelden beschikbaar omdat studies naar zo'n relatie moeilijk zijn op te zetten. Daarom bestaat in de Milaan-hiërarchie ook iets dat 1b heet; analytische prestatiespecificaties op basis van wat artsen nodig vinden, zonder harde uitkomstdata, bijvoorbeeld op basis van simulatiestudies. Dat vraagt een uitspraak van medici over de gewenste zekerheid van hun besluit rondom een diagnose, behandeling of andere actie. Zodra we weten welke onzekerheid zij daarin toelaten, kunnen we uitrekenen welke eisen dit stelt aan de analytische kwaliteit. Om die berekening te kunnen maken, is kennis nodig van de verdeling van de meetresultaten ten opzichte van de beslisgrens. Hoe meer mensen een meetresultaat hebben dat ver weg ligt van de beslisgrens, des te groter de meetfout mag zijn voordat dit leidt tot een beslisfout. Ik kan dit illustreren aan de hand van het bevolkingsonderzoek naar darmkanker.

VOORBEELD DARMKANKERSCREENING

De SKML is als één van de partijen betrokken bij dit bevolkingsonderzoek waarbij Nederlanders vanaf 55 jaar wordt gevraagd om een klein beetje van hun ontlasting te laten onderzoeken op de aanwezigheid van bloed. Nu is het zo dat lang niet iedereen met bloedsporen in de ontlasting darmkanker of een voorloper daarvan heeft; het onderzoek is dus niet heel specifiek. Omdat het wel zo is dat vrijwel iedereen met darmkanker of een voorloper daarvan, bloed in zijn ontlasting heeft, is het onderzoek wel sensitief. Door alleen mensen met bloed in hun ontlasting op te roepen voor een darmonderzoek, kan zo'n coloscopie worden beperkt tot de groep mensen met de grootste kans daar baat bij te hebben. Zo'n 6 procent van de deelnemers aan het bevolkingsonderzoek heeft bloed in de ontlasting en wordt opgeroepen voor een coloscopie. Van hen heeft 8 procent darmkanker en 42 procent een gevorderde poliep (figuur 3). Die getallen zouden veel lager liggen als niet zou worden gescreend met de test om bloed in ontlasting aan te tonen. Maar hoe goed moet die test zijn?

De SKML heeft een rol bij de bewaking van de kwaliteit van de betreffende test bij de vier laboratoria die de test voor dit bevolkingsonderzoek in Nederland uitvoeren. De kans op verwijzing mag immers niet afhangen van het moment waarop iemand wordt gescreend of van het laboratorium waar het monster wordt onderzocht. De laboratoria meten daartoe dagelijks kwaliteitscontrolematerialen van de SKML. De SKML bewaakt dat de uitkomsten voldoende constant zijn in de loop van de tijd en dat de verschillen tussen de vier laboratoria binnen vooraf bepaalde grenzen blijven.



Figuur 3. Resultaten bevolkingsonderzoek darmkanker 2014-2017 in vogelvlucht⁴.

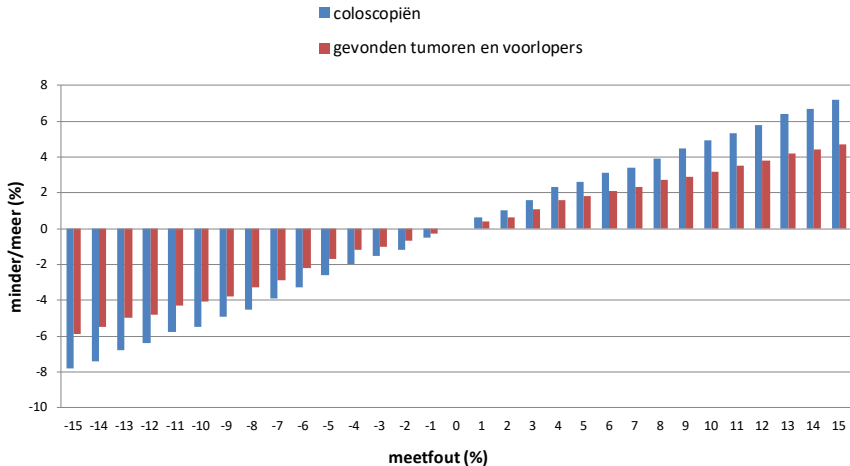
De criteria waarmee de SKML dat momenteel bewaakt, zijn vastgesteld op basis van de eerdere genoemde zogenaamde *state of the art* criteria, oftewel op basis van wat haalbaar is met huidige beschikbare reagentia en apparatuur. Dat komt omdat vooralsnog harde gegevens over de relatie van de test op bloed in feces met medische besliscriteria voor doorverwijzing voor darmonderzoek en de eventuele tumordetectie daarbij ont-

