

De vele gezichten van wetenschappelijke integriteit

Citation for published version (APA):

Penders, B. (2016). De vele gezichten van wetenschappelijke integriteit. *Podium voor Bio-ethiek*, 23(3), 4-7. <https://nvbioethiek.files.wordpress.com/2017/01/podium-16-3-tbv-website.pdf>

Document status and date:

Published: 01/01/2016

Document Version:

Publisher's PDF, also known as Version of record

Document license:

Taverne

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Podium voor Bio-ethiek

Thema: Kwaliteit en integriteit van medisch wetenschappelijk onderzoek

De vele gezichten van wetenschappelijke integriteit

Bart Penders

De invloed van de financieringsbron op de kwaliteit en integriteit van klinisch geneesmiddelenonderzoek

Henk Jan Out

Ontwikkeling van nieuwe geneesmiddelen betekent niet automatisch een verbetering van de kwaliteit van de kindergeneeskunde. Pleidooi voor een Learning Healthcare System

Martine de Vries

Betere kwaliteit en vertaalbaarheid van pre-klinisch wetenschappelijk onderzoek noodzakelijk. Een pleidooi voor een expliciete wereldwijde doelstelling

Merel Ritskes-Hoitinga en Harald Schmidt

Van statistische significantie naar klinische relevantie

Fenneke Blom en Judith J.M. Rijnhart

Toegankelijkheid en hergebruik van medisch-wetenschappelijke onderzoeksdata

Heleen van Luijn

De rol van de CCMO en Medische Ethische Toetsingscommissies bij het bewaken van kwaliteit en integriteit van medisch wetenschappelijk onderzoek

Joop van Gerven

Enkele problemen en dilemma's van METC's bij wetenschappelijke (en dus ethische) toetsing

Rien Janssens en Frits Lekkerkerker

Investeer in het toezicht op de uitvoering van onderzoek!

Jacqueline van Oijen, Kor Grit en Roland Bal

Colofon

De NVBe streeft naar (1) stimulering van de bio-ethiek (humane, dier- en natuurethiek) in relevante sectoren, (2) contacten tussen vertegenwoordigers uit verschillende vakgebieden, instellingen en organisaties die betrokken zijn bij bio-ethische kwesties, (3) erkenning van de waarden van een open discussie over bio-ethische problemen in wetenschap en samenleving, (4) presentatie van discussies in de bio-ethiek in Nederland.

Het Podium voor Bio-ethiek (voorheen de Nieuwsbrief) van de vereniging draagt bij aan deze doelen door publicatie van bio-ethisch nieuws (van binnen en buiten de vereniging) en bondige, voor een breed publiek toegankelijke, interdisciplinaire bijdragen over bio-ethische kwesties.

Het Podium voor Bio-ethiek verschijnt vier keer per jaar en wordt toegezonden aan leden van de NVBe. Het Podium voor Bio-ethiek, mededelingen uit de Vereniging en bio-ethische informatie verschijnen ook op www.nvbe.nl

Redactieadres

Secretariaat NVBe
info@NVBe.nl, t.a.v. Podium-redactie

Redactie

drs. Carla Bal, drs. Beatrijs Haverkamp, dr. Eric van de Laar,
dr. Heleen van Luijn, dr. Hanneke van der Meide,
dr. Niels Nijsingh, mr. drs. Virgil Rerimassie,
dr. Lieke van der Scheer.

Website

dr. Clemens Driessen

Opmaak

drs. Ger Palmboom.

Instructie voor bijdragen

Bijdragen in overeenstemming met de doelstelling van Het Podium voor Bio-ethiek zijn van harte welkom. Voor suggesties en vragen kunt u zich wenden tot de redactie via het e-mailadres. Artikelen bij voorkeur rond de 1500 woorden, boekbesprekingen en verslagen van congressen, conferenties, etc. maximaal 500 woorden.

Bij voorkeur geen uitgebreide literatuurverwijzingen. Bijdragen kunt u per e-mail sturen naar het redactieadres.

De redactie behoudt zich het recht voor bijdragen te weigeren of in te korten.

Bestuur NVBe

dr. Henk van den Belt, drs. Katja ten Cate,
dr. Clemens Driessen, dr. Anke Oerlemans (penningmeester),
dr. Lieke van der Scheer voorzitter a.i.,
dr. Guus Timmerman (secretaris), dr. Margreet Stolper.

WWW.NVBe.NL

Lid worden?

Iedereen die op een of andere manier (op academisch niveau) betrokken is bij de levenswetenschappen en de ethische reflectie daarop, kan lid worden van de Nederlandse Vereniging voor Bio-ethiek (NVBe). Op de website www.nvbe.nl (doorklikken naar 'Lidmaatschap') vindt u een formulier waarmee u zich kunt aanmelden als lid.

Neem contact op met het secretariaat:

NVBe Ledenadministratie
Verenigingen Beheer Nederland B.V.
Spinozalaan33
2273 XC VOORBURG
E-mail: ledenadministratie@nvbe.nl
Telefoon: +31 (0)70 456456
(maandag t/m vrijdag 9:00 uur t/m 17:00 uur)

Inhoudsopgave

Thema: Kwaliteit en integriteit van medisch wetenschappelijk onderzoek

- 2 Inleiding**
Eric van de Laar, Heleen van Luijn, Krista Tromp
- 4 De vele gezichten van wetenschappelijke integriteit**
Bart Penders
- 7 De invloed van de financieringsbron op de kwaliteit en integriteit van klinisch geneesmiddelenonderzoek**
Henk Jan Out
- 11 Ontwikkeling van nieuwe geneesmiddelen betekent niet automatisch een verbetering van de kwaliteit van de kindergeneeskunde. Pleidooi voor een Learning Healthcare System**
Martine de Vries
- 14 Betere kwaliteit en vertaalbaarheid van pre-klinisch wetenschappelijk onderzoek noodzakelijk. Een pleidooi voor een expliciete wereldwijde doelstelling**
Merel Ritskes-Hoitinga en Harald Schmidt
- 18 Van statistische significantie naar klinische relevantie**
Fenneke Blom en Judith J.M. Rijnhart
- 21 Toegankelijkheid en hergebruik van medisch-wetenschappelijke onderzoeksdata**
Heleen van Luijn
- 24 De rol van de CCMO en Medische Ethische Toetsingscommissies bij het bewaken van kwaliteit en integriteit van medisch wetenschappelijk onderzoek**
Joop van Gerven

- 28 Enkele problemen en dilemma's van METC's bij wetenschappelijke (en dus ethische) toetsing**

Rien Janssens en Frits Lekkerkerker

- 32 Investeer in het toezicht op de uitvoering van onderzoek!**

Jacqueline van Oijen, Kor Grit en Roland Bal

Uit de Vereniging

Lid worden?

U bent van harte welkom als lid van de NVBe. Op de website www.nvbe.nl (doorklikken naar 'Lidmaatschap') vindt u een formulier waarmee u zich kunt aanmelden als lid.

Na aanmelding zult u een rekening ontvangen voor de contributie. Als uw betaling binnen is, wordt uw lidmaatschap definitief en zult u het *Podium voor Bio-ethiek* en de uitnodigingen voor NVBe-activiteiten ontvangen.

De jaarlijkse contributie voor individuele leden bedraagt € 42,50. AIO's en studenten betalen € 25. Voor instituten kost het lidmaatschap jaarlijks € 175.

Voordelen?

1. Deelname aan nationaal interdisciplinair bio-ethisch netwerk.
2. Driemaandelijks gratis ontvangst van het Podium voor bio-ethiek.
3. Uitnodiging en gratis toegang tot het NVBe-jaarsymposium en de onderwijsmiddag.
4. Gratis ontvangst jaarlijks gepubliceerde pre-advies.

Als welkomstcadeau ontvangt u – zolang de voorraad strekt – een exemplaar van het boek *Ethiek in discussie. Praktijkvoorbeelden van ethische expertise*.

Thema: Kwaliteit en integriteit van medisch-wetenschappelijk onderzoek

Inleiding

Medisch-wetenschappelijk onderzoek met mensen is een noodzakelijke maar ook delicate aangelegenheid. Noodzakelijk, omdat de geneeskunde niet verder komt zonder onderzoek, delicaat omdat je mensen, patiënten vaak, vraagt om nieuwe geneesmiddelen uit te proberen of experimentele behandelwijzen te ondergaan, waarvan nog niet vaststaat of ze beter zijn dan de bestaande en of ze wel veilig zijn. Deelname aan onderzoek is dus niet altijd zonder risico. Vandaar dat onderzoeken, alvorens ze uitgevoerd mogen worden, grondig getoetst worden op de verhouding tussen verwachte opbrengst aan wetenschappelijk inzicht of therapeutisch effect, en risico en belasting voor de proefpersoon. Vooraleer onderzoekers aan de slag mogen moet er een officiële goedkeuring zijn door een medisch-wetenschappelijke toetsingscommissie (METC) op basis van de Wet medisch-wetenschappelijk onderzoek met mensen (WMO). Gezien deze wettelijke regulering en de toetsing zou je verwachten dat kwaliteit en integriteit voldoende geborgd zijn.

Toch is hier veel discussie over. Met regelmaat worden in de media ernstige zorgen geuit over de kwaliteit van onderzoek, soms zelfs met fikse uithalen naar de integriteit van de uitvoerders. Bert Leufkens, de voorzitter van het College ter beoordeling van Geneesmiddelen, stelde in de NRC van 20 april dit jaar dat een groot deel van de geneesmiddelen studies niet deugt. Deze mening is ook hoogleraar farmacologie Harald Schmidt toegedaan. De Britse arts Ben Goldacre heeft zijn boeken *Bad Science* en *Bad Pharma* laten zien dat zakelijke en commerciële belangen een grote bedreiging vormen voor kwaliteit en integriteit. Het verst in deze kritiek gaat de Deense onderzoeker en medeoprichter van de Cochrane Collaboration Peter Gøtzsche. Volgens hem is alle medisch onderzoek gecorrumpeerd door de financiële belangen van de farmacie. Met de titel van zijn boek *Dodelijke medicijnen en georganiseerde misdaad*, laat hij er geen misverstand over bestaan wat dat volgens hem betekent: stelselmatige en welbewuste manipulatie van onderzoek met desastreuze gevolgen voor patiënten. Hij schrijft

er niet voor terug om de farmacie te vergelijken met de Maffia.

Voor de redactie van het Podium voor Bio-ethiek waren deze kritische uitingen door gezaghebbende insiders reden om een themanummer aan kwaliteit en integriteit van medisch-wetenschappelijk onderzoek te wijden. Het riep bij ons de volgende vragen op: Hoe moeten bovenstaande negatieve oordelen over kwaliteit en integriteit verstaan worden?; Is echt een groot deel van het medisch en farmacologisch onderzoek ondeugdelijk? Kan dat op methodologische gronden beweerd worden of liggen er andere (politieke, ideologische, wetenschapsfilosofische) motieven ten grondslag aan dergelijke oordelen?; Indien het zo gesteld is met de kwaliteit, doen de toetsingscommissies hun werk dan wel goed?; Hebben we te maken met individuele excessen of zitten er structurele fouten in dit systeem?; Kan er iets verbeterd worden? En wie is dan bij machte die verandering teweeg te brengen?

De inzendingen die we hebben gekregen op deze vragen belichten zeer uiteenlopende problemen die bij kwaliteit en integriteit van onderzoek in het geding kunnen zijn, en laten ook zien hoe breed en divers de praktijk van het medisch-wetenschappelijk onderzoek is. Dat leidt dan allereerst natuurlijk tot de vraag: 'wat is kwaliteit van onderzoek eigenlijk?' In zijn bijdrage laat Bart Penders zien dat de visie op kwaliteit niet eenduidig is onder onderzoekers. Niet iedereen verstaat er hetzelfde onder en ook per onderzoeksdomein kan de invulling verschillen. Zodoende moet je misschien ook niet een eenduidige norm opleggen voor heel de waaier aan verschillende onderzoeken die eronder kunnen vallen.

Op de kritieken aan het adres van de farmaceuten en hun belangen, heeft Henk Jan Out een reactie geschreven. Voor de ontwikkeling van nieuwe geneesmiddelen kunnen we niet zonder de farmacie, zij betalen het grootste deel, en willen hun investering terugverdienen. De kwaliteit van de studies van de industrie zijn veel beter dan die uit de academie, en uit de daling van de omvang van boetes die de industrie krijgt opgelegd, maakt Out op dat men de laatste jaren behoedzamer wordt in de marketing van de uitkomsten.

Martine de Vries gaat in op het probleem dat de strenge eisen om onderzoek met kinderen te kunnen doen, ertoe geleid hebben dat veel behandelingen

voor kinderen een wetenschappelijke onderbouwing ontberen. Daardoor is niet duidelijk of kinderen optimale medische zorg krijgen. Met alleen meer wetenschappelijk onderzoek kan dit probleem niet getackeld worden, volgens De Vries, er moet ook gekeken worden of die kennis met andere methoden uit de kliniek gehaald kan worden.

Merel Ristskes Hoitinga en Harald Schmidt vragen in hun bijdrage aandacht voor de kwaliteitskloof tussen pre-klinisch (bijvoorbeeld onderzoek met proefdieren) en klinisch onderzoek (bijvoorbeeld met patiënten). Er zitten veel prikkels in het wetenschappelijk systeem die kwaliteitsondermijnend zijn, zoals de eis om zo veel mogelijk te publiceren in high impact tijdschriften. Gevolgen zijn ondermaatse kennis, slechte kennisdeling en overbodig (vervolg)onderzoek.

Heleen van Luijn vraagt eveneens aandacht voor het probleem van gebrekkige kennisdeling. Zij schrijft dat toe aan de manier waarop ruwe data worden opgeslagen. Dat gebeurt namelijk niet op een centrale plek waardoor de data voor andere onderzoekers voor hergebruik toegankelijk zouden zijn. Toegang en hergebruik zouden volgens haar leiden tot kwaliteitsverbetering, minder fraude en nuttigere besteding van schaars onderzoeksgeld. De manier waarop statistiek geïnterpreteerd wordt is de oorzaak van weer een ander kwaliteitsprobleem. Wetenschappelijke tijdschriften publiceren liefst artikelen met positieve, statistische significante resultaten.

Fenneke Blom en Judith Rijnhart leggen in hun bijdrage uit wat er problematisch is aan het criterium 'statistische significantie'. Klinische relevantie zou volgens hun een veel beter criterium zijn om kwaliteit van onderzoek te taxeren.

Ook op het vlak van de beoordeling van medisch-wetenschappelijk onderzoek is er aandacht voor kritiek en wordt er nagedacht over verbetering. Joop van Gerven beschrijft de positie van de Centrale Commissie Mensgebonden Onderzoek (CCMO) als uitvoerder van de WMO en de taken van METC's. Controle en regulering moeten volgens Van Gerven vooral gericht zijn op het bevorderen van de kernwaarden van verantwoord onderzoek: wetenschappelijke kwaliteit, betrouwbaarheid van de resultaten, respect voor belangen van de patiënt/proefpersoon en publieke verantwoording. Hij noemt diverse maatschappelijke,

medische en persoonlijke dilemma's die deze waarden onder druk zetten.

Rien Janssens en Frits Lekkerkerker, beide bestuursleden van de Nederlandse Vereniging van METC's (NVMETC) wijzen erop dat vooraf lang niet altijd goed in te schatten is hoe reëel de verwachte opbrengsten zijn en hoe deze zich verhouden tot de risico's en belasting. Zij denken ook dat een METC gedurende de uitvoering van onderzoek een monitorende taak moet uitoefenen en van daaruit onderzoekers ertoe kan bewegen om ook negatieve resultaten te publiceren.

Ook Jacqueline van Oijen, Kor Grit en Roland Bal bepleiten een actievere rol voor METC's gedurende de uitvoering van onderzoek. In hun bijdrage beschrijven zij wat het huidige toezicht behelst, wat de onvolkomenheden daarvan zijn en welke oplossingen hen voor ogen staan.

Alles bij elkaar genomen geven de artikelen een goed beeld van wat er speelt rond kwaliteit en integriteit bij medisch-wetenschappelijk onderzoek. Een aantal zaken springen in het oog, vooral de zorg over het niet publiceren van negatieve uitkomsten. Dit lijkt de achilleshiel van het systeem dat door de fixatie op positieve resultaten de kwaliteit van onderzoek structureel dreigt te ondermijnen. Een daarmee samenhangend punt is dat de belangen groot zijn, en dat niet alleen voor de farmaceutische bedrijven maar feitelijk voor elke afzonderlijke participant die in deze sector een wetenschappelijke carrière probeert op te bouwen. De kans dat substantiële veranderingen van binnenuit komen is daarmee niet groot, aangezien zij die afwijken van het gebaande pad waarin positieve resultaten, financiële belangen en carrièremogelijkheden zo nauw aan elkaar geklonken zijn, direct zelf de gevolgen zullen ondervinden. Toezichthoudende organen als de CCMO en METC's hebben een betere positie daarvoor, maar ook de wetenschappelijke tijdschriften zouden hier een uitdrukkelijker rol in moeten spelen. Want uiteindelijk zijn niet de financiële belangen en de carrièrekansen het hoogste goed in deze sector, maar de bescherming van proefpersonen en de gezondheid en het welzijn patiënten.

We wensen u veel leesplezier.

*Eric van de Laar, Heleen van Luijn, Krista Tromp
Themaredactie*

Thema: Kwaliteit en integriteit van medisch-wetenschappelijk onderzoek

De vele gezichten van wetenschappelijke integriteit

Bart Penders

Wetenschappelijke integriteit is geen set uniforme principes die boven de wetenschappelijke praktijk staat. Ze maakt onderdeel uit van de worstelingen op de wetenschappelijke werkvloer. Wetenschappelijke integriteit kent evenveel versies als wetenschap zelf en vooral in interdisciplinaire samenwerkingen kunnen deze smaakverschillen leiden tot het onnodig in twijfel trekken van de integriteit en geloofwaardigheid van wetenschappers en wetenschappelijke beweringen.

Op de zogenaamde *Retraction Watch Leaderboard* staat Diederik Stapel op een niet-zo-eervolle derde plek. *Retractionwatch.com* is een wetenschappelijk blog gewijd aan het documenteren van terugtrekkingen van wetenschappelijke artikelen. Een terugtrekking kan plaatsvinden nadat er fouten zijn ontdekt in een wetenschappelijke publicatie, waardoor de primaire claim van de tekst wordt ondermijnd en uit de literatuur gefilterd dient te worden. Zo'n fout kan een onschuldige, technische vergissing zijn, of, zoals in het geval van de 58 teruggetrokken artikelen van Diederik Stapel, het gevolg van een inbreuk op de wetenschappelijke integriteit. Bovenaan het *Retraction Watch Leaderboard* staat de Japanse anesthesioloog Yoshitaka Fujii, met momenteel 183 terugtrekkingen. Blogs zoals *retractionwatch.com* en analyses van bekende fraudezaken, waaronder Diederik Stapel, Hendrik Schön of Hwang Woo-Suk (die, hoe bekend ook, niet op het *Leaderboard* staat) zijn smakelijk leesvoer en voeden onze sensatiezucht waar het wetenschappelijke misstappen betreft.