breken. Ook zijn er nog onvoldoende betrouwbare gegevens beschikbaar over de biologische variatie van de hoeveelheid bloed in feces bij gezonde personen. Het zou natuurlijk goed zijn om te weten welke criteria zouden gelden als we die zouden baseren op wat nodig is om de verwijsfout voor coloscopie en daarmee de kans op het onnodig missen van darmtumor te voorkomen. Bovendien is het onnodig uitvoeren van coloscopie bij een persoon met een zeer kleine kans op darmkanker ongewenst. Het leidt niet alleen tot onnodige kosten, maar ook tot risico op complicaties van het onderzoek en in ieder geval tot onnodige bezorgdheid bij de personen die het betreft.

Om te komen tot criteria op basis van een acceptabele verwijsfout, dient deze eerst te worden vastgesteld. Momenteel bereidt het Rijksinstituut voor Volksgezondheid en Milieu (RIVM) daartoe een advies aan de minister voor. Het RIVM doet dat op zijn beurt op basis van de data die tot nu in het bevolkingsonderzoek zijn verkregen. Van de mensen die tussen 2014 en nu zijn gescreeend weten we immers hoe dicht hun uitslag bij de beslislans lag en bij welke verandering van die uitslag er voor hen individueel een ander besluit over verwijzing zou zijn genomen. We weten ook bij welk deel van de mensen die nu wel zijn verwezen, maar bij een subtiel lagere uitslag niet zouden zijn verwezen, er toch een tumor is gevonden. Hiermee kunnen we berekenen welke verhoging of verlaging van een testuitslag tot welke consequenties voor de tumordetectiekans leidt. Andersom kan dan op basis van de gewenste tumordetectiekans worden berekend welke analytische prestatiespecificaties daarvoor nodig zijn. Vervolgens kan de kwaliteitsbewaking zodanig worden ingesteld dat de vier laboratoria die de screening uitvoeren ook inderdaad binnen deze grenzen blijven. Zo gaat deze test voldoen aan het Milaan 1b-criterium waarbij de analytische prestatiespecificaties zijn gebaseerd op een professionele beoordeling van wat medisch nodig is.

Van verschillende percentages laboratoriumvariatie is berekend wat het effect is op het aantal coloscopieën en op het detecteren van tumoren en hun voorlopers. De resultaten laten zien dat een 15 procent te lage uitslag leidt tot ongeveer 8 procent minder verwijzingen en het missen van ongeveer 6 procent van tumoren of voorlopers daarvan (figuur 4).

Op basis van eerdere data weten we ook dat 15 procent hogere uitslagen leiden tot het uitvoeren van ongeveer 7 procent meer coloscopieën en detectie van ongeveer 4,5 procent meer tumoren of voorlopers daarvan. Kortom, een meetfout leidt niet 1:1 tot eenzelfde orde van grootte beslislansfout. Het RIVM zal de minister adviseren over de optimale verwijsfout en de SKML zal dat vervolgens doorvertalen naar de daarbij behorende analytische prestatiespecificaties voor de test op bloed in feces in de vier betrokken laboratoria.



Figuur 4. De relatie tussen grootte van de analytische meetfout van de bepaling van bloed in de feces en de kans op oproep voor coloscopie en detectie van een kwaadaardige afwijking⁵.

BESLISSEN IS SPANNEND

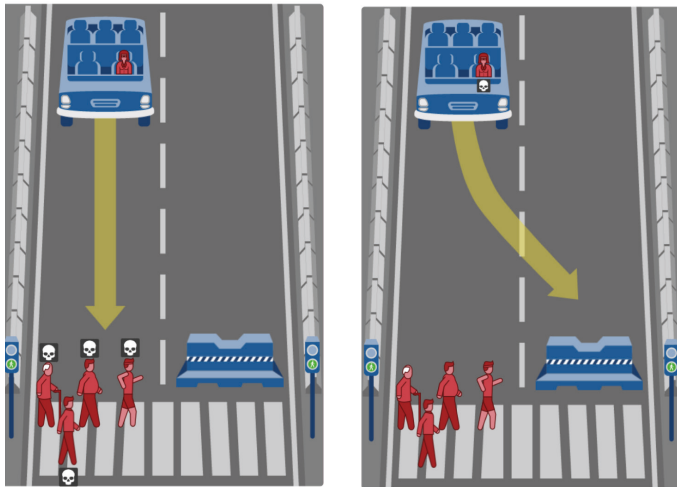
Het beslissen over welke fout aanvaardbaar is, vinden mensen lastig. Uiteraard maken we het liefst helemaal geen fout zodat er in dit geval helemaal geen tumoren worden gemist als gevolg van variatie in de meting van bloed in ontlasting en ook geen mensen onnodig worden onderzocht. Onzekerheid is echter nooit helemaal uit te bannen en hoe kleiner de onzekerheid die we willen accepteren, des te hoger de prijs is die waarvoor moeten betalen. Dat geldt zowel voor de kosten als voor de bijwerkingen die kunnen optreden als gevolg van het onnodig doen van vervolgonderzoek, zoals de coloscopie bij het zojuist genoemde bevolkingsonderzoek. Artsen maken dagelijks keuzes waarbij sensitief en specifiek twee onverenigbare kanten van dezelfde medaille zijn. Als die keuzes onder tijdsdruk worden gemaakt is rechtvaardiging van ongunstige uitkomst achteraf vaak toch goed mogelijk; de arts deed immers naar eer en geweten zijn of haar best en bovendien wordt een beslissing meestal gebaseerd om meer dan slechts één (laboratorium)onderzoek.

Als we willen dat de analytische prestatiespecificaties vaker worden gebaseerd op wat medisch nodig is (conform het Milaan 1-criterium), zullen artsen moeten leren formuleren welke onzekerheden ze bereid zijn te accepteren zodat laboratoriumspecialisten dat kunnen vertalen in analytische prestatiespecificaties voor bepaalde testen. Dat zal voor elk beoogd gebruik van een bepaling tot verschillende analytische prestatiespecificaties leiden.

Dat mensen het moeilijk vinden om een risicoafweging te maken waarbij elke keuze een ander nadeel heeft, is niet uniek voor de geneeskunde. Ook bijvoorbeeld in het verkeer maken we allemaal dagelijks bliksemsnelle afwegingen tussen risico's. Onze afweging wordt gecompliceerder als we dezelfde beslissing moeten nemen zonder tijdsdruk. Dat is bijvoorbeeld nodig bij het programmeren van de algoritmes in zelfrijdende auto's. De software die verantwoordelijk is voor de autonome beslissingen die deze auto's nemen, heeft vooraf natuurlijk input nodig die van mensen moet komen. Dat gaat gepaard met morele dilemma's. Daarbij is een maatschappelijke discussie essentieel om te komen tot draagvlak van deze technologie. Om de maatschappelijke discussie over de morele dilemma's van input te voorzien heeft het Massachusetts Institute of Technology (MIT) een interessante website ontwikkeld met de naam *the moral machine*⁶. Daar kan iedereen die dat wil meedoen om de algoritmes van zelfrijdende auto's te trainen door in bepaalde scenario's te kiezen tussen twee mogelijke aflopen. Daarbij is bewust een confrontatie met de consequentie van de keuze gecreëerd en dat zorgt ervoor dat het maken van de keuze moreel geladen wordt en dus moeilijk is.