Zeker, het identificeren van zulke extreme fraude en andere misstappen is van groot belang, en deze za-

ken hebben een belangrijke voorbeeldfunctie in het debat over wetenschappelijke integriteit, zowel onder wetenschappers als in het onderwijs. Toch zijn deze excessen uitzonderlijk. Wat de cases op het *Retraction Watch Leaderboard* met elkaar delen is een zekere helderheid over goed en fout. Waar het wetenschappelijk veelplegers betreft, gaat het vaak over de drie grote wetenschappelijk doodzonden: plagiaat, falsificatie of fabricatie. De zichtbaarheid van de overtreding is groot, het debat over of er wel of niet een overtreding is begaan is relatief eenvoudig en er is consensus over morele waardering van de overtreding. De bekentenissen van diverse auteurs, maken de analyse van hun daden nog eenvoudiger.

Veel minder spectaculaire gevallen van gebrekkige wetenschappelijke integriteit, die waarschijnlijk vaker voorkomen, zijn de zogenaamde *questionable research practices (QRPs)*. QRPs worden, in tegenstelling tot de wetenschappelijke doodzonden, niet door iedereen grofweg hetzelfde opgevat. In de meeste gevallen is het namelijk niet meteen zichtbaar of er iets fout is gegaan. Sterker nog, er is vaak onenigheid of er überhaupt wel iets fout is gegaan.

Maar hoe kan hier nu onenigheid over bestaan? Hoe ziet zo een onenigheid er dan uit? De antwoorden op deze vragen hangen samen met de sociale en kennistheoretische organisatie van wetenschap, dat wil zeggen met de manier waarop wetenschap is georganiseerd en de manier waarop wetenschap kennis maakt.

Wetenschappelijke praktijken zijn namelijk niet homogeen. Wetenschap kent vele onderverdelingen: in disciplines, epistemische culturen, denkstijlen en meer. Dat houdt in dat wetenschap zich bezig houdt met verschillende objecten van studie (moleculen, organen, mensen, culturen), gebruik maakt van verschillende instrumenten, data, publicatiestijlen, materiële onderzoekssettingen, dat er verschillende normen gelden voor wat telt als wetenschappelijk, interessant, significant, belangrijk of triviaal; dat er verschillende culturen bestaan voor persoonlijke omgangsvormen, carrièreverloop, prestatiedrang, etc. (Hackett et al., 2016). Deze verschillen lijken voor een buitenstaander wellicht klein en onbeduidend, maar zijn soms wel de bron van onenigheid als het gaat om 'goede wetenschap'. Dergelijke onenigheid

zien we vooral waar interdisciplinair gewerkt wordt en verschillende standpunten met betrekking tot 'goede wetenschap' bij elkaar komen.

Komen eiwitstructuren uit een lab, of uit een computer?

Wetenschappers werken steeds vaker met elkaar samen, en doen dat in steeds grotere, steeds internationalere, en steeds meer interdisciplinaire groepen (Wuchty et al., 2007). Dat leidt tot de nodige worstelingen met betrekking tot de inhoud en organisatie van wetenschap en wetenschappelijke kennis. Wetenschappelijke integriteit is geen set uniforme principes die hier boven staat. Ze maakt onderdeel uit van de worstelingen van de dagelijkse praktijk van wetenschappers.

Interdisciplinair werken is een manier om complexe vraagstukken te bestuderen. Psychologen en neurowetenschappers slaan de handen (en hun perspectieven) ineen in de studie van de hersenen. De genetica kan in een laboratorium bestudeerd worden (door experimenten met erfelijkheid), op papier (door het maken van stambomen), maar ook in een computer (door de analyse van genoomsequenties en erfelijkheidsberekeningen). Dat geldt ook voor werk aan, bijvoorbeeld, eiwitstructuren. Waar verschillende wetenschappelijke stijlen samenkomen, ontstaan worstelingen - ook met betrekking tot wetenschappelijke integriteit. Wanneer moet iets in een laboratorium aangetoond zijn, en wanneer mag het in een computer? Waar ligt de grens tussen het domein van het experiment en dat van software? Of, anders geformuleerd: welke normen zijn er voor verantwoord gebruik van software? Geoffrey Chang is structureel bioloog en houdt zich bezig met hoe aminozuurketens zich vouwen om een eiwit te vormen, zodat dit eiwit functioneel kan zijn. Hij zag zich in 2006 geconfronteerd met een collega die een volledig andere 3d-structuur voor een eiwit publiceerde dan zijn lab een jaar eerder had gedaan. Op zoek naar een verklaring voor het verschil, vond Chang een fout in de code van een computerprogramma dat zijn groep zelf had gemaakt en veelvuldig had gebruikt. Het gevolg: vijf artikelen moesten uit de literatuur worden teruggetrokken en de jonge Chang werd overspoeld met kritiek op zijn handelen en beschuldigingen van zogenaamde *sloppy*

science. Toch waren er ook veel steunbetuigingen en begrip voor zijn handelen. Een gedetailleerdere analyse van de commentaren onthulde dat wetenschappers uit de laboratorium-hoek (biochemici, fysisch biologen, etc.) Chang aanvielen, en dat wetenschappers uit de algoritme-hoek (bioinformatici, statistici) zijn felste verdedigers waren. De affaire Chang legde een verschil in opvatting over goede wetenschap bloot tussen twee verschillende denkstijlen, twee verschillende groepen wetenschappers. Een belangrijk verschil betrof de waarde van en het vertrouwen in software. Waar de ene groep beweerde dat *té* veel vertrouwen in een computerprogramma alleen maar desastreuus kon eindigen en bij gevolg wetenschappelijk onverantwoord was, gaven anderen aan dat ook software feilbaar was en het een onschuldige fout was, eenvoudig op te lossen door de code te repareren. Waar één groep een overtreding van de grens van fatsoenlijke wetenschap zag, zag de ander een methodologische imperfectie. Waar de ene groep Changs integriteit en geloofwaardigheid in twijfel trok, kreeg hij van de andere complimenten voor het snel corrigeren van de fout (Penders et al., 2009).

Wanneer ben je auteur?

Een tweede voorbeeld van onenigheid over 'goede wetenschap' in interdisciplinair onderzoek is te herkennen in de verdeling van auteurschappen bij het publiceren van nieuwe studies. Onjuiste of onverantwoorde toekenningen van auteurschap vertalen zich in het onjuist weergeven van intellectueel eigendom van de wetenschappelijke bijdrage, en het onrechtmatig toe-eigenen van erkenning. In de medische wetenschappen is het dominante model van verdeling van erkenning als volgt: de meeste erkenning is er voor de eerste auteur, veelal een junior onderzoeker; daarna volgt de laatste auteur, meestal een senior onderzoeker met eindverantwoordelijkheid voor de studie; vervolgens de tweede auteur, die experimenteel of anderszins veel heeft bijgedragen aan de studie. Andere disciplines hanteren andere normen: alfabetisch auteurschap (vaak bij rechten) of met een dalende hoeveelheid erkenning van links naar rechts (in veel, maar zeker niet alle sociale wetenschappen). Deze grove onderverdeling van perspectieven kan bij een interdisciplinair project al tot wrijving leiden bij beslissingen over

auteursvolgorde. Andere, fijnmazigere verschillen bemoeilijken dit verder. Veel richtlijnen voor auteurschap, het zij van een instituut of tijdschrift, het zij (inter)nationale integriteitsrichtlijnen, schrijven een minimumbijdrage voor om je te kwalificeren als auteur. Deze richtlijnen bevatten de nodige open normen. In de richtlijnen van de *International Committee of Medical Journal Editors* staat bijvoorbeeld dat er een *significante* bijdrage geleverd moet zijn aan de studie (uitvoering, ontwerp, enzovoort) of een *kritieke* bijdrage aan het schrijfsproces (niet in aantallen woorden, maar conceptueel). De operationalisering van zulke normen verschilt radicaal tussen landen, disciplines, universiteiten, en zelfs werkgroepen binnen dezelfde afdeling. Laboranten mogen in het ene lab wel auteur zijn en in het andere niet (Shapin, 1989). Hetzelfde geldt voor onderzoeksmanagers of studenten. Wanneer is een bijdrage *significant genoeg*? Bij verschillende antwoorden op deze vraag kan het oordeel over een concrete situatie verschuiven van legitiem naar illegitiem auteurschap – een veelvoorkomende QRP (Tijdink et al, 2016) – met alle bijkomende morele kwalificaties van dien.

Worstelen met goede wetenschap

De worstelingen met integriteit in interdisciplinair onderzoek beperken zich niet tot algoritmes en auteurschappen, maar omvatten elke stap in het kennisproductieproces. Welke metingen kan en mag ik vertrouwen? Welke criteria zijn er voor het uitsluiten van bepaalde proefpersonen of datapunten uit een cohort? Welk statistische toets pas ik toe? Kan dat ook een andere zijn? Wie heeft het meeste werk verricht? Weten we dat eigenlijk wel zeker? Dergelijke vragen uit de alledaagse wetenschappelijke praktijk zijn niet eenduidig te beantwoorden en er bestaan meerdere parallelle antwoorden waarbij er niet één per definitie beter is dan de ander. Dat wetenschappers het er over eens zijn dat je geen meetresultaten uit je duim mag zuigen en vervolgens publiceren, betekent niet dat iedereen hetzelfde denkt over wat wetenschappelijk integer handelen nu precies is. Elk oordeel over de kwaliteit en integriteit van wetenschap wordt gearticuleerd vanuit één visie op wetenschappelijke integriteit. Wanneer ik een onderzoek *sloppy* vind, kunnen vele anderen dat met mij oneens zijn, zonder dat één

van ons ongelijk heeft.

Versillen in perspectieven op ‘goede wetenschap’ hebben potentieel grote gevolgen, waaronder het onnodig in twijfel trekken van de integriteit en geloofwaardigheid van individuele onderzoekers en hele onderzoeksgroepen. Om dit te voorkomen, dienen we ons te realiseren dat de grenzen tussen slodderwetenschap en fatsoenlijke wetenschap lokale kenmerken van wetenschap kunnen zijn. Dat ze lokaal zijn, betekent niet dat ze niet geldig zijn of dat iedereen er zijn eigen perspectief op goede wetenschap op na mag houden. Iedereen dient zich te conformeren aan de geldende normen binnen zijn of haar coherente en homogene groep wetenschappers die een object, methode en denkstijl delen. Dat maakt goede wetenschap een lokaal kenmerk van wetenschap, maar daarmee ook beter toepasbaar, en relevanter voor de praktijk. Het betekent echter ook dat de grens tussen slodderwetenschap en fatsoenlijke wetenschap niet overal identiek getrokken kan worden of dat deze aan iedereen uniform opgelegd kan worden. Het betekent dat we in de analyse van, en in het morele oordeel over wetenschap en wetenschappelijke integriteit, gevoelig moeten zijn voor deze verschillen, en dat wanneer we de wetenschap bestuderen, deze verschillen *zelf* onderwerp van onderzoek moeten laten zijn.

Dr. Bart Penders is universitair docent aan de vakgroep Health, Ethics & Society van de Universiteit Maastricht. Hij onderzoekt processen van wetenschappelijke samenwerking, waaronder de totstandkoming van geloofwaardigheid in collectieve kennisproductie.

Referenties

- Hackett, E.J., J.N. Parker, N. Vermeulen & B. Penders (2016). The Social and Epistemic Organization of Scientific Work. In: Handbook of Science and Technology Studies, 4th edition. Cambridge: MIT Press, p. 733-764.
- Penders, B., R. Vos & K. Horstman (2009). A question of style: method, integrity and the meaning of proper science. *Endeavour* 33 (3), 93-98.
- Shapin, S. (1989). The Invisible Technician. *American Scientist* 77 (6), 554-563.
- Tijdink, J.K., K. Schipper, L.M. Bouter, P. Maclaine

- Pont, J. de Jonge & Y.M. Smulders (2016). How do scientists perceive the current publication culture? A qualitative focus group interview study among Dutch biomedical researchers. *BMJ Open* 6, e008681 doi:10.1136/bmjopen-2015-008681
- Wuchty, S., B. F. Jones & B. Uzzi (2007). The increasing dominance of teams in production of knowledge. *Science* 316 (5827), 1036-1039.

Thema: Kwaliteit en integriteit van medisch-wetenschappelijk onderzoek

De invloed van de financieringsbron op de kwaliteit en integriteit van klinisch geneesmiddelenonderzoek

Henk Jan Out

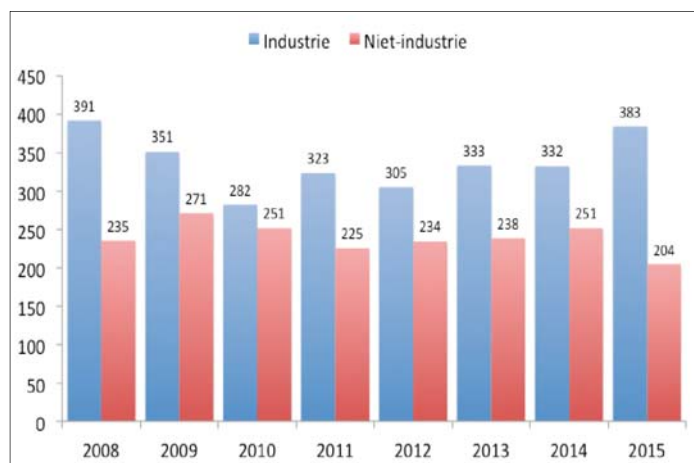
“Wie betaalt, bepaalt” luidt een bekend gezegde. Het klinisch onderzoek naar nieuwe geneesmiddelen door de farmaceutische industrie is daarom per definitie gecompromitteerd, volgens sommigen. De commerciële motieven van deze bedrijven maken het onmogelijk om objectieve kennis te vergaren naar de werkzaamheid en veiligheid van medicijnen. Dergelijk onderzoek zou dan ook veel beter gedaan kunnen worden door onafhankelijke onderzoekers aan academische instellingen. Alleen op die wijze kunnen we belangenverstremming voorkomen en de objectiviteit van de wetenschapsbeoefening garanderen. Kloppen deze overwegingen wel en hoe kunnen we het beste de kwaliteit en integriteit van geneesmiddelenonderzoek borgen zodat patiënten erop kunnen vertrouwen dat de medicijnen die zij innemen werkzaam en veilig zijn?

Onderzoek en ontwikkeling

De farmaceutische industrie ontdekt en ontwikkelt verreweg de meeste geneesmiddelen. Dit proces duurt meer dan tien jaar en kost miljarden per goedgekeurd geneesmiddel. Van alle nieuwe medicijnen die voor het eerst bij de mens worden getest, komt maar tien procent op de markt. De complexiteit van de humane biologie, het gebrek aan inzicht in de exacte oorzaken van ziektes en de enorme regelgeving maken het klinisch onderzoek zeer kostbaar. Ondanks het toe-

nemende gebruik van biomarkers die de effectiviteit van medicinale behandelingen pogen te voorspellen, zijn grote klinische testen noodzakelijk om in relatief grote populaties statistisch onderbouwde uitspraken te kunnen doen over de werkzaamheid van nieuwe geneesmiddelen op goed gedefinieerde, klinisch relevante eindpunten. Het is evident dat alleen private ondernemingen investeringen van soms honderden miljoenen euro met een hoog risico op mislukking kunnen opbrengen.

Ten onrechte menen sommige critici dat geneesmiddeleninnovatie vooral afkomstig is van de universiteiten. Uit de grootste studie waarin dit onderzocht is, blijkt dat van de 1541 door de FDA goedgekeurde registratieaanvragen tussen 1970 en 2009, slechts 143 (9,3%) ontdekt waren in de publieke sector (Stevens et al., 2011). Ook in Europa was maar 17% van de 94 goedgekeurde geneesmiddelen tussen 2010 en 2012 van academische origine (Lincker et al., 2014). Zo'n 70-80% van de totale onderzoekskosten maakt de farmaceutische industrie tijdens de laatste, zeer kostbare, klinische onderzoeksfase. Farmaceutische bedrijven financieren in Nederland ongeveer 60% van het interventionele geneesmiddelenonderzoek (zie Figuur 1).



Figuur 1: Bron van financiering klinisch geneesmiddelenonderzoek in Nederland tussen 2008 en 2015 (Bron: CCMO jaarverslagen)

Kwaliteit

Kwaliteit van klinisch onderzoek heeft een drietal componenten: (1) De wetenschappelijke relevantie van de vraagstelling en de keuze voor de juiste onderzoeksmethode, (2) de zorgvuldigheid van de uitvoering en (3) de compleetheit van de rapportage. Hoe is

anno 2016 de stand van zaken en hoe onderbouwd is de visie van sommige critici dat bedrijven zich schuldig zouden maken aan grootschalige manipulatie van dit onderzoek? En hoe verhoudt zich dat tot het academische, het zogenaamde onderzoeker-geïnitieerde klinische onderzoek?

Vraagstelling

De doelstelling van het meeste onderzoek dat de farmaceutische industrie verricht, is een nieuw middel te registreren bij de Amerikaanse, Europese en Japanse overheden. Na registratie wordt vaak nog aanvullend onderzoek gedaan om extra indicaties te verkrijgen. Ook vindt er dan veel observationeel onderzoek plaats om de alledaagse werkzaamheid te monitoren en zeldzame bijwerkingen te detecteren. Verreweg de meeste bedrijven overleggen vooraf met registratieautoriteiten zoals de *Food and Drug Administration* (FDA) en *European Medicines Agency* (EMA) over het vereiste klinische programma. Bij middelen voor de meest gangbare ziektes en indicaties bestaan formele richtlijnen over de noodzakelijke studies. Placebo-gecontroleerde studies zijn alleen toegestaan als er nog geen geregistreerd alternatief middel beschikbaar is of indien valide bewijs over werkzaamheid niet anders kan worden verkregen. In dat geval zal altijd een opzet gekozen worden die medisch nog steeds verantwoord is voor de onderzochte populatie. In de praktijk kiest de industrie steeds vaker voor vergelijkingen met bestaande middelen om de kans op vergoeding door ziektekostenverzekeraars te vergroten. De protocollen van klinische studies inclusief de statistische analyseplannen moeten worden goedgekeurd door gezondheidsautoriteiten van elk deelnemend land en de medisch-ethische toetsingscommissies (METC's). Het is volstrekt onduidelijk hoe in een dergelijk gereguleerd landschap bedrijven structureel in staat zouden zijn om de FDA, de EMA, en/of METC's om te tuin te leiden met bewust manipulatieve onderzoekdesigns. Om nog maar niet te spreken van de vele onderzoeksinstellingen en haar medewerkers die de studies uitvoeren.