Figuur 5 toont een voorbeeld van zo'n scenario. Kies je voor een mogelijk fatale afloop van vier door groen licht overstekende voetgangers (links) of die van jezelf (rechts)?

What should the self-driving car do?



Figuur 5. Voorbeeld uit the Moral Machine⁶.

Als elke keuze onvermijdelijk ongelukkig afloopt, wordt de keuze een stuk ingewikkelder wanneer je die rustig achter je pc of iPad mag maken, dan wanneer je in een fractie van seconde achter het stuur moet beslissen. Zo werkt dat ook voor andere keuzes. Als we de tijd krijgen, willen we zorgen dat er een goede onderbouwing voor de rechtvaardiging van de keuze is.

Hebben zelfrijdende auto's iets met medische laboratoriumdiagnostiek te maken? Nee, maar net als bij het programmeren van zelfrijdende auto's moeten bij het formuleren van de zekerheid waarmee artsen hun beslissingen willen nemen, ongemakkelijke keuzes worden gemaakt. In dit geval over sensitiviteit en specificiteit. Toch is dat nodig om tot analytische prestatiespecificaties te komen op basis van wat medisch nodig is.

Het goede nieuws is dat mensen veel beter zijn in het nemen van beslissingen die ze langzaam en doordacht nemen dan in beslissingen die ze snel en impulsief nemen. Hoogleraar Mariëlle Stoeltinga van deze universiteit refereerde in haar oratie⁷ over kwantitatieve risicoanalyse aan het werk van Nobelprijswinnaar Kahneman⁸. Kahneman beschrijft dat er twee systemen in onze hersenen aan het werk zijn: systeem 1 is snel en automatisch en erg geschikt voor zaken als lokaliseren en het inschatten van gelaatsuitdrukkingen; systeem 2 is langzaam, systematisch en rationeel en veel beter in het inschatten van kansen en risico's. Kortom, ook al vinden we dat razend spannend, we moeten leren vertrouwen op de kracht van onze eigen rationaliteit.

ONDERZOEK NAAR DE OORZAAK VAN DE VERSCHILLEN

We kunnen dus vaker dan tot nu toe de analytische prestatiespecificaties baseren op wat medisch nodig is conform het Milaan 1-criterium. Dan is het natuurlijk interessant om te zien welke laboratoria met behulp van welke methoden nu reeds aan die nieuwe analytische prestatiespecificaties voldoen. Zulke informatie kan worden gehaald uit de data die de SKML heeft van deelnemerresultaten van de afgelopen jaren. In die data kan ook worden bestudeerd wat de gemeenschappelijke factor is van laboratoria die beter of slechter dan gemiddeld presteren.

Eline van der Hagen publiceerde recent een studie op basis van SKML-data naar geschiktheid van de huidige cardiaal-troponinetesten voor de diagnostiek naar myocardinfarcten zonder ST-elevatie (zogenaamde non-STEMIs). Dat zijn hartinfarcten zonder de typische ECG-afwijkingen waaraan een hartinfarct kan worden herkend⁹. Bij deze myocardinfarcten is tijdige herkenning dus volkomen afhankelijk van goede laboratoriumdiagnostiek. In de meest recente richtlijn over de diagnostiek naar non-STEMIs wordt een beslisregel gebruikt om te beoordelen of de troponineconcentratie oploopt tussen het moment van opname van de patiënt en een uur later. Om ook een kleine toename te kunnen aantonen is het natuurlijk van belang dat de analytische variatie,

die er onvermijdelijk altijd is, verwaarloosbaar klein is in vergelijking met het verschil van de troponineconcentratie in het bloed van de patiënt gedurende dat uur.