Uitvoering

Patiënten die deelnemen aan klinische studies dienen er van verzekerd te zijn dat hun veiligheid niet

in gevaar komt. Daarnaast is het belangrijk dat de verzamelde gegevens valide zijn. Data-integriteit en patiëntveiligheid worden geborgd in de *Good Clinical Practice* (GCP) richtlijnen. Autoriteiten keuren alleen nieuwe geneesmiddelen goed als de studies volgens GCP uitgevoerd zijn. GCP is inmiddels verankerd in Europese wetgeving nadat de industrie al jaren volgens deze standaard werkte. Universiteiten hebben hier moeite mee. Daar waar bedrijven werken met *electronic data capturing*, gevalideerde databases, interactieve webapplicaties voor randomisatie en voor-geprogrammeerde statistische analyses, werkt menig universiteit nog met papieren formulieren, Excel en verzegelde enveloppes met de randomisatiecode. De statisticus wordt vaak pas betrokken na afloop van de studie. In zo'n setting is het minder moeilijk eindpunten tussentijds te veranderen, zonder enige controle gegevens in databases aan te passen, en post-hoc analyses te doen net zo lang totdat er een statistisch significante uitkomst is. Er is dus geen enkele twijfel over dat de farmaceutische industrie hogere kwaliteitsstandaarden in de uitvoering van klinisch onderzoek hanteert dan universiteiten, de meeste onderzoekers geven dat ook gewoon toe. Dat wil natuurlijk niet zeggen dat onderzoeker-geïnitieerd onderzoek niet valide zou zijn. Ook is de vraag terecht of de regelgeving niet doorgeschooten is en de soms overdreven aandacht op de uitvoering niet ten koste gaat van de inhoud. Maar in ieder geval is voor wat betreft de uitvoering van trials de commerciële financieringsbron eerder een garantie voor hoge dan lage kwaliteit en kunnen integriteitsissues en wetenschappelijke fraude sneller gedetecteerd worden.

Rapportage

Veel is te doen over publicatiebias, het verschijnsel dat negatieve studies niet gepubliceerd worden in tegenstelling tot positieve trials. Dit verschijnsel is bewezen met name bij nieuwe antidepressiva waarbij de beschikbare gegevens bij de FDA werden vergeleken met de gepubliceerde artikelen. Volgens velen het bewijs dat commerciële financiering niet deugt. Aan de andere kant, met overheidsgeld gesponsorde studies naar het effect van cognitieve psychotherapie bij depressie, liet eenzelfde publicatiebias zien, waardoor ook hier het effect van interventie werd overschat. Ongeacht de

sponsor, wetenschappers lopen dus niet warm om negatieve resultaten te publiceren. Medische tijdschriften trouwens ook niet, want die melden liever spectaculaire nieuwtjes. Inmiddels hebben de FDA en de grote internationale tijdschriften belangrijke veranderingen afgedwongen. Onderzoekers moeten hun klinische studies bij aanvang in een trialregister registreren. In de Verenigde Staten is dat zelfs wettelijk verplicht, inclusief het plaatsen van de resultaten op het register *clinicaltrials.gov*. Universiteiten en bedrijven hielden zich er maar mondjesmaat aan, maar inmiddels hebben de bedrijven het voortouw genomen. John Ioannidis, de gerenommeerde methodoloog, heeft weliswaar kritiek op industriestudies maar zegt ook: "They do them very well, they score better on "quality" checklists, and they are more prompt than nonindustry trials to post or publish results." (Ioannidis, 2016).

Industrie versus academia

In vele opzichten is de kwaliteitsborging van industriegesponsorde studies beter dan van onderzoeker-geïnitieerde trials. Zo is inmiddels goed gedocumenteerd dat industriestudies beter *gepowered* zijn, waardoor de kans om een werkelijk bestaand verschil tussen behandelingen te detecteren, groter is. Deze studies vinden dus vaker statistisch significante verschillen op primaire eindpunten en zijn derhalve vaker positief. Dat is geen teken van publicatiebias maar van een intrinsiek betere methodologie. Ook behalen industriestudies vaker dan academische het benodigde aantal patiënten. Het is niet onlogisch. Bedrijven hebben immers meer financiële armslag om studies uit te breiden met nieuwe onderzoekscentra in andere landen dan academische instellingen.

Conclusie

De uitspraak "Wie betaalt, bepaalt" over farmaceutische bedrijven, is tot op zekere hoogte correct. De sponsor bepaalt de prioriteiten op grond van medische noodzaak, biologische kans op succes en commerciële haalbaarheid c.q. aantrekkelijkheid. Het is de enige manier om nieuwe geneesmiddelen op de markt te krijgen. Voor andere noodzakelijke, maar verliesgevende geneesmiddelen zal dus een alternatief financieringsmodel gevonden moeten worden. Antibiotica zijn een goed voorbeeld. De behoefte aan

nieuwe middelen is groot, maar voor bedrijven per definitie verliesgevend aangezien artsen ze alleen als laatste redmiddel voorschrijven. Nieuwe antibiotica zullen er dus alleen komen als er (ook) publiek geld aangewend wordt.

De registratie-autoriteiten bepalen grotendeels de aard van het klinische ontwikkelingsprogramma en daarmee de wetenschappelijke activiteiten van bedrijven. Gezien de enorme regelgeving is de vrees van de samenleving dat private financiering de kwaliteit compromitteert, volstrekt onterecht. Integendeel, dergelijke kwaliteitssystemen en de controle daarop door de Inspectie voor de Gezondheidszorg (IGZ) dragen bij tot publieke geruststelling dat nieuwe geneesmiddelen werkzaam en veilig zijn. Dat betekent niet dat industriestudies niet kritisch moeten worden geïnterpreteerd. Soms zijn de gevonden voordelen van nieuwe middelen in de mega-studies wel statistisch significant, maar klinisch irrelevant. De commerciële prikkels na markttoegang leiden soms tot overdreven effectiviteitsclaims. Met name de claims die de industrie in het verleden maakte voor niet-goedgekeurde indicaties (*offlabel marketing*) hebben terecht geleid tot forse boetes, met name in de Verenigde Staten. Inmiddels is het tij gekeerd en betaalde de industrie in de VS \$44 miljoen aan boetes in de periode 2014-2015 vergeleken met \$2,7 miljard in 2012-2013, een laagterecord (Almashat et al., 2016). Nog steeds teveel, maar een structurele trend lijkt ingezet. Critici van de sector beperken zich vaak tot het uitvergroten van incidenten, het generaliseren van voorvallen uit het (verre) verleden en het ontkennen van recente substantiële veranderingen binnen de bedrijfstak. Daarmee wordt een karikatuur van de farmaceutische industrie gecreëerd en de kwaliteit van het geneesmiddelenonderzoek miskend.

Ten slotte, een kritische evaluatie van het onafhankelijk onderzoek blijft ook nodig. Het ontbreken van financiële relaties betekent immers niet dat er geen belangenverstrengeling zou kunnen bestaan. Uiteindelijk bepaalt de inhoud de kwaliteit van wetenschappelijk onderzoek en niet de bron van financiering. Zo'n inhoudelijke beoordeling is het minste dat de deelnemers aan risicovol klinisch onderzoek door farmaceutische bedrijven verdienen.

Dr Henk Jan Out is Farmaceutisch geneeskundige; visiting professor Department of Pharmacy, Health and Wellbeing, University of Sunderland, UK. www.henkjanout.nl

Aanbevolen en geciteerde literatuur

- Out HJ, Van Meurs P, Van Olden RW, redactie (2014). *Handboek farmaceutische geneeskunde*. Houten: Bohn Stafleu van Loghum.
- Van Lent, M (2015). *Influence of the funding source on quality of clinical drug trials*. Radboudumc, Nijmegen: Proefschrift.
- Out, HJ (2016). *Leve het geneesmiddel! Over de verdiensten van de farmaceutische industrie*. Amsterdam: Prometheus.
- Ioannidis, JPA (2016). *Evidence-Based Medicine Has Been Hijacked: A Report to David Sackett*. *J Clin Epidemiol*, 73, 82-86.
- Stevens AJ et al. (2011). *The role of public sector research in the discovery of drugs and vaccines*. *N Engl J Med*, 364, 535-541.
- Lincker H. et al. (2014). *Regulatory watch: where do new medicines originate from in the EU?* *Nature Reviews Drug Discovery*, 13, 92-93.
- Almashat S. et al. (2016). *Twenty-five years of pharmaceutical industry criminal and civil penalties: 1991 through 2015*. Public Citizen.

Thema: Kwaliteit en integriteit van medisch-wetenschappelijk onderzoek

Ontwikkeling van nieuwe geneesmiddelen betekent niet automatisch een verbetering van de kwaliteit van de kindergeneeskunde. Pleidooi voor een Learning Healthcare System

Martine de Vries

Liefst 40% van de medicijnen die door huisartsen aan kinderen worden voorgeschreven zijn niet geregistreerd en niet goed onderzocht voor daadwerkelijk gebruik bij kinderen. Dat geldt zelfs voor 80% van de medicijnen die op een neonatologie IC worden voorgeschreven. Er zijn talloze voorbeelden van medicijnen die vanwege het ontbreken van gedegen klinisch onderzoek bij kinderen verkeerd gebruikt zijn. Wetenschappelijk onderzoek bij kinderen is noodzakelijk voor vooruitgang in de medische zorg voor kinderen. Kinderen mogen daarin niet achtergesteld worden ten opzichte van volwassenen. De vraag is echter met welk soort onderzoek de kindergeneeskunde het meest geholpen is.

Verschillende soorten onderzoek

Wanneer een nieuw geneesmiddel wordt getest gebeurt dit in een aantal fasen (zie tabel 1). Na eventuele dierexperimenten vindt "first-in-human" onderzoek (fase 1) meestal plaats bij volwassen proefpersonen. Ook de volgende fasen van onderzoek gebeuren over

het algemeen met volwassenen. Onderzoek met kinderen is gebonden aan zeer strikte voorwaarden. Veel geneesmiddelen zijn derhalve niet goed onderzocht voor daadwerkelijk gebruik bij kinderen, maar worden wel aan kinderen voorgeschreven.

Uitbreiden van de mogelijkheden voor onderzoek met kinderen

Al sinds 2006 is er discussie in Nederland over uitbreiding van de mogelijkheden van wetenschappelijk onderzoek met kinderen. In dat jaar signaleerde de Centrale Commissie Mensgebonden Onderzoek (CCMO) in haar jaarverslag dat veel geneesmiddelen die aan kinderen worden voorgeschreven, alleen bij volwassenen zijn getest op werkzaamheid en veiligheid (Centrale Commissie Mensgebonden Onderzoek, 2006). Dit is een ongewenste situatie. Een kind zou theoretisch veel sterker, maar soms ook minder sterk op een nieuw geneesmiddel kunnen reageren dan een volwassene. Ook als de dosis voor het verschil in lichaamsgewicht wordt gecorrigeerd.

De signalering van de CCMO werd direct gevolgd door een reactie uit het veld, zowel van kinderartsen als patiëntenverenigingen.¹ De oorzaak van de slechte wetenschappelijke onderbouwing van de zorg aan kinderen zou liggen in het feit dat de Wet medisch-wetenschappelijk onderzoek met mensen (WMO) bepaald onderzoek met kinderen onmogelijk maakt, namelijk onderzoek waarvan kinderen geen baat kunnen hebben, en waarbij de bezwaren meer dan minimaal zijn en de risico's meer dan verwaarloosbaar. Belangrijk (vooral fase 1 en 2-) onderzoek zou daardoor niet kunnen plaatsvinden in Nederland.

Vanuit de Nederlandse Vereniging voor Kindergeneeskunde (NVK) en de patiëntenverenigingen is vervolgens actief meege gedacht over een oplossing voor het tekort aan gedegen wetenschappelijk bewijs voor behandelingen. Die oplossing werd gezocht in het uitbreiden van de mogelijkheden voor wetenschappelijk onderzoek met kinderen, zoals uiteindelijk beschreven in het rapport van de Commissie Doek (Commissie Doek, 2009). Er werd opgetrokken met politieke partijen (VVD, D66) om de Tweede Kamer ertoe aan te zetten om een voorstel tot wetwijziging van de WMO te realiseren, en om op Europees niveau te lobbyen voor aanpassingen in de EU Verordening betreffende

Fase	Beschrijving
Fase 1	Eerste toediening bij mensen van het middel, meestal bij een klein aantal gezonde vrijwilligers. Er wordt gekeken naar verdraagzaamheid en veiligheid, meestal door middel van oplopende doseringen. Er wordt gekeken hoe het middel werkt in het lichaam, hoe het wordt omgezet en uitgescheiden.
Fase 2	Exploratief therapeutisch: onderzoek van de werkzaamheid bij (een klein aantal) patiënten met de betreffende aandoening. Ook verificatie van het werkingsmechanisme bij patiënten of gezonde vrijwilligers en van de veiligheid bij kortdurend gebruik.
Fase 3	Bevestigend therapeutisch: testen in de medische praktijk van de uiteindelijke dosering op bruikbaarheid en effectiviteit bij grotere aantallen patiënten. Vergelijking met een bestaand middel of placebo. Ook wordt gekeken naar de veiligheid op korte en langere termijn.
Fase 4	Onderzoek na registratie (post-marketing) en gerelateerd aan het indicatiegebied waarvoor het middel is geregistreerd. Deze onderzoeken zijn niet nodig voor registratiedoeleinden, maar wel belangrijk voor het optimaliseren van het gebruik van het geneesmiddel.

Tabel 1: fasen van wetenschappelijk onderzoek

wetenschappelijk onderzoek (EU Regulation, 2014). Beide wetswijzigingen zijn nu bijna een feit. Er is een relatieve bovengrens gekomen voor belasting en risico voor een kind in onderzoek, gerelateerd aan de standaardbehandeling voor de aandoening van het kind. Dit maakt meer niet-therapeutisch (fase 1 en 2-) onderzoek mogelijk. Voor veel kindergeneeskundige onderzoekers is het een mooi resultaat en een voorbeeld van de kracht van effectief lobbyen.

Is uitbreiding de oplossing?

De vraag is echter, of de ingeslagen weg daadwerkelijk de oplossing zal blijken voor het verbeteren van de behandeling van kinderen. De realiteit is dat interventies die getest worden in (pre-)klinische fase 1 en 2 trials slechts een fractie zijn van wat er uiteindelijk in de praktijk gebruikt wordt. Voor het ontwikkelen van nieuwe geneesmiddelen of therapieën is de wetswijziging een goede stap. Maar niet voor de bulk aan bestaande therapieën die zijn gebaseerd op weinig bewijs, maar ook niet snel opnieuw onderzocht zullen worden, juist omdat ze al “standaardzorg” zijn. Het evalueren van de effectiviteit van standaardzorg is niet sexy en is daardoor moeilijk te financieren. Hierdoor blijven kinderen in de dagelijkse zorg risico's lopen door inadequaat bewijs, onbewezen traditionele praktijken en bias in klinische oordelen. De zelden uitgesproken, controversiële, maar mijns inziens centrale vraag is: wat heeft een algemeen kinderarts eraan dat een nieuw cytostaticum wordt getest, terwijl de groep die hiervan profiteert zo klein is? Of dat er zoveel fo-

cus ligt op zeer zeldzame ziekten als Duchenne, terwijl dit om zo jongens per jaar gaat? Terwijl in een groot kinderziekenhuis jaarlijks zo'n 80.000 kinderen passeren die therapieën krijgen voorgeschreven die niet getest zijn. Als we kijken naar de sociale waarde van het onderzoek dat tot nu toe de meeste aandacht en financiële ondersteuning krijgt, is onze focus dan niet totaal verkeerd? Of zoals Nancy Kass het zo mooi verwoordde in een speciale bijlage van de Hasting Center Report over kwaliteit en risico's in de zorg: “we have to reconsider what is morally important” (Kass et al, 2013).

Een Learning Healthcare System als oplossing

Al in 2007 benadrukte het invloedrijke Amerikaanse Institute of Medicine dat patiënten dagelijks risico's liepen door onbewezen praktijken (Olsen et al. 2007). Het instituut kwam met een oplossing in de vorm van “The Learning Healthcare System” (LHS), een systeem waarin structurele evaluatie van zorg centraal staat. Continu leren is de morele mindset. Postmarketing (fase 4-) onderzoek zoals Comparative Effectiveness Research, pragmatische trials en quality improvement studies zijn een integraal onderdeel van de zorg in een LHS. Standaard evaluatie van de dagelijkse zorg en daarmee het genereren van de tot nu toe zo gemiste data is een directe oplossing voor het tekort aan gedegen wetenschappelijk bewijs voor alledaagse behandelingen. LHS is een systeem waarin we dagelijks onze zorg evalueren, verbeteren en veilig

maken.

Dat klinkt ideaal, maar er zitten ook haken en ogen aan. Want is de continue evaluatie een uitvloeisel van zorg, of is het wetenschappelijk onderzoek? (Kass et al 2013) Volgens de CCMO kan ook fase 4 onderzoek aangemerkt worden als onderzoek dat getoetst moet worden volgens de WMO als personen aan handelingen worden onderworpen of hen gedragsregels worden opgelegd. Het maakt dus niet uit of het onderzoek met een wel of niet-geregistreerd middel wordt uitgevoerd, maar of er op een of andere manier inbreuk wordt gemaakt op de lichamelijke en/of psychische integriteit van de proefpersoon, bijvoorbeeld door een extra bloedafname of het langdurig bijhouden van een dagboek om de werking van een middel te evalueren.

Als evaluatief onderzoek aangemerkt wordt als WMO-onderzoek, betekent dit langdurige trajecten langs METC's voor goedkeuring, en eindeloze informed consent procedures, en dat is nu juist vaak de reden dat dit soort onderzoek niet plaatsvindt. Bovendien blijft daarmee een status quo behouden, namelijk "overprotectie" bij dit relatief risicoloos onderzoek en "onderprotectie" van de risico's die kinderen lopen in de dagelijkse klinische praktijk. Als continue evaluatie wordt gezien als uitvloeisel van de dagelijkse zorg, hoe zit het dan met vrijwilligheid van deelname, en met valkuilen m.b.t. gebruik van Big Data? Dit zijn allemaal aspecten die onderzocht moeten worden. In plaats van, of in ieder geval naast, de energie die gestoken wordt in de huidige WMO wijzigingen, zal nagedacht moeten worden over een fundamenteel andere wijze van toetsing en informed consent, bijvoorbeeld over een opt-out systeem² voor pragmatische trials en quality improvement studies, of zelfs over een verplichting voor patiënten om deel te nemen aan evaluatie-onderzoek. Die verplichting zou gerechtvaardigd kunnen worden via het argument dat iedereen die gebruikmaakt van zorg ook zou moeten bijdragen aan de verbetering van de kwaliteit van die zorg.

Volledig nieuw zijn deze gedachten niet. In de EU Verordening (2014) is er in de Overwegingen al aandacht voor 'low-intervention trials' ('klinische proeven met beperkte interventie'). Deze 'klinische proeven met beperkte interventie worden gezien als van cruciaal belang om standaardbehandelingen en -diagnoses te beoordelen en zo het gebruik van geneesmiddelen

te optimaliseren. Daardoor wordt bijgedragen aan een hoog niveau van volksgezondheid. Voor deze klinische proeven is in de EU Verordening ruimte voor een beperktere toetsing en informed consent procedure. Jammer genoeg heeft de lobby van kinderartsen en patiëntenverenigingen deze opening niet gebruikt, maar de nadruk vooral gelegd op een ander deel van de Verordening, namelijk de niet-therapeutische studies.

Voor het verbeteren van de wetenschappelijke onderbouwing van de kindergeneeskunde is innovatief onderzoek nodig naar nieuwe behandelmethoden. Maar net zo belangrijk is het streven naar structurele evaluatie van zorg in Lerende Systemen. Ook daaraan zouden de NVK en de verschillende patiëntenorganisaties moeten werken.

Martine de Vries, MD, PhD, is kinderarts en medisch ethicus bij de Sectie Ethiek & Recht, Leids Universitair Medisch Centrum.

Noten

1. Zie voor verschillende documenten waarin het standpunt wordt verwoord de website van de Nederlandse Vereniging voor Kindergeneeskunde (NVK): <http://www.nvk.nl/Nieuws/Dossiers/Medischoonderzoekbijkinderen.aspx>
2. Een opt-out systeem houdt in dat patiënten niets hoeven te doen om mee te doen met het onderzoek, maar de mogelijkheid hebben ervan af te zien door dit actief te melden.

Literatuur

- Centrale Commissie Mensgebonden Onderzoek. (2006) CCMO jaarverslag 2006. Beschikbaar via <http://www.ccmo.nl/attachments/files/ccmo-jaarverslag-2006.pdf>.
- Commissie Doek. (2009) Advies medisch-wetenschappelijk onderzoek met kinderen. Beschikbaar via: <https://www.rijksoverheid.nl/documenten/kamerstukken/2009/11/26/advies-commissie-doek>.
- EU Regulation no 536/2014 of the European parliament and of the council of 16 april 2014 on clinical trials on medicinal products for human use, and repealing Directive 2001/20/EC.
- Kass NE. Faden RR. Goodman SN. Pronovost P. Tunis

S. Beauchamp TL. (2013) The Research-Treatment Distinction: A Problematic Approach for Determining Which Activities Should Have Ethical Oversight. *Hastings Cent Rep.*;43(s1):S4-S15.

Olsen L. Aisner D. McGinnis JM., eds. (2007) *Institute of Medicine, IOM Roundtable on Evidence-Based Medicine, The Learning Healthcare System: Workshop Summary*. Washington, D.C.: National Academies Press. Beschikbaar via: <http://www.iom.edu/Reports/2007/The-Learning-Healthcare-System-Workshop-Summary.aspx>.

Thema: Kwaliteit en integriteit van medisch-wetenschappelijk onderzoek

Betere kwaliteit en vertaalbaarheid van pre-klinisch wetenschappelijk onderzoek noodzakelijk. Een pleidooi voor een expliciete wereldwijde doelstelling

Merel Ritskes-Hoitinga en Harald Schmidt

De huidige beoordelingscriteria van pre-klinisch onderzoek leiden ertoe dat er door onderzoekers veel en snel gepubliceerd dient te worden in zogenaamde 'high impact factor' journals om goed beoordeeld te blijven worden, carrière te maken en fondsgelden binnen te kunnen halen. Dit werkt averechts voor het streven naar wetenschap van goede kwaliteit en goede vertaalbaarheid van preklinische uitkomsten naar de kliniek. De criteria ter beoordeling van onderzoek worden daarom ook wel 'perverse prikkels' genoemd. Het wordt hoog tijd dat hier verandering in komt, maar wie zit er in de cockpit?

Onderzoeksbeoordeling en -financiering

Waarom kiest een jong iemand ervoor om onderzoeker te worden? In de meeste gevallen is dit om maatschappelijk belangrijke en relevante nieuwe kennis te creëren. Deze maatschappelijke relevante nieuwe kennis betreft hier onderzoeksresultaten die tot succesvolle nieuwe klinische en relevante toepassingen leiden. Het huidige systeem van beoordelen van onderzoek en onderzoekers is erop gebaseerd snel en veel wetenschappelijke artikelen te publiceren, het

liefst in zogenaamde *high impact factor* tijdschriften. Veel publiceren betekent dan ook weer veel subsidiegeld binnenhalen voor verder onderzoek etc. De impactfactor staat voor hoe frequent de artikelen van de laatste 2 jaar in een bepaald tijdschrift worden geciteerd, hetgeen daarmee als een maat voor het belang wordt gezien. Of deze maat voor het belang wel echt valide is, valt te bezien, omdat iets citeren een subjectieve keuze kan zijn die voortkomt uit (uiteenlopende) belangen. Tijdschriften zijn vooral geïnteresseerd in zogenaamd positieve resultaten, waardoor negatieve of neutrale resultaten nogal eens ongepubliceerd blijven of in tijdschriften met een lagere impactfactor worden gepubliceerd, en daarmee automatisch minder invloed hebben. Dit geeft een vertekend beeld van de werkelijkheid, met alle consequenties van dien. Er kunnen bijvoorbeeld klinische studies gestart zijn op basis van gepubliceerde positieve resultaten in dierstudies, waarbij vervolgens blijkt dat het geneesmiddel in de patiënt geen effect heeft (Horn, 2001).

Daarnaast is het noodzakelijk voor onderzoekers om steeds zelf geld voor onderzoek binnen te slepen, om het eigen salaris te kunnen blijven betalen en het gewenste onderzoek uit te kunnen voeren. Als een onderzoeker veel publicaties haalt, liefst in tijdschriften met een hoge impactfactor, dan krijgt de individuele onderzoeker een goede beoordeling, en dit resulteert weer in een hogere kans op nieuw geld voor onderzoek. Het is in dit systeem van belang dat de onderzoeker veelbelovende, dus positieve resultaten scoort en grote beloftes voor de toekomst doet, anders is de kans op nieuw geld voor onderzoek zeer klein, met als risico dat je geen baan en werk meer hebt.

Dit systeem leidt ertoe dat er vooral naar kwantiteit wordt gestreefd, en de kwaliteit het onderspit delft: in veel preklinische publicaties worden bijvoorbeeld de twee uitgangsvoorwaarden voor wetenschappelijk onderzoek, randomiseren en blinderen, niet vermeld. Het 'voordeel' van niet randomiseren en blinderen is dat dit de kans op positieve resultaten vergroot, en daarmee de kans op publicatie in tijdschriften met een hoge impactfactor verhoogt. Ook vele andere belangrijke details van de studie worden niet of onvoldoende in de publicaties beschreven. Dit wordt ondertussen breed erkend en de 'perverse prikkels' van het onderzoekssysteem genoemd (Macleod,

2014). Het probleem met deze inzet op veel publiceren in tijdschriften met een hoge impactfactor is dat dit niet noodzakelijkerwijs gericht is op het creëren van klinisch relevante kennis.

Herhaalbaarheid

Voor geneesmiddelen houdt klinisch relevante kennis bijvoorbeeld in dat een nieuwe behandeling of een nieuwe diagnostiek voor patiënten wordt gerealiseerd. Voordat een geneesmiddel of test naar de kliniek kan, moet preklinisch onderzoek en validatie plaatsvinden. Echter voor het preklinische onderzoek (dit omvat dierproeven en ander preklinisch onderzoek, zoals bijvoorbeeld onderzoek met kankercellijnen) is meermaals vastgesteld dat 51-89% van dit onderzoek niet herhaalbaar is (Freedman, 2015). De overgrote meerderheid van publicaties vermeldt namelijk niet de noodzakelijke details van de uitgevoerde proeven, zodat het voor een ander vrijwel onmogelijk is een dergelijke proef te kunnen herhalen, laat staan hetzelfde resultaat te kunnen genereren. En wanneer bijvoorbeeld bedrijven door academici gepubliceerde studies herhalen, blijkt dat die niet tot dezelfde resultaten komen (Prinz, 2011). Eén van de geconstateerde majeure problemen is dat de kwaliteit van publicaties ernstig tekortschiet, want zij volgen niet de richtlijnen die voorschrijven welke details in een publicatie over dierproeven vermeld dienen te worden (ARRIVE guidelines for reporting). Sinds 2010 staan deze richtlijnen nota bene in de voorschriften van zo'n 600 wetenschappelijke tijdschriften voor de auteurs, maar vrijwel geen van deze tijdschriften dwingen hun eigen richtlijnen af (Baker, 2014). Waarschijnlijk is dit uit vrees dat veel auteurs zullen afhaken bij hun tijdschrift en naar een ander, 'gemakkelijker' tijdschrift gaan. Er is immers haast bij om te publiceren. Daarnaast worden zelfs de twee uitgangspunten van wetenschappelijk onderzoek, blinderen en randomiseren, frequent nog niet toegepast bij dierproefonderzoek, waardoor de kans op positieve resultaten wordt vergroot, en daarmee dus ook de kans op publicatie. ARRIVE vraagt om te vermelden of er geblindeerd en gerandomiseerd is.

Vertaalbaarheid

Omdat veel hoge impactfactor-tijdschriften vooral positieve resultaten publiceren, terwijl publicaties van

negatieve resultaten vaak worden afgewezen, lijken resultaten vaak veelbelovender dan ze in de realiteit zijn. Hierdoor kunnen patiënten worden blootgesteld aan achteraf toch niet-effectieve behandelingen, en/of ernstige bijwerkingen (Horn, 2001; Pound, 2004). Meestal lukt het niet de resultaten van dierproeven te vertalen naar de patiënt. Voor de aandoeningen hersenbloedingen en multiple sclerose is bijvoorbeeld aangetoond dat de vertaalbaarheid van dierproeven naar de mens 0.4% is, dat wil zeggen, dat van 1000 positieve behandelresultaten uit dierproeven er slechts 4 goed werken bij de mens. Dat wordt zeer waarschijnlijk mede veroorzaakt door bovengenoemde factoren. Bovendien wordt het beschikbare bewijs uit dierproeven in zijn algemeenheid onvoldoende geanalyseerd voorafgaand aan klinische studies (Hooijmans, 2013). Dat pre-klinische studies van hogere kwaliteit ook zouden kunnen leiden tot betere translatie wordt gesuggereerd door de uitkomsten van zogenaamde 'carcinogenicity bioassays'. Dat zijn wettelijk verplichte testen om te onderzoeken of stoffen mogelijk kankerwekkend zijn. Deze langlopende hoog kwalitatieve studies, voldoen aan de zogenaamde richtlijnen voor Good Laboratory Practise (GLP). Het is gebleken dat stoffen die in epidemiologische studies kankerwekkend bleken te zijn bij de mens, dit ook in eerdere carcinogeniciteits dierstudies waren (Huff, 1999). Behalve dat de kwaliteit per studie verhoogd dient te worden, zijn ook multicenter studies geïndiceerd, om aan te tonen dat een resultaat 'echt' is, dat wil zeggen op meerdere locaties kan worden gereproduceerd (Kleikers, 2015).

Noodzaak tot verbetering

Het huidige onderzoeksbeoordelingssysteem lijkt uit de bocht gevlogen te zijn en zijn doel voorbij te schieten. Publiceren, fondsenwerving en PR voor veelbelovende resultaten zijn een doel op zich geworden. De zoektocht naar 'de waarheid' is naar de achtergrond geraakt. De huidige beoordelingscriteria van onderzoek lijken helaas haaks te staan op het streven naar de succesvolle vertaling van onderzoek naar de patiënt. Voor het valide kunnen beoordelen en interpreteren van onderzoek is het echter essentieel dat de wetenschappelijke kwaliteitscriteria nauwgezet worden gevolgd en vermeld in publicaties. Dit vergt veel tijd

en aandacht en kritische wetenschappelijke discussies, waarvoor nu 'geen tijd is'.

De *rankings* van individuele onderzoekers en van universiteiten houden ons ondertussen in de greep, zodat er vooral veel positieve resultaten worden gepubliceerd. Niemand durft als eerste het huidige systeem van *rankings* gebaseerd op kwantitatieve productieparameters te doorbreken, vanwege het risico dat de onderzoeker of de universiteit daalt in de *rankings*. Dat heeft meteen een negatief effect op het imago, met alle mogelijke risico's van dien, zoals dalende financiering van eigen onderzoek. Echter, in het kader van wetenschappelijke kwaliteit en integriteit is het noodzakelijk dat er snel algehele verbeteringen komen. In de Lancet (Macleod, 2014) zijn voorstellen gedaan om te komen tot nieuwe vormen van beoordeling van onderzoek, zoals het toetsen van de reproduceerbaarheid van resultaten, de kwaliteit van de rapportage, of een onderzoeker wel ál zijn/haar onderzoek heeft gepubliceerd (en niet alleen de positieve resultaten) en of een onderzoeker de gewenste methodologische kwaliteit volgt.

Maar, om met de woorden van Joris Luijendijk de vergelijking met de financiële wereld maar eens te maken: wie zit er hier nu in de cockpit? Wie gaat dit nu veranderen?

Er rust hoe dan ook een grote verantwoordelijkheid bij diegenen die onderzoek financieren en de editors van wetenschappelijke tijdschriften. Op het moment dat onderzoeksfinanciers een hogere kwaliteit eisen, dan gaat het immers meteen gebeuren. En ook wanneer editors van tijdschriften een hogere publicatiekwaliteit gaan afdwingen. Nu al moeten wij in publicaties wel nauwkeurig de instructies volgen om punten en komma's in de tekst aan te brengen conform de eisen van het tijdschrift, anders wordt het niet gepubliceerd. Het wordt dan ook hoog tijd dat de tijdschriften gaan eisen dat de inhoudelijk wezenlijke punten over hoe een proef is uitgevoerd, in detail worden opgeschreven. Dat is ten slotte veel relevanter dan waar de punten en komma's in de tekst staan.

Expliciete doelstellingen noodzakelijk

In de verantwoordingsrede¹ van 23 april jongstleden heb ik (MRH) een pleidooi gehouden voor een mondiale doelstelling, te weten ½N 2Q 4P. Dit betekent:

mondiaal halveren van het proefdiergebruik (aantal N), verdubbelen van de kwaliteit (Q) en verviervoudigen van de patiëntrelevantie (P) http://www.ru.nl/publish/pages/793780/merel_ritskes_verantwoordingsrede_-_april_2016.pdf.

De impactfactor en het aantal publicaties staan niet in deze doelstelling en moeten dus verdwijnen uit de beoordelingscriteria voor goed onderzoek en voor subsidies.

Kan dat wel, doelen stellen in de wetenschap? Economie werkt al lang met expliciet geformuleerde doelstellingen. De wetenschap hanteert geen expliciet gekozen doelstellingen, maar wel criteria, zoals het aantal publicaties in high impact factor journals om ranking lijsten van individuen en instellingen te kunnen maken. Deze criteria lijken momenteel averechts te werken in het streven naar goede kwaliteit en naar vertaalbaarheid van preklinisch onderzoek naar de kliniek. Er dient derhalve gekozen te worden voor een expliciet doel, namelijk hoge kwaliteit en vertaalbaarheid van biomedische wetenschap, zodat de inzet en gelden voor onderzoek op de juiste manier geïnvesteerd worden. Het nut van doelstellingen is dat het transparantie creëert, dat het helpt om keuzes te maken en dat er getoetst kan worden of we op de goede weg zijn of niet. Die toetsing kan weer leiden tot bijstelling. Hiertoe dienen behalve doelstellingen, ook criteria te worden geformuleerd om die doelstellingen te kunnen realiseren, en er dient te worden gezorgd voor het naleven van die criteria. Kortom, met doelstellingen kan er beleid gemaakt worden en kan de cockpit functioneel worden bemenst, ook mondiaal in de wetenschap. Het *world wide web* is een voorbeeld van een succesvolle wereldwijde samenwerking, waarbij internationale consortia gemeenschappelijke webstandaarden hebben opgesteld en permanent onderhouden (Freedman, 2015). Dus het kan. We moeten het nu 'alleen' ook in de wetenschap nog doen.

Merel Ritskes-Hoitinga is hoogleraar Proefdierkunde, SYRCLE, Radboudumc

Harald Schmidt is hoogleraar Pharmacology & Personalised Medicine, Maastricht University

Noot

1. Een verantwoordingrede is een voordracht waarmee

een hoogleraar publiek verantwoording aflegt van hoe zij of hij in de afgelopen jaren onderzoeksgeld besteed heeft en wat dit heeft opgeleverd. De publieke verantwoordingrede van professor Ritskes-Hoitinga vond 23 april 2016 plaats op de Radbouduniversiteit Nijmegen.

Literatuur:

- Baker D, Lidster K, Sottomayor A, Amor S (2014) Two years later: journals are not yet enforcing the ARRIVE guidelines on reporting standards for pre-clinical animal studies. *PLoS Biol*, 12(1), e1001756
- Freedman LP, Cockburn IM, Simcoe TS (2015) The economics of reproducibility in preclinical research. *PLoS Biology* 13(6), e1002165
- Hooijmans CR, Ritskes-Hoitinga M (2013) Progress in using systematic reviews of animal studies to improve translational research. *PLoS Med*, 10(7), e1001482
- Horn J, de Haan RJ, Vermeulen M, Luiten PG, Limburg M (2001) Nimodipine in animal model experiments of focal cerebral ischemia: a systematic review. *Stroke*, 32(10), 2433-2438
- Huff J (1999) Long-term chemical carcinogenesis bioassays predict human cancer hazards. Issues, controversies and uncertainties. *Annals of the NY Academy of Sciences* 1, 56-79.
- Kleikers PW, Hooijmans C, Göb E, Langhauser F, Rewell SS, Radermacher K, Ritskes-Hoitinga M, Howells DW, Kleinschnitz C, Schmidt HH (2015) A combined pre-clinical meta-analysis and randomized confirmatory trial approach to improve data validity for therapeutic target validation. *Sci Rep* 27(5), 13428
- Macleod MR, Michie S, Roberts I, Dirnagl U, Chalmers I, Ioannidis JPA, Al-Shahi Salman R, Chan A-W, Glasziou P (2014) Biomedical Research: increasing value, reducing waste. *Lancet*, 383, pp. 101-104
- Pound P, Ebrahim S, Sandercock P, Bracken MB, Roberts I (2004) Where is the evidence that animal research benefits humans? *BMJ* 328(7438), 514-517
- Prinz F, Schlange T, Asadullah K (2011) Believe it or not: how much can we rely on published data on potential drug targets? *Nat Rev Drug Discov* 10:712-713.