In de meeste laboratoria zijn meerdere exemplaren van dezelfde meetapparaten aanwezig om te waarborgen dat op elk moment voldoende analysecapaciteit beschikbaar is. Dat geldt zeker voor een belangrijke bepaling als troponine voor de diagnose van een myocardinfarct. De bijdrage van de analytische variatie wordt daarbij natuurlijk groter indien de concentratie na één uur op een ander apparaat wordt gemeten dan die bij opname. Uit Elines onderzoek blijkt dat met de in Nederland meest gebruikte troponinetest voor dit doel 90 procent van de laboratoria een non-STEMI voldoende zeker kan uitsluiten. Dit percentage zakt echter naar ongeveer 80 indien er op meer dan één apparaat wordt gemeten. Om dit te verbeteren is eerst onderzoek nodig naar de verschillen in werkwijze tussen laboratoria die wel en niet voldoen. Dat is één van de onderzoeksdoelen waar deze leerstoel zich op richt. We willen niet alleen weten hoe goed, goed genoeg is, maar ook hoe dat kan worden bereikt. Alleen zo wordt meten immers weten.

Marith van Schroyen Lantman zal als zojuist benoemde eerste promovenda bij deze leerstoel een vervolg geven aan de troponinestudie van Eline en gaan onderzoeken wat goed presterende laboratoria anders doen om vervolgens in interventiestudies te onderzoeken of andere laboratoria inderdaad beter gaan presteren als ze een dergelijke werkwijze implementeren. Daarnaast zijn er vele andere bepalingen waarvan kan worden onderzocht wat de toelaatbare meetfout is op basis van wat dokters nodig vinden. Zo kan ook van die bepalingen, op basis van informatie in de SKML-database, worden onderzocht welke laboratoria aan die criteria voldoen, wat die anders doen dan de andere laboratoria en of implementatie van die werkwijze bij andere laboratoria tot een kwaliteitsverbetering leidt.

HET BELANG VAN ONDERWIJS

Onderzoek levert kennis en inzichten op. Om impact te hebben op de kwaliteit van de medische laboratoriumzorg is het nodig dat die kennis en inzichten worden toegepast in het dagelijks handelen van de mensen die betrokken zijn bij medische laboratoriumdiagnostiek. Die kennisoverdracht moet plaatsvinden in het universitair onderwijs. Ik zal me sterk maken dat de inzichten uit het onderzoek van mijn leerstoel via onderwijs zowel de professionals bereiken in de laboratoria als de artsen die het onderzoek aanvragen en de uitkomsten interpreteren en toepassen.

De professionals in de laboratoria die keuzes maken over de selectie en bewaking van de analytische prestatiespecificaties, zullen moeten leren waarop ze hun keuzes het best kunnen baseren en dat ze daartoe in dialoog moeten met de aanvragers. De aanvragend

artsen zullen daartoe in hun opleiding moeten leren om eisen te formuleren over de zekerheid waarmee ze hun beslissingen willen nemen die op medische laboratoriumdiagnostiek zijn gebaseerd. Alleen zo kan een professionele dialoog ontstaan tussen zorgprofessionals die samen tot keuzes komen over de meetkwaliteit die optimaal is voor de beslissing die ermee genomen gaat worden.

Dit betekent dat zowel studenten Geneeskunde als studenten Biomedische Wetenschappen van de Radboud Universiteit in hun opleiding vertrouwd zullen worden gemaakt met relatie tussen de kwaliteit van meten en de kwaliteit van beslissen en leren wat hun professionele verantwoordelijkheid is om daar op een zinvolle manier met elkaar de dialoog over aan te gaan.

ZO WORDT METEN WETEN

Ik heb betoogd dat de norm vaker dan nu kan worden bepaald door wat medisch nodig is. Met onderzoek binnen mijn leeropdracht zal voor specifieke combinaties tussen medische beslissingen en bepaalde laboratoriumbepalingen worden uitgezocht bij welke analytische prestatiespecificaties van de laboratoriumtesten de medische beslissing optimaal wordt ondersteund. Op basis van onderzoek kan ook inzicht ontstaan op welke wijze de noodzakelijke analytische prestatiespecificaties het beste kunnen worden behaald. Op basis van onderwijs kunnen de met onderzoek verkregen kennis en inzichten worden overgebracht op zowel artsen als laboratoriumprofessionals.