Thema: Kwaliteit en integriteit van medisch-wetenschappelijk onderzoek

Van statistische significantie naar klinische relevantie

Fenneke Blom en Judith J.M. Rijnhart

Menno Oosterhof beschrijft in zijn column ‘Waardeloos onderzoek’ in Medisch Contact dat de meerderheid van de medische onderzoeken direct de prullenbak in kan, omdat de bevindingen van deze onderzoeken onjuist zijn. Eén van de oorzaken die Oosterhof hiervoor aandraagt is een publicatiebias. Artikelen met positieve statistisch significante resultaten hebben namelijk een grotere kans om gepubliceerd te worden dan artikelen met onverwachte en niet statistisch significante resultaten. Statistische significantie is blijkbaar een belangrijke factor voor het wel of niet publiceren van onderzoeksresultaten. We bespreken in dit artikel waarom de focus op statistische significantie onterecht is en bepleiten een shift van statistische significantie naar klinische relevantie.

Inleiding

Uit een onderzoek van Wager en Williams blijkt dat redacteurs van wetenschappelijke tijdschriften de voorkeur geven aan artikelen met een hoge nieuws-waarde (Wager & Williams, 2013). Doorgaans zijn dit artikelen met positieve statistisch significante resultaten. Ook bij onderzoekers speelt de statistische significantie van de resultaten een belangrijke rol bij de keuze om over te gaan tot publicatie van de resultaten. Zij passen bijvoorbeeld analyses zodanig aan dat deze significante resultaten opleveren (Ferguson & Heene, 2012). In deze zogenaamde ‘fishing expeditions’ gaan onderzoekers bewust opzoek naar statistisch significante effecten in de data. Zowel redacteurs als onderzoekers hechten dus veel waarde aan de statistische significantie van de gevonden resultaten. Statistisch significante resultaten zijn echter niet per

definitie relevant voor de klinische praktijk. Het is dus de vraag of het terecht is dat onderzoekers en redacteurs zoveel waarde hechten aan statistische significantie.

Statistische significantie

De term ‘*statistische significantie*’ wordt gebruikt om aan te geven of men de nulhypothese behoudt of verworpt. Onder de nulhypothese gaat men er van uit dat er *geen* relatie is tussen twee variabelen. Indien er sprake is van een statistisch significante relatie, verworpt men de nulhypothese en is er *wel* een relatie tussen twee variabelen. De ene groep heeft bijvoorbeeld een betere of slechtere gezondheidsuitkomst ten opzichte van de andere groep.

Om te bepalen of er sprake is van statistische significantie maakt men gebruik van de *p*-waarde. De *p*-waarde is de kans dat je de gevonden relatie, of een nog extremer resultaat, vindt terwijl in werkelijkheid de nulhypothese juist is (Sterne & Smith, 2001). Deze kans is gelijk aan 1 als de gevonden relatie overeenkomt met de nulhypothese en deze kans is (bijna) 0 als de gevonden relatie afwijkt van de nulhypothese. Bij een *p*-waarde van bijna 0 is er dus waarschijnlijk een relatie tussen de twee variabelen en wordt de nulhypothese, oftewel de hypothese dat er geen relatie tussen de variabelen is, verworpen.

Bij een *p*-waarde van bijvoorbeeld 0,1 is het echter minder duidelijk of men de nulhypothese moet verwerpen of behouden. In dit geval zijn er grofweg twee verklaringen voor het gevonden resultaat, namelijk: 1) het gaat om toevalsbevinding, of 2) de nulhypothese dient verworpen te worden (Sterne & Smith, 2001). In dergelijke situaties maakt men gebruik van het arbitraire afkappunt van 0,05 als richtlijn voor statistische significantie. Bij een *p*-waarde van 0,05 of lager wordt de nulhypothese verworpen. Men denkt vaak ten onrechte dat een statistisch significant resultaat betekent dat er 95% kans is dat het gevonden effect overeenkomt met het daadwerkelijk onderliggende effect. Een *p*-waarde van 0,05 of lager betekent echter dat de kans op een toevalsbevinding zodanig klein wordt geacht, dat het resultaat waarschijnlijk geen toevalsbevinding is, maar de kans op toevallige statistisch significante resultaten blijft aanwezig.

Klinische relevantie

De p -waarde is een relatief simpele maat om te gebruiken, omdat men enkel hoeft te bepalen of de p -waarde hoger of lager dan 0,05 is om vast te stellen of een effect statistisch significant is. De p -waarde zegt echter niks over de richting, noch de 'klinische relevantie' van het effect (Sterne & Smith, 2001). Een klinisch relevant effect is een effect dat in de medische praktijk betekenisvol is. Het volgende fictieve voorbeeld illustreert dit: uit een onderzoek blijkt dat mensen met overgewicht die een interventie ontvangen gericht op het verlagen van de Body Mass Index (BMI), gemiddeld 0,4 BMI-punten dalen ten opzichte van mensen die de interventie niet ontvangen, met een bijbehorende p -waarde van 0,02. Dit effect is statistisch significant, maar een reductie in BMI van 0,4 is waarschijnlijk niet klinisch relevant, omdat een daling van 0,4 BMI-punten voor veel mensen met overgewicht zal betekenen dat zij na de interventie nog steeds een ongezond gewicht hebben. Het tegenovergestelde is ook mogelijk, bijvoorbeeld als mensen in de interventiegroep gemiddeld 3 BMI-punten dalen ten opzichte van mensen die de interventie niet ontvangen, met een bijbehorende p -waarde van 0,10. Dit effect is niet statistisch significant, maar mogelijk wel klinisch relevant, aangezien de BMI van de mensen in de interventiegroep aanzienlijk is gedaald. De klinische relevantie van een relatie kan dus niet afgeleid worden van de statistische significantie van een relatie.

Onderzoekers laten de keuze om hun resultaten aan te bieden voor publicatie vaak afhangen van de statistische significantie. Idealiter zouden onderzoekers voordat zij hun analyses uitvoeren, al moeten bedenken welk effect zij klinisch relevant vinden. Bij het beschrijven van hun resultaten kunnen zij dan aangeven of de gevonden resultaten daaraan voldoen. In de praktijk laten ook beleidsmakers en artsen zich vaak leiden door de statistische significantie van resultaten in de keuze om tot een behandeling of interventie over te gaan. Ook zij zouden voorafgaand aan het lezen van wetenschappelijke artikelen moeten bedenken welk effect zij klinisch relevant vinden in de keuze om tot een behandeling of interventie over te gaan.

Betrouwbaarheidsintervallen

Een maat waarmee de richting en sterkte van een ef-

fect beter beoordeeld kan worden, is het betrouwbaarheidsinterval. Dit interval van uitkomstwaarden wordt meestal gebaseerd op een betrouwbaarheidsniveau van 95% (Gardner & Altman, 1986). In ons fictieve voorbeeld loopt het betrouwbaarheidsinterval voor de daling van 0,4 BMI punten van 0,1 tot 0,7. Dit houdt in dat als het onderzoek oneindig vaak herhaald wordt, de daling in BMI punten in 95% van deze onderzoeken tussen 0,1 en 0,7 zal liggen. De kans dat de daling in BMI punten bij herhaling van het onderzoek lager is dan 0,1 of hoger is dan 0,7 is slechts 5%. Hoe smaller het betrouwbaarheidsinterval is, hoe groter het vertrouwen dat het daadwerkelijke effect overeenkomt met het gevonden resultaat in het huidige onderzoek. De klinische relevantie van een resultaat kan worden vastgesteld door naar de waarden in het betrouwbaarheidsinterval te kijken. Indien de waarden in het betrouwbaarheidsinterval voorbij het afkappunt voor klinische relevantie liggen, kan het effect als klinisch relevant worden beschouwd. Het te hanteren (arbitraire) afkappunt voor klinische relevantie hangt af van de uitkomstmaat, de populatie en de onderzoeksvraag. Zo zal het afkappunt voor de klinische relevantie van een daling in BMI-punten een ander getal zijn dan voor een daling in bloeddruk. Het afkappunt voor de klinische relevantie dient door deskundigen, zoals onderzoekers, artsen en beleidsmakers, per uitkomstmaat bepaald te worden, hiervoor bestaat geen statistische oplossing.

In de praktijk worden betrouwbaarheidsintervallen helaas veelvuldig gebruikt als statistische significantie-testen. Er is sprake van een statistisch significant effect als de nulhypothese niet in het betrouwbaarheidsinterval ligt (Gardner & Altman, 1986). In ons voorbeeld over de interventie gericht op BMI reductie is de nulhypothese dat er geen verschil is in BMI reductie tussen mensen die de interventie wel of niet ontvingen, het verschil is dus nul. Indien het getal nul niet in het betrouwbaarheidsinterval ligt, is er sprake van een statistisch significant effect. In ons voorbeeld loopt het betrouwbaarheidsinterval van 0,1 tot 0,7, en ligt de nulhypothese dus niet in het interval en is er sprake van een statistisch significant effect. Onderzoekers die dit weten, zijn snel geneigd het betrouwbaarheidsinterval op deze manier te interpreteren en zich dus toch te focussen op statistische significantie.

Oplossingen

Zowel redacteuren van wetenschappelijke tijdschriften als docenten van statistische en methodologische vakken spelen een belangrijke rol bij de totstandkoming van een shift van statistische significantie naar klinische relevantie.

Redacteuren

Steeds meer redacteuren van biomedische wetenschappelijke tijdschriften zijn zich bewust van de beperkingen van de p -waarde. Om een shift van statistische significantie naar klinische relevantie te realiseren wijzen zij artikelen waarin p -waardes gerapporteerd worden af, en vragen de auteurs om betrouwbaarheidsintervallen rondom de gevonden effecten te rapporteren. Het gevaar hiervan is echter wel dat men de betrouwbaarheidsintervallen vervolgens gebruikt om de statistische significantie van resultaten te bepalen. Redacteuren dienen dus ook alert te zijn op de interpretatie van de betrouwbaarheidsintervallen in het artikel (Sterne & Smith, 2001).

In de strijd tegen publicatiebias zijn er ook tijdschriften die bereid zijn om de publicatie van de resultaten van een studie te overwegen vóórdat de statistische analyses zijn uitgevoerd (Nuzzo, 2014). Deze tijdschriften bepalen op basis van de relevantie van de onderzoeksvraag en methodologische kwaliteit of een onderzoek van voldoende methodologische kwaliteit is om de resultaten ervan te willen publiceren. Hierdoor hebben studies met niet statistisch significante resultaten en resultaten die niet overeenkomen met de oorspronkelijke hypothesen evenveel kans om gepubliceerd te worden als studies met positieve statistisch significante resultaten.

Docenten

Docenten van statistische en methodologische vakken, nascholingen en intercollegiale besprekingen spelen een belangrijke rol in de wijze waarop (toekomstige) onderzoekers, artsen en beleidsmakers betrouwbaarheidsintervallen interpreteren. De interpretatie van betrouwbaarheidsintervallen door lezers is namelijk moeilijk te beïnvloeden door redacteuren. Het is van belang dat docenten, en begeleiders, en andere rolmodellen van jonge professionals benadrukken op welke wijze betrouwbaarheidsintervallen gebruikt kunnen worden om de klinische relevantie van resultaten te bepalen. Ook het

bepalen van het afkappunt voor klinische relevantie voor verschillende uitkomstwaarden, populaties en onderzoeksvragen dient hierbij aan bod te komen, zodat men handvatten krijgt voor het bepalen van de klinische relevantie (Nuzzo, 2014). Om docenten van het belang van deze boodschap te doordringen en hen de ruimte te geven er aandacht aan te besteden, zullen ook de samenstellers van (bio-)medische onderwijscurricula, zich hierachter moeten scharen.

Conclusie

De p -waarde is een veelvuldig gebruikte maat binnen de biomedische wetenschappen. De p -waarde zegt echter niks over de klinische relevantie van een effect, en zou daarom geen grote rol moeten spelen in de keuze voor een behandeling of interventie. Statistiek is namelijk niet in staat om te bepalen of een geobserveerd effect relevant is voor de klinische praktijk. Betrouwbaarheidsintervallen geven daarentegen wel informatie over de klinische relevantie van een effect. Het is aan de onderzoekers, artsen en beleidsmakers om te bepalen of het geobserveerde effect en de waarden in het bijbehorende betrouwbaarheidsinterval klinisch relevant zijn. Zowel redacteuren van wetenschappelijke tijdschriften, als docenten van statistische en methodologische vakken spelen een belangrijke rol in de totstandkoming van een shift van statistische significantie naar klinische relevantie.

Fenneke Blom, PhD^{1,2} is coördinator van het Netherlands Research Integrity Network, geeft onderwijs over wetenschappelijke integriteit en is betrokken bij diverse onderzoeksprojecten. Judith J.M. Rijnhart, MSc^{1,3} is werkzaam als PhD-student op de afdeling Epidemiologie en Biostatistiek van het VUmc en houdt zich bezig met de toepassing van methodes voor statistische mediatie analyse binnen epidemiologische studies.

Noten

1. VU Medisch Centrum, Afdeling Epidemiologie en Biostatistiek
2. Netherlands Research Integrity Network
3. EMGO Institute for Health and Care Research

Literatuur

- Ferguson, C. J., & Heene, M. (2012) A vast graveyard of undead theories publication bias and psychological science's aversion to the null. *Perspectives on Psychological Science*, 7(6), pp.555-561.
- Gardner, M. J., & Altman, D. G. (1986) Confidence intervals rather than P values: estimation rather than hypothesis testing. *Br Med J (Clin Res Ed)*, 292(6522), pp.746-750.
- Nuzzo, R. (2014) Statistical errors. *Nature*, 506(7487), pp.150-152.
- Sterne, J. A., & Smith, G. D. (2001) Sifting the evidence—what's wrong with significance tests? *Physical Therapy*, 81(8), pp.1464-1469.
- Wager, E., & Williams, P. (2013) "Hardly worth the effort"? Medical journals' policies and their editors' and publishers' views on trial registration and publication bias: quantitative and qualitative study. *BMJ*, 347, pp.f5248.

Thema: Kwaliteit en integriteit van medisch-wetenschappelijk onderzoek

Toegankelijkheid en hergebruik van medisch-wetenschappelijke onderzoeksdata

Heleen van Luijn

Ruwe data van medisch-wetenschappelijk onderzoek worden niet op een centrale plaats opgeslagen en zijn niet systematisch toegankelijk voor hergebruik door derden. Dit gebeurt wel in toenemende mate in de sociale en geesteswetenschappen en het is ook een eis van NWO om geld te krijgen voor onderzoek. Hierdoor kunnen ook andere onderzoekers de data gebruiken en heranalyses of eigen analyses doen op de data. Daarmee wordt de kans op fraude (welbewuste manipulatie van onderzoeksdata, maar ook slordigheden) kleiner, de ontdekking van fraude groter, en wordt de kwaliteit van het onderzoek beter. In dit artikel staat de vraag centraal hoe opslag en hergebruik van medisch-wetenschappelijke data in ons land is geregeld. Bestaat hier regelgeving voor? En wat zijn de voor- en nadelen van toegankelijkheid en hergebruik van medisch-wetenschappelijke data door derden?

Wetgeving en richtlijnen

In regelgeving, ook Europese, voor medisch-wetenschappelijk onderzoek, zijn momenteel nog geen eisen opgenomen over opslag en hergebruik van ruwe data die in medisch-wetenschappelijk onderzoek verzameld worden. Wel is naar aanleiding van de affaire Stapel en andere onderzoeksfraudezaken (bijv. de Leidse reumatologe die data manipuleerde in het lab) het denken over openbaarmaking van medisch-wetenschappelijke gegevens in een stroomversnelling gekomen.¹ In Nederland bestaat de verplichting om

geneesmiddelenstudies te registreren in een database (bijv. EudraCT). Ook bestaat het Nederlands Trialregister waarin ook dergelijke studie geregistreerd worden. Een EU-portal en database zullen naar verwachting in 2017 gereed zijn. In 2014 is de European Clinical Trial Regulation aangenomen in het Europees parlement, deze treedt in werking in 2018. In deze richtlijnen is opgenomen dat in principe alle gegevens over klinische geneesmiddelentrials openbaar gemaakt moeten worden, tenzij hier bezwaren tegen zijn met het oog op privacy en commercieel belang (denk aan: bescherming van persoonsgegevens en van commercieel vertrouwelijke informatie).² Het gaat hier om regelgeving over het vastleggen van karakteristieken van de studies (bijv. wanneer is onderzoek gestart, hoeveel proefpersonen nemen deel, wat zijn de resultaten e.d.?) (European Medicines Agency, 2015). Het doel van deze richtlijnen (nog geen wetgeving) is om Europees geneesmiddelenonderzoek transparanter te maken. Dit komt zowel de patiënten ten goede die willen deelnemen aan trials als de voortgang van de wetenschap, omdat hierdoor bijvoorbeeld voorkomen wordt dat trials dubbel worden gedaan (European Medicines Agency, 2015). Echter, in deze richtlijnen is niets opgenomen over toegankelijkheid voor, en hergebruik van ruwe medisch-wetenschappelijke data door derden voor heranalyse of nieuwe eigen analyses.

Verspreide opslag

Op dit moment vindt geen centrale opslag van ruwe medisch-wetenschappelijke data plaats in ons land. Dergelijke data worden opgeslagen binnen het instituut of de onderzoeksgroep waar de arts-onderzoekers werkzaam zijn. Hoe lang medische onderzoeksgegevens bewaard mogen worden, is wisselend. Sommige richtlijnen spreken van maximaal 15 jaar, andere van minimaal 15 jaar; voor lichaamsmateriaal gelden weer andere regels. Onderzoeksdata worden veelal ofwel bewaard door medisch-wetenschappelijke centra zoals academische ziekenhuizen of gespecialiseerde centra (bijv. Nederlands Kanker Instituut/Antoni van Leeuwenhoek ziekenhuis), ofwel door de farmaceutische industrie. Maar ook in perifere ziekenhuizen, huisartsenpraktijken en mogelijk op diverse andere plaatsen (bijv. patiëntenorganisaties) worden data opgeslagen. Waar de data bewaard worden hangt af van

wie het onderzoek geïnitieerd en/of betaald heeft. Is dit de farmaceutische industrie, dan worden data door die farmaceutische bedrijven bewaard; is de financier NWO, het ministerie van VWS of Europese subsidiegelden of collectiefondsen (bijv. Koningin Wilhelmina fonds), dan worden data vaak in het instituut bewaard van waaruit het onderzoek is geïnitieerd en veelal ook uitgevoerd. In hoeverre de opgeslagen onderzoeksdata toegankelijk zijn voor andere onderzoekers om de onderzoeksresultaten te repliceren dan wel nieuwe secundaire analyses te verrichten op de verzamelde data met bijvoorbeeld nieuwe onderzoeksvraagstellingen, is de vraag. Hiervoor bestaan geen eenduidige regels. Het lijkt erop dat dit per instituut of zelfs per studie kan worden bepaald. Wel bestaat, zoals gezegd, het Nederlands Trialregister (NTR). Het NTR is een openbaar toegankelijk en vrij te doorzoeken trialregister, waarin studies staan geregistreerd die in Nederland lopen of worden uitgevoerd door Nederlandse onderzoekers. Daar zijn echter enkel gegevens over trials opgenomen en niet de ruwe data zelf die met de trials zijn verzameld. Bij de CCMO worden alle voortgangsrapportages en studierapporten van geneesmiddelenstudies opgeslagen¹. In deze laatste rapportages en rapporten gaat het echter om de geanalyseerde data en de beschrijving van hoe het onderzoek is verlopen, eveneens niet om de ruwe data.

Voordelen

Wat zijn de voordelen van centrale opslag en hergebruik van medisch-wetenschappelijke data? Als dergelijke data op een centrale plaats zouden worden opgeslagen en toegankelijk zijn voor hergebruik door derden, kunnen ook andere onderzoekers de data gebruiken en heranalyses of eigen analyses doen op de data. Hierdoor wordt de kans op fraude (welbewuste manipulatie zoals datafabricatie, maar ook slordigheden) kleiner, de ontdekking van fraude groter, en wordt de kwaliteit van het onderzoek beter. De kans op fraude wordt kleiner, omdat van het toegankelijk maken van de ruwe data een duidelijk preventieve werking uitgaat. Als andere onderzoekers achteraf kunnen kijken in de dataset en heranalyses kunnen doen, wordt men voorzichtiger. Hierdoor kan bijvoorbeeld het (opzettelijk) verkeerd gebruik van statistische methoden worden tegengegaan (de 'pakkans'

is groter). Maar ook kan bij replicatie van de studie waarbij opnieuw dezelfde soort ruwe data worden verzameld, deze nieuwe data worden vergeleken met de al eerder verzamelde opgeslagen ruwe data. Hierdoor kan men bijvoorbeeld (sterke) onregelmatigheden in de eerder verzamelde opgeslagen ruwe data op het spoor komen, en datafabricatie (het verzinnen van data en manipuleren van data) op het spoor komen. Bovendien kan dit in de praktijk betekenen dat er meerdere vraagstellingen beantwoord kunnen worden op basis van de data. Hierdoor wordt onderzoeksgeld (vaak belastinggeld) beter besteed.

Ook zou dit kunnen betekenen dat de CCMO en METC's op één centrale plaats kunnen zien wat er daadwerkelijk aan data is verzameld met goedgekeurd onderzoek en of er publicaties uit zijn voortgekomen. Dit kan het niet publiceren van negatieve resultaten mogelijk voorkomen. Hierdoor kan de bias van het enkel publiceren van positieve resultaten afnemen. Ook kan het mogelijk de toetsing van onderzoek verbeteren, omdat snel en gemakkelijk terugkoppeling kan plaatsvinden over afgerond onderzoek.

Nadelen

Uiteraard zijn er ook haken en ogen aan toegankelijkheid en hergebruik van medisch-wetenschappelijke data door derden. Denk aan onderzoekers die geen inzage willen geven in hun data; het regelen van bescherming van intellectueel eigendom; het mogelijk verkrijgen van patent op data van anderen. Maar ook spelen uiteraard kwesties rond de privacy van de proefpersoon een rol. Kunnen persoonsgegevens nog afdoende beschermd worden als onderzoeksdata voor andere onderzoekers toegankelijk zijn? Een mogelijk ander nadeel zou kunnen zijn dat de incentive om zelf data te verzamelen kleiner wordt, omdat men wacht op data die andere onderzoekers verzamelen om te gebruiken. Hierdoor kan een tweedeling ontstaan tussen onderzoekers die veel tijd steken in het verzamelen van data en anderen die dit niet doen en publiceren op basis van dat van anderen. Ook is het de vraag wie er bepaalt waarvoor de data gebruikt mogen worden, en wie daarop toeziet. Het gaat tenslotte om data die in het kader van een bepaalde specifieke studie zijn verzameld. Tot slot is uiteraard ook van belang dat proefpersonen besluiten om mee te doen met

een specifiek onderzoek - bijvoorbeeld omdat ze dát willen steunen - en dat ze daarvoor informed consent hebben gegeven. Wanneer door hergebruik van de data andere onderzoekers vrij zijn om onderzoek te doen op basis van andere doelen (bijv. mogelijk ook commerciële doelen), is het originele informed consent dan nog geldig? Of dient dan heel anders te worden nagedacht over de informed consentprocedure? De farmaceutische industrie zal zeker tegen toegankelijkheid en hergebruik van ruwe medisch-wetenschappelijk data zijn. Hun belangrijkste argumenten zullen uiteraard te maken hebben met commerciële belangen. Echter, dergelijke kwesties spelen deels (behalve patent, privacy van persoonsgegevens en de veel striktere informed consent) ook in de sociale en geesteswetenschappen (Van Luijn, 2012). Voor de haken en ogen die kleven aan toegankelijkheid en hergebruik van data in de sociale en geesteswetenschappen, zijn inmiddels verschillende bruikbare oplossingen bedacht. Met het oog op het regelen van bescherming van intellectueel eigendom, kan bijvoorbeeld gedacht worden aan een moratoriumperiode (bijv. enkele maanden tot enkele jaren) waarin de ruwe en geanalyseerde data nog niet toegankelijk zijn voor derden. In die periode wordt eerst de onderzoekers die de studie hebben bedacht en opgezet de gelegenheid geboden om nog extra analyses te doen op de data en publicaties te realiseren. Voor de specifieke nadelen die gelden voor medisch-wetenschappelijke data, zullen mogelijk ook bruikbare oplossingen gevonden kunnen worden.

Tot slot

Het is de vraag of individuele excessen van onderzoeksfraude – ook in het medisch-wetenschappelijk onderzoek – die bij tijd en wijle in de publiciteit komen, wijzen op structurele fouten in het wetenschappelijk systeem. Het ontbreken van eisen tot openbaarmaking van de ruwe en geanalyseerde onderzoeksdata, en de daarmee samenhangende ontoegankelijkheid van deze data voor andere medisch-wetenschappelijk onderzoekers en andere geïnteresseerden, zijn wel degelijk te kwalificeren als zwakke plekken in het systeem. Gezien de mogelijke nadelen van toegankelijkheid en hergebruik van medisch-wetenschappelijke onderzoeksdata, is het wel de vraag of en hoe dit het beste geregeld zou kunnen worden. Een goede analyse

van de voor- en nadelen is daarvoor dringend gewenst. Het lijkt echter hoog tijd dat de medische wetenschap het voorbeeld van de sociale en geesteswetenschappen volgt. Niet alleen om fraude te voorkomen, maar ook om onderzoeksgeld zo nuttig mogelijk te besteden.

Dr. Heleen van Luijn is directeur en senioronderzoeker bij het Nederlands Instituut voor Onderzoek en Promotiebegeleiding (NIOP). HeleenvanLuijn@promotiebegeleiding.nl

Noten

1. Denk aan voormalig hoogleraar Sociale Psychologie Diederik Stapel van de Universiteit van Tilburg die onderzoeksdata verzon. Een ander voorbeeld is een reumatoloog aan het LUMC. De vrouw had een test bedacht waarmee de aanwezigheid van bepaalde antistoffen bij patiënten met reumatoïde artritis mogelijk kon worden aangetoond. Dat lukte eerst een paar keer, maar al snel niet of nauwelijks meer. Daarop besloot ze zelf in te grijpen: zij sloop 's avonds en 's nachts het laboratorium in en maakte daar negatieve bloedmonsters positief door er antistoffen van muizen aan toe te voegen. (De Volkskrant, 23 januari 2014).
2. Persoonlijke mededeling medewerker CCMO, 30 september 2016.

Literatuur

European Medicines Agency (2015).

(www.ema.europa.eu/docs/en_GB/document_library/Other/2015/01/WC500180632.pdf.)

Luijn, H. van (2012). De toegankelijkheid van onderzoeksdata in de academische psychologie. *De Psycholoog*, 47 (2): 43-53.

Thema: Kwaliteit en integriteit van medisch-wetenschappelijk onderzoek

De rol van de CCMO en Medische Ethische Toetsingscommissies bij het bewaken van kwaliteit en integriteit van medisch-wetenschappelijk onderzoek

Joop van Gerven

Er is de laatste jaren veel te doen over de betrouwbaarheid en het nut van medisch-wetenschappelijk onderzoek met mensen. Met enige regelmaat verschijnen in de media kritische stukken over wetenschappelijke kwaliteit, integriteit, patiëntenbelangen en maatschappelijke opbrengsten. Dit roept vragen op over de controle van medisch-wetenschappelijk onderzoek. Hoe is dit geregeld en wie is daar verantwoordelijk voor? En kan dit niet beter? In dit stuk wordt nader ingegaan op de positie van de Centrale Commissie Mensgebonden Onderzoek (CCMO) als uitvoerder van de Wet medisch-wetenschappelijk onderzoek met mensen (WMO). De CCMO heeft een aantal wettelijke taken, onder andere bij de erkenning van medisch-ethische toetsingscommissies (METC's), de beoordeling van onderzoek als toetsingscommissie op een aantal specifieke gebieden, en de registratie van en voorlichting over medisch-wetenschappelijk onderzoek in Nederland.

Wetenschappelijke kwaliteit

Nederland neemt internationaal een sterke positie in op het gebied van medisch-wetenschappelijk onder-

zoek. Dit blijkt onder meer uit analyses van publicaties en uit ranglijsten van universiteiten en medische faculteiten, waarbij Nederland al jaren hoog scoort. Uitmakende wetenschappelijke opleidingen, een relatief groot aantal topinstituten, en een goed georganiseerde gezondheidszorg bieden een kader voor efficiënt en hoogwaardig klinisch onderzoek. De beoordeling van onderzoek door verschillende instanties (zoals de wetenschapscommissies van lokale onderzoeksinstituten en subsidiegevers, en de 'peer review' van publicaties) is ook bedoeld om de onderzoekskwaliteit te bevorderen. De METC's (inclusief de CCMO), maken hier onderdeel van uit, door zich bij de beoordeling van protocollen nadrukkelijk ook te buigen over de kwaliteit van de wetenschappelijke vraagstellingen en de methodologie. Het belang van de patiënt/proefpersoon, dat bij de toetsing van medisch-wetenschappelijk onderzoek met mensen centraal staat, wordt immers geschaad als de kwaliteit van het onderzoek slecht is. De METC's en de CCMO zien veel protocollen, waardoor kruisbestuiving van inzichten uit andere gebieden ontstaat en trends of systeemfouten aan het licht kunnen komen. De CCMO probeert deze ervaringen steeds vaker te delen met de onderzoekers, door terugkerende problemen of vragen te bespreken met het veld, los van de beoordelingsprocedures. De interacties en discussies tussen toetsingscommissies en onderzoekers kunnen een belangrijke bijdrage leveren aan de wetenschappelijke vorming van onderzoekers (en ook van de METC-leden zelf) en aan het onderzoeksklimaat in Nederland.

Data integriteit

De betrouwbaarheid van onderzoeksgegevens is een belangrijke voorwaarde voor wetenschappelijke validiteit. De toetsingscommissies controleren de wijze waarop de onderzoeksgegevens volgens het protocol worden verzameld, maar ze kunnen de betrouwbaarheid van deze gegevens maar in beperkte mate inschatten. Er wordt vooral gekeken naar de staat van dienst van de onderzoekers, de aanwezigheid van controlerende systemen in de betrokken instellingen en uiteraard naar de onderzoeksplannen. Tijdens de uitvoering kunnen echter allerlei onvoorziene omstandigheden aanleiding geven tot aanpassingen van het oorspronkelijke studieplan. Als aanpassingen van een

WMO-plichtig onderzoek structureel en substantieel zijn, dient de onderzoeker de wijzigingen als amendement te laten goedkeuren door de toetsingscommissie. Bij incidentele veranderingen is er een grijs gebied, tussen geringe afwijkingen ('deviations') en grovere schendingen ('violations') van het protocol. Daarnaast zijn er allerlei statistische variaties in de onderzoeksgegevens. De manier waarop de data worden geanalyseerd of gerapporteerd heeft grote gevolgen voor de conclusies die eruit worden getrokken. Zowel academische als farmaceutische onderzoekers staan daarbij bloot aan diverse externe invloeden (registratie- of publicatie-eisen), maar ook persoonlijke beweegredenen (inclusief verleidingen) spelen een rol. Al deze factoren hebben een invloed op de betrouwbaarheid van wetenschappelijke publicaties. De METC's en de CCMO kunnen vragen om interimrapportages om de voortgang van de studie in het oog te houden, maar ze hebben geen wettelijke taak bij het controleren van de uitvoering, daarvoor is de Inspectie voor de Gezondheidszorg (IGZ) verantwoordelijk. Waar mogelijk werken CCMO en IGZ samen om de data integriteit te bevorderen.

Maar uiteindelijk maakt de zorg voor betrouwbare gegevens deel uit van de onderzoeksethiek. Het is daarmee de verantwoordelijkheid van academische en farmaceutische instellingen. Zij (moeten) zorgen voor goed onderwijs aan research-professionals en voor een juiste organisatie van het onderzoek.

Het belang van de proefpersoon

Elke onderzoeker zal beamen dat het belang van de proefpersoon een wezenlijk aspect vormt van medisch-wetenschappelijk onderzoek met mensen. Toch is het voor een professional moeilijk om zich volledig in te leven in de positie van de patiënt die meedoet aan een onderzoek of daarvoor wordt gevraagd, al was het maar vanwege de verschillen in kennis en ervaring op het terrein van het onderzoek. Dit wordt in zekere zin weerspiegeld door de samenstelling van toetsingscommissies. In de CCMO zitten meerdere artsen, een embryo-deskundige, een klinisch farmacoloog, meerdere ziekenhuisapothekers, een verpleegkundige, een gedragswetenschapper, een jurist, een methodoloog en een ethicus, alsmede één persoon die het wetenschappelijk onderzoek specifiek beoordeelt vanuit de

invalshoek van de proefpersoon. Het lid van de commissie dat vanuit de proefpersoon een studie beoordeelt, neemt een belangrijke plaats in. Dat lid heeft een onafhankelijke taak door juist geen deel uit te maken van het medisch establishment. Hij/zij dient voldoende maatschappelijke statuur en ervaring te hebben om vanuit het perspectief van de proefpersoon over het onderzoek te kunnen oordelen. Het is voor onderzoekers soms moeilijk om een goed evenwicht te vinden tussen de belangen van de proefpersoon en de doelstellingen van het onderzoek.

Dit dilemma wordt het duidelijkst bij onderzoek met kwetsbare proefpersonen, zoals kinderen en demente of psychiatrische patiënten. Met regelmaat ontstaan discussies over de afhankelijkheid en wils-onbekwaamheid van deze patiënten bij onderzoek, tegenover de noodzaak om juist voor deze groep medische vorderingen te boeken. Het is uiteraard van groot belang om de patiënten en hun vertegenwoordigers zelf intensief te betrekken bij deze afwegingen. Maar tegelijkertijd is het voor de directe betrokkenen moeilijk om niet beïnvloed te worden door minder realistische verwachtingen en hoop op nieuwe behandelingen. Er is bij alle partijen (patiënten, behandelaars, onderzoekers en ontwikkelaars van behandelingen) behoefte aan enige sturing op dit gebied, en er worden momenteel dan ook door de CCMO en de wetenschappelijke verenigingen verschillende documenten opgesteld die een leidraad of een toetsingskader moeten vormen voor onderzoek bij kwetsbare personen. Er wordt ook gewerkt aan het proefpersonen-informatieformulier, dat kort en duidelijk genoeg moet zijn om ook door een persoon met een gemiddelde opleiding (VMBO-4-niveau) begrepen te kunnen worden.

Opbrengsten ('social value')

De vraag wat de maatschappij heeft aan wetenschappelijk onderzoek ligt voor de hand, maar is niet eenvoudig te beantwoorden. Hopelijk vooruitgang, maar dat biedt weinig houvast. In het geval van medisch-wetenschappelijk onderzoek lijkt de opbrengst concreter. Dit type onderzoek richt zich toch vooral op het vergaren van kennis over (menselijk) leven, gezondheid en welzijn, ook al verschillen de meningen over de wegen daar naartoe. De algemene verwachting is dat medisch-wetenschappelijk onderzoek uitein-

delijk leidt tot een betere behandeling en preventie van ziekten. De vruchten van dit onderzoek zijn de afgelopen decennia evenwel kleiner, duurder en zuurder geworden, althans volgens sommigen. Vaak wordt beschuldigend gewezen naar de onmacht van medische onderzoekers en de hebzucht van de farmaceutische industrie. We moeten echter niet vergeten dat de maatschappij er vooralsnog voor heeft gekozen om de ontwikkeling van geneesmiddelen grotendeels in handen van commerciële partijen te leggen. Daardoor worden economische krachten gemobiliseerd die niet automatisch samenvallen met medische en wetenschappelijk belangen. Universiteiten en andere onderzoeksinstituten krijgen te weinig publieke middelen om zelf geneesmiddelen te ontwikkelen volgens de kwaliteitsnormen die de samenleving oplegt aan de farmaceutische industrie.

De rol van METC's en CCMO in deze discussie is beperkt, maar er zijn wel mogelijkheden om academische onderzoekers op weg te helpen bij de ontwikkeling van potentiële geneesmiddelen, die vaak in universitaire instituten hun oorsprong vinden. Toetsingscommissies kunnen bijvoorbeeld hun ervaring met studies van farmaceutische industrieën gebruiken om onderzoekers, die een zelfontwikkelde behandeling willen bestuderen, te helpen om hun protocol op een hoger niveau te krijgen. De CCMO wil haar mogelijkheden uitbreiden om onderzoekers hierover te adviseren, maar hier ligt zeker ook een kans voor de onderzoeksinstituten zelf.