ZO ZAL DE SKML VAN METEN WETEN MAKEN

De kennis en inzichten over de noodzakelijke analytische prestatiespecificaties zullen ook impact hebben op het beleid van de SKML. Ook bij de externe kwaliteitscontrole die de SKML met haar rondzendingen onderzoekt, zal de norm vaker moeten worden bepaald door wat medisch nodig is.

Om dat proces richting te geven zal de SKML haar verantwoordelijkheid nemen door een aantal daartoe essentiële stappen te zetten waar ik met deze leerstoel een belangrijke impuls aan zal geven:

- analytische prestatiespecificaties formuleren per testtoepassing in plaats van per test;
- de specificaties baseren op wat nodig is voor medische beslissingen die met die test worden genomen;
- in de SKML-database onderzoeken welke laboratoria met welke methoden die specificaties halen;
- onderzoeken met welke aanpassingen de laboratoria en methoden die nu nog niet voldoen, wel kunnen voldoen;

- in onderwijs uitdragen hoe deze inzichten kunnen worden verankerd in optimalisatie van de medische laboratoriumdiagnostiek;
- de daarvoor noodzakelijke werkwijzen vastleggen in richtlijnen en veldnormen.

Ik zie het als mijn verantwoordelijkheid om vanuit deze leerstoel leiding te geven aan deze ontwikkelingen. Want pas als deze zaken zijn opgelost, kan met gerust hart worden gezegd; meten is weten.

WAAROM DIT VOOR IEDEEREEN RELEVANT IS

Ik heb uiteengezet hoe de kwaliteit van medisch laboratoriumonderzoek beter kan worden afgestemd op de kwaliteit van medische beslissingen die artsen mede op basis van die laboratoriumbepalingen nemen. Met onderzoek kunnen kennis en inzichten worden verkregen die met onderwijs kunnen worden overgebracht en in richtlijnen kunnen worden verankerd. Onderzoek en onderwijs richten zich tot de doelgroepen artsen en laboratoriumprofessionals, maar de consequenties hebben impact voor een veel grotere doelgroep namelijk de gehele maatschappij.

Zelfs wie zelf helemaal geen gezondheidszorg gebruikt, heeft er toch mee te maken. Al was het maar omdat we allemaal bijdragen aan de kosten die inmiddels 13 procent van ons bruto nationaal product beslaan¹⁰. Optimalisatie van de kwaliteit van medische laboratoriumdiagnostiek kan helpen de kwaliteit van zorg te optimaliseren, maar ook de kosten die daarmee gemoeid zijn te beheersen.

Als de kwaliteit van medische laboratoriumdiagnostiek wordt gebaseerd op de zekerheid waarmee artsen een diagnose willen stellen, kan dat niet alleen leiden tot een snellere route naar de juiste diagnose, maar worden ook kosten voorkomen om via een andere weg bij de juiste diagnose te komen.

Als de kwaliteit van medische laboratoriumdiagnostiek wordt gebaseerd op de zekerheid waarmee een arts kiest voor een bepaalde ingreep en het beoordelen van het succes daarvan, dan leidt dat niet alleen tot een grotere kans op een juiste behandeling, maar voorkomt het ook het herstellen van een eventuele verkeerde keuze en de kosten daarvan.

Als de kwaliteit van medische laboratoriumdiagnostiek wordt gebaseerd op de mogelijkheid om data van verschillende laboratoria te mogen combineren, kan dat ertoe bijdragen dat laboratoriumonderzoek bij verwijzing niet wordt herhaald. Dat spaart niet alleen tijd, maar ook geld en wat tegenwoordig ook steeds meer relevant is: het voorkomt verspilling van grondstoffen en energie.

Ik hoop dat ik u vanmiddag iets meer inzicht heb kunnen geven in wat mij bezig houdt bij de kwaliteit van meten en waarom onderzoek en onderwijs nodig zijn om die kwaliteit zinvoller te maken. Door optimalisatie van de kwaliteit van medische laboratoriumdiagnostiek, maken we deze diagnostiek doeltreffender, doelmatiger en duurzamer. Dat betekent voor de maatschappij een bijdrage aan betere zorg, tegen lagere kosten en met minder verspilling.