Publieke verantwoording

De CCMO hanteert het formele standpunt dat onderzoeksresultaten gepubliceerd moeten worden. Onredelijke beperkingen in contracten zijn een reden om een onderzoek niet goed te keuren, en dat wapent de onderzoekers in de onderhandelingen met sponsors. In 2014 is de EU Verordening Geneesmiddelenonderzoek vastgesteld, ter vervanging van de eerdere richtlijn uit 2001. Deze nieuwe verordening zal binnen enkele jaren moeten leiden tot meer openheid over de resultaten van geneesmiddelenonderzoek. Hierbij worden vooral voor fase I-onderzoek (de eerste studies met een nieuw geneesmiddel bij gezonde vrijwilligers) ook enkele beperkingen in de timing van publicaties geïntroduceerd die nu in Nederland niet

bestaan en waartegen de CCMO zich tevergeefs heeft verzet. De mate waarin studies werkelijk worden gepubliceerd, is overigens niet goed bekend maar wel een bron van zorg. Uit analyses van goedgekeurde protocollen blijkt dat een aanzienlijk percentage van de studies grote vertraging oploopt of uiteindelijk zelfs niet wordt afgemaakt, meestal vanwege problemen met de rekrutering van proefpersonen. De oorzaken hiervoor moeten worden gezocht bij alle betrokken partijen: de onderzoekers die hun eigen mogelijkheden overschatten om naast alle andere verplichtingen effectief studiedeelnemers te werven; de instellingen die hierbij te weinig structurele steun bieden; de subsidiegevers die misschien te hoge eisen stellen; maar ook de patiënten zelf die niet altijd bereid zijn (beter: gemotiveerd worden) om aan onderzoek deel te nemen; en de toetsingscommissies die bij hun beoordelingen niet altijd scherp genoeg op de haalbaarheid letten.

Conclusies

Medisch-wetenschappelijk onderzoek biedt veel mogelijkheden voor vergroting van het welzijn van de samenleving. Daardoor ligt het ook onder een vergrootglas, waarbij allerlei onvolkomenheden aan het licht komen. De roep om strengere wet- en regelgeving is de laatste decennia luider geworden. De CCMO en de METC's hebben een aantal wettelijke taken op deze gebieden, en die nemen de komende jaren door de nieuwe Europese regelgeving nog verder toe. Er moet gewaakt worden voor bureaucratie en administratieve overbelasting van onderzoekers en commissies. Controle en regulering moet vooral zijn gericht op het bevorderen van de kernwaarden van verantwoord onderzoek: wetenschappelijke kwaliteit, betrouwbaarheid van de resultaten, respect voor belangen van de patiënt/proefpersoon en publieke verantwoording. De beschreven maatschappelijke, medische en persoonlijke dilemma's zetten deze waarden onder druk. Uiteindelijk wordt de verwezenlijking van deze kernwaarden het meest gediend door een open en eerlijke dialoog tussen alle betrokken partijen, gericht op verdere verbetering van het onderzoeksklimaat in Nederland.

Prof. dr. Joop van Gerven is voorzitter ad interim van de Centrale Commissie Mensgebonden Onderzoek (CCMO). De CCMO is uitvoerder van de Wet medisch-

wetenschappelijk onderzoek met mensen (WMO).

Deze bijdrage vertegenwoordigt de visie van de auteur en niet noodzakelijkerwijs de standpunten van de CCMO.

Literatuur

CCMO. Taken CCMO. (<http://www.ccmo.nl/nl/taken/>)

Damen, L., Van Agt, F., De Boo, T., Huysmans, F. (2012) Terminating clinical trials without sufficient subjects. *Journal of Medical Ethics* 38:413-416

De Volkskrant, interview Harald Schmidt. (1 september 2015) Veel medisch onderzoek kan direct prullenbak in. (<http://www.volkskrant.nl/wetenschap/-veel-medisch-onderzoek-kan-direct-prullenbak-in~a4132845/>)

EU. Ethical considerations for clinical trials on medicinal products conducted with minors. Recommendations of the expert group on clinical trials for the implementation of Regulation (EU) No 536/2014 on clinical trials on medicinal products for human use. (http://ec.europa.eu/health/files/clinicaltrials/2016_06_pc_guidelines/gl_1_consult.pdf)

EU. Verordening nr. 536/2014 van het Europees Parlement en de Raad van 16 april 2014 betreffende klinische proeven met geneesmiddelen voor menselijk gebruik en tot intrekking van richtlijn 2001/20/EG. (http://ec.europa.eu/health/files/eudralex/vol-1/reg_2014_536/reg_2014_536_nl.pdf)

European Medicines Agency. Appendix, on disclosure rules, to the "Functional specifications for the EU portal and EU database to be audited EMA/42176/2014". (http://www.ema.europa.eu/docs/en_GB/document_library/Other/2014/12/WC500179339.pdf)

NOS. (23 september 2013) Overzicht fraude in de wetenschap. (<http://nos.nl/artikel/554459-overzicht-fraude-in-de-wetenschap.html>)

NRC, interview Peter Gøtzsche. (18 november 2015) Geneesmiddelen zijn gevaarlijk. (<http://www.nrc.nl/nieuws/2015/11/18/geneesmiddelen-zijn-gevaarlijk-1558376-a662526>)

Pijnenborg, M.A.A., Van der Marck, M.A., Olde Rikert, M.G.M. (2015) Recht Oud, (G)een Probleem.

Ruimere mogelijkheden voor onderzoek met ouderen. Aanpassing Wet Medisch-wetenschappelijk Onderzoek met Mensen. *Nederlands Tijdschrift Geneeskunde* 159: A9118.

Zembla. (15 april 2015) Omstreden medisch experiment bij GGZ Noord-Holland Noord. (<http://zembla.vara.nl/dossier/uitzending/omstreden-medisch-experiment-bij-ggz-noord-holland-noord>)

Thema: Kwaliteit en integriteit van medisch-wetenschappelijk onderzoek

Enkele problemen en dilemma's van METC's bij wetenschappelijke (en dus ethische) toetsing

Rien Janssens en Frits Lekkerkerker

In Nederland wordt veel en goed medisch-wetenschappelijk onderzoek gedaan. Internationaal staat dat onderzoek in hoog aanzien. Toch is er kritiek. Sommige critici beweren dat een hoog percentage van het onderzoek direct de prullenbak in kan (Harald Schmidt). Is dat bij de beoordeling van studievoorstellen al te zien? Kunnen METC's dergelijk onderzoek voorkomen? Zijn METC's voldoende kritisch?

Nederland kent, in tegenstelling tot ons omringende landen, een geïntegreerd toetsingssysteem. Dit betekent dat onderzoeksprotocollen door één enkele commissie zowel op hun wetenschappelijke als ook op hun ethische merites worden getoetst. Terwijl in veel landen, in het bijzonder voor studies met geneesmiddelen en medische hulpmiddelen, expliciet onderscheid wordt gemaakt tussen wetenschappelijke toetsing (deel één) en ethische toetsing (deel twee) is dit in Nederland niet het geval (Kenter 2009). En hier is veel voor te zeggen. Bij een ethische toetsing hoort nu eenmaal ook een wetenschappelijke toetsing. Sterker nog, een wetenschappelijke toetsing is een ethische toetsing. Een protocol dat wetenschappelijk niet valide is, is per definitie onethisch. Wanneer een methodoloog in een commissie oordeelt dat de primaire vraagstelling van een protocol met deze sample, of met deze methode, niet betrouwbaar beantwoord kan worden, velt hij of zij daarmee een ethisch oordeel. De onderzoekers zullen het proto-

col moeten aanpassen op ethische, methodologische gronden. Ten eerste vanwege het feit dat de beoogde proefpersonen onnodig worden belast en aan risico's worden blootgesteld. Ten tweede omdat toekomstige patiënten in gevaar worden gebracht als de niet betrouwbare resultaten zouden leiden tot aanpassing van de standaardzorg.

In deze bijdrage worden enkele problemen en dilemma's beschreven rondom wetenschappelijke (en dus ethische) toetsing waar (leden van) METC's in de praktijk mee worden geconfronteerd. Achtereenvolgens wordt ingegaan op: (1) de inschatting van de te verwachten voordelen; (2) het principe van subsidiariteit; (3) monitoring van lopend onderzoek door METC's; (4) de vraag wat eigenlijk 'medisch-wetenschappelijk' is; en (5) dilemma's rondom studies die al een wetenschappelijke toetsing hebben ondergaan. De auteurs schrijven vanuit hun ervaring als lid (RJ) en voorzitter (FL) van twee toetsingscommissies. Beiden zijn tevens lid (RJ) en voorzitter (FL) van het bestuur van de Nederlandse Vereniging voor METC's (NVMETC). Dit artikel is echter op persoonlijke titel geschreven.

Vooruitgang in de medische wetenschap?

Een geïntegreerde toetsing door één METC vraagt nogal wat. Zo zal de commissie er in ieder geval van overtuigd moeten worden dat het protocol bijdraagt aan de vaststelling van nieuwe inzichten op het gebied van de medische wetenschap (art. 3a van de Wet medisch-wetenschappelijk onderzoek met mensen (WMO)). Dit oordeel kan om specialistische kennis vragen die niet altijd voorhanden is. In zulke gevallen dient de METC haar grenzen te erkennen en expertise van buitenaf in te schakelen. Soms komt het voor dat reviewers uit de METC in de referentielijst van het protocol een trial aantreffen die wel heel veel lijkt op het onderhavige protocol. Heeft het onderzoek niet elders al eens plaatsgevonden? Is niet al veel meer bekend dan de onderzoekers willen doen geloven? Wat is hier eigenlijk nieuw? En, als de risico's en de belasting aanzienlijk zijn, weegt de wetenschappelijke waarde van het onderzoek hier dan wel tegen op? Eén van de voorwaarden waar elk onderzoek aan moet voldoen is, wat in een veel geciteerd artikel een 'favourable risk-benefit ratio' wordt genoemd (Emanuel et al. 2000). 'Benefit' kan hier de deelnemende proefpersoon be-

treffen (therapeutisch onderzoek) maar betreft ook altijd toekomstige patiënten die, volgens de hypothese, van de nieuwe interventie of het nieuw opgedane inzicht kunnen profiteren.

Op voorhand blijft het echter lastig een toekomstig voordeel te voorspellen. Sommige vergelijkende studies onderzoeken de waarde van een algemeen geaccepteerde standaardbehandeling. De vraag hoeveel risico's en belasting je voor de huidige proefpersoon acceptabel acht in relatie tot het te verwachten voordeel voor toekomstige patiënten is in laatste instantie niet te beantwoorden. Het gaat hier immers niet om een mathematische calculus. Een inschatting maken van het te verwachten voordeel is al lastig, laat staan deze inschatting af te wegen tegen de risico's en belasting (vgl. Stukart et al. 2012). De nog immer voortdurende discussies over de fase I oncologische trials waar nieuwe medicijnen voor het eerst aan patiënten worden toegediend, spreken in dit verband boekdelen (hoge belasting, veel risico's en geen direct voordeel voor de huidige proefpersonen vs. een uiterst lastig in te schatten verwacht voordeel voor toekomstige populaties). Gelukkig bestaat een METC uit een multidisciplinair gezelschap, waaronder ook een proefpersoonlid en een ethicus, dat tot een consensus moet komen over goed- of afkeuring. Aan deze consensus is, als het goed is, enig moreel gewicht toe te kennen.

Subsidiariteit

Daarnaast dient de METC zich ervan te vergewissen dat dezelfde vooruitgang van de medische wetenschap niet ook zonder inclusie van proefpersonen bereikt kan worden of door onderzoekshandelingen van minder ingrijpende aard (art. 3b WMO). De wetgever doelt hier op het principe van subsidiariteit dat voorschrijft dat de minst belastende methode altijd de voorkeur heeft om het doel te bereiken. Met het oog op dit principe dient het aantal te includeren proefpersonen juist groot genoeg te zijn voor een betrouwbaar antwoord op de vraagstelling: niet te groot en niet te klein. Immers, te klein betekent dat er waarschijnlijk geen conclusies te trekken zijn, te groot leidt tot onnodige blootstelling aan mogelijk risicovol onderzoek. Als dossieronderzoek de inclusie van proefpersonen kan vervangen heeft dat de voorkeur. En als de vraagstelling beantwoord kan worden door bijvoorbeeld

kwetsbare proefpersonen te excluderen heeft dat eveneens de voorkeur. Zo gebeurt het dat onderzoekers een internationale placebo-gecontroleerde geneesmiddelenstudie in Nederland willen laten toetsen terwijl er in Nederland al een standaardbehandeling wordt voorgeschreven. Terwijl, mits methodologisch goed onderbouwd, in andere landen een placebogroep acceptabel geacht kan worden, wordt deze dan in Nederland als problematisch gezien. De onderzoekers zullen met steekhoudende argumenten moeten komen om hiervoor toestemming te krijgen.

Monitoring van lopend onderzoek

Goedkeuring van een protocol is geen blanco cheque. Recentelijk is hier ook vanuit de toezichthoudende Centrale Commissie Medisch Onderzoek (CCMO) en de Inspectie voor de Gezondheidszorg (IGZ) op gewezen (vgl. Grit en van Oijen, 2015). METC's hebben een taak om de voortgang van studies te monitoren. Tussentijdse aanpassingen dienen voorgelegd te worden aan de toetsende METC evenals zogenaamde SAE's (serious adverse events, ernstige bijwerkingen, ziekenhuisopname en dood) waarbij lang niet altijd een oorzakelijk verband bestaat tussen de nieuwe interventie en de SAE. In jaarlijkse voortgangsrapportages dienen onderzoekers aan te geven of het beoogde aantal proefpersonen daadwerkelijk geïnccludeerd kan worden. In enkele gevallen vragen METC's terecht meer frequente updates hiervan. Het antwoord op de primaire vraagstelling dient te worden gepubliceerd, ook als de experimentele interventie niet bewezen beter is gebleken dan de standaardzorg. In dit verband is het verontrustend te noemen dat niet uit alle WMO-studies publicaties komen. Dit impliceert immers dat de proefpersonen voor niets zijn belast en aan risico's zijn blootgesteld. Ook gebeurt het dat de primaire vraagstelling niet wordt beantwoord (wellicht omdat het resultaat negatief was of dat een nieuwe behandeling geen voordeel heeft opgeleverd), maar dat wel over secundaire vraagstellingen wordt gepubliceerd. De focus op positieve resultaten, niet alleen bij onderzoekers maar ook bij tijdschriften, geeft een enorme bias in de wetenschappelijke literatuur. Monitoring kan onderzoekers ertoe aanzetten om ook negatieve resultaten te publiceren. Dit kan deels ondervangen worden door een verplichting onderzoeksresultaten te publiceren.

In de nieuwe geneesmiddelen verordening is dit ook opgenomen (EU Verordening Nr. 536/2014).

Wat is eigenlijk 'medisch-wetenschappelijk'?

Volgens de WMO toetsen METC's medisch-wetenschappelijk onderzoek waarbij proefpersonen een gedragswijze krijgen opgelegd of waarbij ze aan handelingen worden onderworpen. In de literatuur is al veel gezegd over wat nu precies valt onder de in de wet opgenomen woorden "gedragswijze en handeling". Minder is gezegd over wat verstaan moet worden onder "medisch-wetenschappelijk". Studies die door middel van interviews en vragenlijsten bedoeld zijn om inzicht te verschaffen, bijvoorbeeld over veranderingen bij het ouder worden, missen soms een hypothese of een duidelijke wetenschappelijke vraagstelling. Als je zo'n studie als 'niet wetenschappelijk' zou kwalificeren, zou dat impliceren dat ze ook niet getoetst hoeft te worden. Dit, terwijl er wel degelijk sprake is van het opleggen van een gedragswijze en van de deelnemers een flinke tijdsinvestering wordt gevraagd. Ook zijn onderzoekers nogal eens van mening studies bij vrijwilligers als 'niet toetsingsplichtig' te zien omdat risico's als minimaal worden beschouwd. METC's zullen dan, terecht of niet, geneigd zijn om de belasting te betrekken in de reikwijdte van wat geldt als wetenschappelijk (vgl. ZonMW, 2012): hoe hoger de belasting, hoe breder de definitie van 'wetenschap'. Bij kwalitatief onderzoek doet zich dit probleem ook voor. Soms wordt dit 'niet toetsingsplichtig' verklaard en onder de noemer geschaard van 'niet-wetenschappelijk'. Bij de definitie van 'medisch' spelen soortgelijke problemen. Zo worden soms protocollen met een duidelijk psychologische vraagstelling of studies met voedingssupplementen toch getoetst door METC's. Enerzijds omdat METC's protocollen met een hoge belasting zelf willen toetsen, anderzijds omdat onderzoekers dit eveneens willen, met het oog op internationale publicaties. Wederom geldt (terecht of niet): hoe hoger de belasting, hoe breder de definitie van 'medisch' die dan gehanteerd wordt.

Onderzoek dat eerder wetenschappelijk is beoordeeld

In enkele academische centra, zoals het VUmc, is het vereist om protocollen eerst door een wetenschapscommissie te laten beoordelen alvorens ze kunnen

worden voorgelegd aan de METC. Wanneer studies die door wetenschappelijke organen zoals ZonMW en NWO zijn gehonoreerd, worden voorgelegd aan METC's, betekent dit niet altijd dat de wetenschappelijke toetsing gemakkelijk kan worden afgedaan. Protocollen, voorgelegd aan METC's, bevatten noodzakelijkerwijs meer details van de studies dan de subsidieaanvragen. Iets dergelijks geldt ook voor grootschalige internationale studies die door farmaceutische bedrijven worden gesponsord. Veel van deze studies worden vereist door registratie-autoriteiten als een voorwaarde bij het op de markt brengen van een nieuw geneesmiddel. Soms blijken er, ondanks een grondige eerdere wetenschappelijke toetsing door *peers*, toch ethische, methodologische of andere wetenschappelijke bezwaren te zijn.

Voor onderzoekers is het niet goed mogelijk om ingrijpende aanpassingen door te voeren in het onderzoeksprotocol. Ze hebben immers subsidie verkregen voor het voorliggende protocol of het protocol is in andere landen al goedgekeurd en de studie loopt daar al. METC's worden zo voor dilemma's geplaatst. Enerzijds willen METC's met onderzoekers meedenken en de wetenschappelijke vooruitgang die met het protocol gemoeid is bevorderen. Anderzijds zijn METC's er ook om proefpersonen te beschermen en hen niet noodloos te belasten. Het kan dus niet zo zijn dat bij door externe financiers gehonoreerde protocollen of protocollen die in andere landen al zijn goedgekeurd, de METC voor het blok wordt gezet door de indieners om het protocol om die reden dan ook maar goed te keuren. Onafhankelijke toetsing blijft één van de belangrijkste vereisten voor elke WMO-studie (vgl. Emanuel et al. 2000). Maar dat een eventuele afkeuring op wetenschappelijke en ethische gronden van een reeds gehonoreerde studie een moeizaam en pijnlijk proces is, moge duidelijk zijn. Dit komt gelukkig niet vaak voor, mede omdat onderzoeksvoorstellen worden aangepast naar aanleiding van kritische vragen van METC's.