DANKWOORD

Nu ik aan het einde van mijn rede ben gekomen, wil ik graag enkele woorden van dank uitspreken.

Ik dank het college van bestuur van deze universiteit en de raad van bestuur van het Radboudumc voor het in mij gestelde vertrouwen. Ik voel de verantwoordelijkheid om deze leeropdracht tot een succes te maken.

Graag wil ik ook de mensen bedanken die zich hebben beijverd voor deze leerstoel. Samen met Hans Willems, de voorzitter van de raad van advies van de SKML, heeft met name SKML-voorzitter Fred Sweep veel energie gestoken in het tot stand komen van deze leerstoel. Ik dank ook Dorine Swinkels voor haar steun en coaching.

Dan zijn er altijd voorbeeldrollen; mensen die mij hebben geïnspireerd om te werken aan de wetenschappelijke ontwikkeling van de SKML in dienst van zingevende medische laboratoriumdiagnostiek. Graag dank ik Christa Cobbaert en Cas Weykamp; Christa voor de lessen in volharding en energie, Cas voor zijn grote talent om alles wat moeilijk is gemakkelijk te maken.

Dan wil ik de mensen danken die mij hebben gevormd en opgeleid en me het vertrouwen hebben gegeven om deze verantwoordelijkheid aan te gaan. Mijn vader bracht mij de kracht van vlijt bij, van mijn moeder leerde ik hoe belangrijk belangstelling en aandacht zijn.

In mijn professionele ontwikkeling hebben drie mensen een bijdrage geleverd waardoor ik ze voor altijd als vriend zal koesteren. Dop Simonides was mijn copromotor en heeft mij het plezier van wetenschappelijk denken en schrijven bijgebracht en plezier hebben we nog steeds samen. Rob Jansen beschouw ik als visionair in de klinische chemie. Zijn lezing over kalibratie 2000 in datzelfde jaar heeft een verlangen in mij aangewakkerd om die visie mede te gaan waarmaken en gelukkig heb ik dat vaak en lang met Rob samen mogen doen. Rob is mijn vader in de klinische chemie en gids op het SKML-pad en als deze leerstoel er eerder was gekomen, was het terecht de zijne geweest. Daarom draag ik deze leerstoel aan je op Rob. Ik ben vereerd hier te mogen staan in de toga van je in 2018 overleden vader Ad Jansen, hoogleraar Klinische Chemie aan de Radboud Universiteit en grondlegger van de SKML. Aan Hans Meij, mijn eerste directeur in Amphia, dank ik mijn vertrouwen dat optimisme niet naïef is, maar een levenskunst waarmee je ook anderen plezier kunt geven.

Ook wil ik de mensen danken die het mogelijk maken dat ik een dergelijke functie kan vervullen als nevenfunctie en daarmee accepteren dat er altijd minder tijd is voor elkaar

dan we zouden willen. Dat betreft in ieder geval de medewerkers van het centraal facilitair bureau (CFB) van de SKML. Gelukkig denken jullie allemaal hard mee hoe we de beschikbare interactietijd zo doeltreffend en gezellig mogelijk kunnen inrichten.

Twee mensen bij de SKML verdienen een speciale vernoeming: Helma Rikken is als hoofd van het secretariaat tevens het hart en het geweten van CFB en een belangrijke factor voor het werkplezier van iedereen die daar werkt. Fred Sweep heb ik als SKML-voorzitter al eerder genoemd. Dank je voor je warme empathische betrokkenheid. Ondanks al je klussen heb je altijd tijd om met aandacht voor mens en detail iets waardevols bij te dragen.

Ik dank de leden van het SKML-bestuur en de SKML-secties voor het enthousiasme waarmee jullie participeren in projecten om zo de SKML internationaal voorop te laten blijven lopen. De SKML is daardoor meer dan kwaliteitsbewaking; de SKML wordt terecht gezien als gids in harmonisatie en standaardisatie van medische laboratoriumdiagnostiek.

Op het klinisch chemisch laboratorium van Amphia dank ik op de eerste plaats mijn collega vakgroepleden omdat jullie me de ruimte geven om tijd te besteden aan activiteiten die niet altijd onmiddellijk iets opleveren voor ons eigen laboratorium. Mijn frequente afwezigheid vraagt ook van anderen veel. Gelukkig heeft bedrijfsmanager Erna Verstraten niet veel tijd nodig om samen veel te bereiken.

De raad van bestuur van Amphia dank ik voor het vertrouwen en de ruimte die ik krijg om me ook buiten Amphia in te zetten voor kwaliteitsdoelen waar we samen trots op kunnen zijn. In dat kader dank ik jullie ook voor het mogelijk maken van de aanstelling van mijn eerste promovendus. Marith zal ook jullie trots maken.

Thuis

In het dankwoord van mijn proefschrift heb ik mijn familie en vrienden bedankt dat ze de verwaarlozing niet wederzijds maakten. Ik kan het ook nu niet beter zeggen. Toen beloofde ik nog beterschap. Nu weten jullie beter. Dank dat het altijd weer gezellig is als we elkaar zien.

Tenslotte mijn echte thuis

Eleanor en Fiona, jullie wonen al niet meer thuis, maar voelen je er gelukkig wel nog thuis. Jullie ontwikkeling vol zelfbewuste keuzes en plezier maakt ons als ouders gelukkig en trots. Lieve Anne, wat heb ik veel aan jou; je bent mijn rustpunt en bron van energie tegelijk. Bij jou laad ik op.

Ik heb gezegd.

REFERENTIES

- 1 Forsman RW, Why is the laboratory an afterthought for managed care organizations? *Clin Chem* 1996;42:813-6
- 2 BIPM, IEC, IFCC, ILAC, IUPAC, IUPAP, ISO, OIML The international vocabulary of metrology—basic and general concepts and associated terms (VIM), 3rd edn. JCGM 200:2012. <http://www.bipm.org/vim>
- 3 Sandberg S, Fraser FG, Horvath AR, Jansen R, Jones G, Oosterhuis W, et al. Defining analytical performance specifications: consensus statement from the 1st Strategic Conference of the European Federation of Clinical Chemistry and Laboratory Medicine. *Clin Chem Lab Med* 2015;53:833-5
- 4 RIVM infographic Bevolkingsonderzoekdarmkanker 2014-2017, zoals uitgevoerd door de vijf samenwerkende regionale screeningsorganisaties van bevolkingsonderzoek (BVO) voor vroege opsporing van kanker.
- 5 Op basis van berekeningen van Lansdorp-Vogelaar I, Afdeling maatschappelijke gezondheidszorg (MGZ), Erasmus MC, Universitair Medisch Centrum Rotterdam, in werkgroep 2 kwaliteitsborging fecal immunochemical test (FIT) waarin RIVM, BVO, SKML, MGZ en de klinisch chemici van de screeningslaboratoria samenwerken voor optimalisatie van de kwaliteitsbewaking van de FIT binnen het Nederlandse bevolkingsonderzoek darmkanker.
- 6 Awad E, Dsouza S, Kim R, Schulz J, Henrich J, Shariff A, Bonnefon J-F, Rahwan L, The Moral Machine experiment, *Nature* 2018; 563:59-64, <http://moralmachine.mit.edu>
- 7 Stoelinga M, A certain risk: quantitative risk management for high-tech systems, inaugurele rede, Radboud Universiteit, Nijmegen, 2018
- 8 Kahneman D. A perspective on judgement and choice: mapping bounded rationality. *American Psychologist* 2003, 58: 697-720
- 9 van der Hagen E, Cobbaert C, Meijer R, Thelen M, Fast 0/1-h algorithm for detection of NSTEMI: are current high-sensitivity cardiac troponin assays fit for purpose? An EQA-based evaluation. *Clin Chem Lab Med*. 2019;57:1999-2007
- 10 Centraal Bureau voor de Statistiek, Statline: <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/84047NED/table?dl=1D372>