Conclusie

METC's hebben een grote verantwoordelijkheid bij het toetsen van protocollen op wetenschappelijke merites, maar worden daarbij wel geconfronteerd met problemen en dilemma's: is het onderzoek voldoende doordacht? Is de vraagstelling met het onderzoek-

voorstel te beantwoorden? Wat zijn de risico's van het onderzoek? Is de veiligheid van de proefpersonen voldoende gewaarborgd? En, ja, als bij de beoordeling al blijkt dat, in de woorden van Schmidt, het onderzoek de prullenbak in kan, is het per definitie onethisch om het uit te voeren. Zoals hierboven al is aangegeven is het te verwachten voordeel van een studie op het moment van de beoordeling vaak lastig in te schatten. Het is mede daarom zaak voor METC's om, samen met andere instanties, ook tijdens de uitvoering van het onderzoek betrokken te blijven. De ethische toets kan niet goed van de wetenschappelijke toets worden onderscheiden, beide gaan hand in hand. Het Nederlandse geïntegreerde toetsingssysteem is daarom een groot goed en moet worden gehandhaafd ook na invoering van de nieuwe harmoniserende Europese regelgeving over geneesmiddelen.

Dr. Rien Janssens is Universitair Docent Medische ethiek, afdeling Metamedica, VUMC en ethicus-lid van de METC VUmc. Dr. Frits Lekkerkerker was voorzitter van het College ter Beoordeling van Geneesmiddelen en is voorzitter van de METC Medisch Spectrum Twente

Literatuur

- Emanuel, E., Weindler, D., & Grady, C. (2000) What makes clinical research ethical? *JAMA* 283, pp. 2701-11.
- EU. VERORDENING (EU) Nr. 536/2014 van het Europees parlement en de Raad van 16 april 2014 betreffende klinische proeven met geneesmiddelen voor menselijk gebruik.
- Grit, K. & van Oijen, J. (2015) Toezicht op het wetenschappelijk onderzoek met mensen. Instituut Beleid en Management Gezondheidszorg: Erasmus Universiteit Rotterdam.
- Kenter, M. (2009) Regulating human participants protection in medical research and the accreditation of medical research ethics committees in the Netherlands. *Journal of Academic Ethics* 7, pp. 33-43.
- Schmidt H. (2015) Veel medisch onderzoek kan direct prullenbak in. *De Volkskrant*.
- Stukart, M., Olsthoorn, E., van de Vathorst S., et al. (2012) Tweede evaluatie Wet medisch-wetenschappelijk onderzoek met mensen. Den Haag: ZonMW.

Thema: Kwaliteit en integriteit van medisch-wetenschappelijk onderzoek

Investeer in het toezicht op de uitvoering van onderzoek!

Jacqueline van Oijen, Kor Grit en Roland Bal

Proefpersonen die deelnemen aan medisch-wetenschappelijk onderzoek door ziekenhuizen worden beschermd door wet- en regelgeving. Het protocol van een onderzoek moet goedgekeurd worden door een medisch-ethische toetsingscommissie (METC) en ook het ziekenhuisbestuur moet toestemming geven voor de uitvoering van het onderzoek. Incidenten laten echter zien dat het toezicht op de uitvoering van onderzoek onvolkomenheden kent. Hierbij kunnen proefpersonen onnodige schade ondervinden. Tijd daarom voor een zoektocht naar wat het toezicht op de uitvoering omvat, waarom er onvolkomenheden in dit toezicht bestaan en wat mogelijke oplossingsrichtingen zijn. Wij betogen dat toezicht niet alleen vooraf maar ook na de start van een onderzoek hard nodig is. Dit geldt zeker voor onderzoek dat door de (arts)onderzoeker zelf is geïnitieerd, het *investigator-initiated* onderzoek.

Medisch-wetenschappelijk onderzoek kenmerkt zich door een hoge regeldichtheid, zoals de Wet medisch-wetenschappelijk onderzoek met mensen (WMO). Tevens dragen tal van publieke (zoals de Inspectie voor de Gezondheidszorg (IGZ) en medisch-ethische toetsingscommissies (METC's)) en private actoren (opdrachtgevers, onderzoekers en monitors) bij aan de bescherming van proefpersonen en het borgen van de kwaliteit en integriteit van het onderzoek. Desondanks gaat het soms grondig mis, zoals de dramatisch verlopen Nederlandse Propatria-studie in 2008 en de recente Franse test met een pijnstillertje.

De Propatria-studie in een Universitair Medisch Centrum (UMC), waarbij sprake was van een verhoogde sterfte bij patiënten, leverde na een grondig onderzoek tal van verbeteringen op in het toezicht op de uitvoering. Besloten werd de taakverdeling tussen de IGZ en de Centrale Commissie Mensgebonden Onderzoek (CCMO) te formaliseren. Bij incidenten onderzoekt de CCMO de toetsing van de studie door de betrokken METC en de IGZ neemt de uitvoering van de studie ter hand. In het eindrapport werd onder meer geconstateerd dat de melding van ernstige ongewenste voorvallen voor niet-geneesmiddelenonderzoek niet duidelijk was verankerd in de wet (IGZ, CCMO, VWA, 2009). Dit is in oktober 2015 via een wijziging in de WMO geregeld. Daarnaast was gebleken dat ziekenhuizen niet altijd beseffen dat zij formeel de opdrachtgever/verrichter zijn van het onderzoek. Om deze reden is in het eindrapport aanbevolen de bestuurlijke verantwoordelijkheid voor *investigator-initiated* onderzoek neer te leggen bij de Raden van Bestuur van ziekenhuizen (IGZ, CCMO, VWA 2009). Dit is in 2010 bekrachtigd door voormalig minister Klink'.

Echter, ondanks deze – en andere – verbeteringen doen zich nog steeds onvolkomenheden in het toezicht op de uitvoering voor. Wij betogen dat deze onvolkomenheden te relateren zijn aan een drietal problemen, te weten [1] de beperkte capaciteit en middelen bij instellingsgebonden METC's om toezicht te houden op de uitvoering, [2] onduidelijkheden in wet- en regelgeving qua taakverdeling, en [3] het interne toezicht op *investigator-initiated* onderzoek in ziekenhuizen is nog in ontwikkeling.

Probleem 1: Beperkte capaciteit en middelen bij instellingsgebonden METC's

Wettelijk is geregeld dat erkende METC's worden geïnformeerd door de opdrachtgever/verrichter over bijzonderheden zoals ernstige onverwachte voorvallen in het lopende onderzoek (WMO art. 10). Dit is van belang om de veiligheid van proefpersonen ook tijdens de uitvoering van het onderzoek te waarborgen evenals de validiteit van de ingevoerde data. Uit (evaluatie)onderzoeken (o.a. Stukart et al., 2012; Grit & Van Oijen, 2015) blijkt dat niet alle METC's intensief toezicht houden op de uitvoering van het onder-

zoek. De METC's die verbonden zijn aan één of meerdere ziekenhuizen zijn vaak passief en afwachtend en besteden weinig aandacht aan meldingen en rapportages die zij ontvangen. Deze METC's geven aan beperkt toezicht te kunnen houden op de uitvoering aangezien zij geen capaciteit en middelen hebben om de veelheid aan meldingen en documenten te beoordelen. Bovendien missen ze vaak nadere informatie om een goed oordeel te kunnen geven en is het lastig te bepalen of de afwijking zodanig is dat een nieuw oordeel over het protocol wenselijk is. Er kan een vacuüm in het toezicht ontstaan als er wel een melding wordt gedaan door de verrichter, maar geen inhoudelijke beoordeling plaatsvindt door een METC. Dit kan leiden tot een situatie waarin de verrichtten onrechte in de veronderstelling verkeert dat een melding geen probleem vormt of een herbeoordeling van het protocol niet nodig is omdat de METC geen actie onderneemt (Grit & Van Oijen, 2015). Hierdoor ontstaan onvolkomenheden, mazen in de structuur, waardoor belangrijke signalen onopgemerkt kunnen blijven. Wij constateren dat er hier een fundamenteel probleem aan ten grondslag ligt: er is geen helder omschreven rol voor METC's bij de uitvoering van het onderzoek. Daarnaast hebben instellingsgebonden METC's door de beperkte capaciteit en middelen bescheiden mogelijkheden om te controleren of het onderzoek – bijvoorbeeld het verkrijgen van toestemming van proefpersonen – conform het protocol verloopt.

Probleem 2: Onduidelijkheden in wet- en regelgeving qua taakverdeling

De IGZ dient toezicht te houden op de uitvoering, zo blijkt uit bovengenoemde werkafspraken na het incident in 2008. De IGZ komt in actie als er een incident door een toezichthoudend orgaan of door een proefpersoon wordt gesignaleerd. De METC speelt echter ook een rol bij dit toezicht, al is deze rol niet helder geformuleerd in wet- of regelgeving (zie probleem 1). Daarnaast heeft de IGZ een overkoepelende toezichtstaak omdat zij verantwoordelijk is voor het toezicht op de naleving van de WMO (WMO art. 28). Met andere woorden: de IGZ vervult een dubbelrol in het toezicht op medisch-wetenschappelijk onderzoek, namelijk als toezichthouder die functioneert naast

diverse andere toezichthoudende organen en als een meta-toezichthouder. Deze dubbele rol scheidt onduidelijkheden, waar nog bij komt dat de IGZ voor haar toezichtstaak op de naleving van de WMO afhankelijk is van andere publiekrechtelijke toezichthoudende organen.

Probleem 3. Intern toezicht op *investigator-initiated* onderzoek in ziekenhuizen is nog in ontwikkeling

In het eindrapport van de Propatria-studie is uiteengezet dat bij *investigator-initiated* onderzoek de Raad van Bestuur (RvB) volgens de wet opdrachtgever/verrichter is. Een opdrachtgever/verrichter is er verantwoordelijk voor dat er zowel een adequaat kwaliteitssysteem als een goede infrastructuur aanwezig is voor de uitvoering van het onderzoek (IGZ, CCMO & VWA, 2009). Eén van de taken van de opdrachtgever/verrichter is het monitoren van de onderzoeken. Hierbij wordt beoordeeld of de uitvoering veilig geschiedt en of het vastleggen van gegevens gebeurt in overeenstemming met het protocol en de wettelijke vereisten. Uit de tweede wetsevaluatie van de WMO blijkt dat de wettelijk verplichte *on-site* monitoring voor *investigator-initiated* geneesmiddelenonderzoek een grote belasting is. Onderzoekers hebben namelijk moeite onafhankelijke monitors te vinden en kampen met de hoge kosten van monitoring (Stukart et al., 2012). Deze problemen spelen minder bij het commerciële geneesmiddelenonderzoek waar veelal door de farmaceutische industrie voldoende middelen worden vrijgemaakt om elk onderzoek frequent en nauwgezet te monitoren.

We zien dat ziekenhuizen voor het interne toezicht op de uitvoering van onderzoek onder meer afhankelijk zijn van het eigen standpunt van de RvB, de ondersteuning via de brancheorganisatie en externe prikkels zoals een inspectiebezoek. De mate waarin dit interne toezicht in ziekenhuizen is uitgekristalliseerd, verschilt.

Na het incident rond de Propatria-studie hebben de meeste Raden van Bestuur van de acht UMC's duidelijke stappen genomen om dit interne toezicht vorm te geven, zeker in hun rol als opdrachtgever/verrichter bij *investigator-initiated* onderzoek. Per slot van rekening is één van de kerntaken van UMC's het verrichten

van medisch-wetenschappelijk onderzoek. Als steun in de rug heeft de Nederlandse Federatie van Universitaire Medische Centra (NFU) in 2012 een advies uitgebracht over hoe kwaliteitsborging van onderzoek kan worden vormgegeven wanneer de RvB opdrachtgever/verrichter is (NFU, 2012).

Van de ongeveer 80 algemene ziekenhuizen hebben de 26 opleidingsziekenhuizen onderzoek als speerpunt benoemd. Deze ziekenhuizen hebben zich verenigd in de Samenwerkende Topklinische opleidingsZiekenhuizen (STZ). STZ stelt minimumeisen aan het doen van wetenschappelijke onderzoek gericht op o.a. het stimuleren van een gedegen onderzoeksklimaat (STZ Ziekenhuizen, 2014). Concrete handvaten voor *investigator-initiated* onderzoek ontbreken. We constateren dat het interne toezicht op *investigator-initiated* onderzoek, zeker in de opleidingsziekenhuizen, zich nog in een ontwikkelfase bevindt.

Oplossingsrichtingen

In het licht van de geschetste probleemgebieden kunnen we constateren dat het toezicht op de uitvoering van het onderzoek naast het beschermen van proefpersonen ook het toezien op het kwaliteitsniveau van de uitvoeringspraktijk omvat.

Voor de eerste twee probleemgebieden geldt dat de IGZ een voorname rol kan spelen bij het bespreekbaar maken van geconstateerde leemtes in de taakverdeling. De IGZ is vanwege haar overkoepelende toezichtstaak bij uitstek het orgaan om knelpunten in het toezicht op de uitvoering bespreekbaar te maken bij de CCMO, de wetgever (Ministerie VWS en Tweede Kamer) en de overige betrokken toezichthoudende organen. Er is reflectie nodig op de taakopdracht van de METC's zodat een heldere taakverdeling tussen METC en de IGZ bij het toezicht op de uitvoering ontstaat. Daarnaast dient er aandacht te zijn voor problemen rond de capaciteit en de middelen van METC's. Wij zijn van mening dat een METC een grotere rol verdient bij het toezicht op de uitvoering. De METC heeft het protocol beoordeeld en is daarom het beste in staat de uitvoering en de meldingen over het lopende onderzoek te beoordelen. Indien er sprake is van een ernstig incident kan deze worden doorgeleid naar de IGZ. Consequentie van dit voorstel is dat er meer staf nodig is bij METC's. Het realiseren van extra financie-

ring is geen sinecure bij instellingsgebonden METC's aangezien zij een laagdrempelige kostenstructuur hanteren voor *investigator-initiated* onderzoek (Grit & Van Oijen, 2015).

Voor het derde probleemgebied geldt dat er bij niet-commercieel *investigator-initiated* onderzoek in ziekenhuizen minder financiële middelen beschikbaar zijn waardoor zaken als monitoring vaak onder druk staan. Als er weinig geld beschikbaar is voor een studie is de onderzoeker geheel afhankelijk van hoe het ziekenhuis(bestuur) de ondersteuning binnen het ziekenhuis heeft geregeld. Het ontbrekende geld maakt dit soort studies dubbel zo kwetsbaar. Ziekenhuisbesturen zijn verantwoordelijk voor het interne toezicht en de kwaliteitsborging. Dit betekent dat zij tijd, geld en capaciteit dienen te regelen om onderzoek gedegen te monitoren. Hierbij kan een creatief proces nodig zijn om bijvoorbeeld budgetneutrale oplossingen te bedenken in samenwerking met andere afdelingen binnen het ziekenhuis.

Investeer in toezicht op de uitvoering

Een goedgekeurd protocol is niet een afdoende waarborg dat de uitvoering daadwerkelijk goed verloopt. Aan medisch-wetenschappelijk onderzoek blijven altijd risico's verbonden en bovendien blijft het uitvoeren van medisch-wetenschappelijk onderzoek mensenwerk. Dit maakt dat ook het toezicht op de uitvoering prioriteit moet hebben bij de IGZ, de METC's en de (brancheorganisaties van) ziekenhuizen, om zwakke schakels te versterken. Proefpersonen, en latere potentiële gebruikers van onderzochte middelen, moeten erop kunnen vertrouwen dat ook tijdens het onderzoek met voldoende afstand (lees met 'vreemde ogen') wordt meegekeken zodat signalen tijdig kunnen worden waargenomen. Voor elk onderzoek dienen dan ook voldoende middelen vrij gemaakt te zijn voor het bewaken van de uitvoering van het onderzoek en het beschermen van proefpersonen.

Drs. Jacqueline van Oijen, externe promovendus. Dr. Kor Grit, universitair docent Beleid van de gezondheidszorg. Prof. dr. Roland Bal, hoogleraar Bestuur en beleid van de gezondheidszorg. De auteurs zijn verbonden aan het instituut Beleid en Management

Gezondheidszorg, Erasmus Universiteit Rotterdam

Noot

1. Kamerstukken II, 2010/11, 32 123, XVI nr. 130

Literatuur

IGZ, CCMO & VWA (2009) *Onderzoek naar de Propaganda-studie. Lessen voor het medisch-wetenschappelijk onderzoek met mensen in Nederland*. Den Haag/Utrecht: IGZ, CCMO, VWA.

Grit, K. & van Oijen, J. (2015) *Toezicht op het medisch-wetenschappelijk onderzoek met mensen: het in kaart brengen van een multi-centered speelveld*. Rotterdam: Erasmus Universiteit Rotterdam, Instituut Beleid & Management Gezondheidszorg

NFU (2012) *Kwaliteitsborging van mensgebonden onderzoek 2.0*. Utrecht: Nederlandse Federatie van Universitair Medische Centra.

Stukart, M.J., Olsthoorn-Heim, E.T.M., van de Vathorst, S., van der Heide, A., Tromp, K., de Klerk, C. (2012) *Tweede evaluatie Wet medisch-wetenschappelijk onderzoek met mensen*. Den Haag: ZonMw.

STZ ziekenhuizen (2014) *Toelatings- en hervisitatiecriteria STZ lidmaatschap*. Vastgesteld in de Ledenvergadering STZ d.d. 11 april 2014.

Verslag

Tijd voor Kwaliteit! Verslag van een symposium

*Roos van Westrhenen en
Teun van Gelder (organisatoren)*

Op 8 juli 2016 vond in het Erasmus MC te Rotterdam het symposium “Tijd voor Kwaliteit!” plaats. Doelstelling was het bevorderen van kwaliteit van klinisch onderzoek door het vergroten van kennis bij onderzoekers, apothekers en medici. Het ging daarbij over zowel praktische als inhoudelijke zaken van klinisch onderzoek en het bevorderen van discussie over het onderwerp.

Eerste spreker was Gemma Voss, directeur van het Clinical Trial Centrum Maastricht, met ruime ervaring als consultant bij mensgebonden studies met name op het gebied van Doelmatigheidsstudies. Ter verbetering van de kwaliteit van medisch-wetenschappelijk onderzoek met mensen heeft het Clinical Trial Centrum Maastricht (CTCM) een methodiek voor risicoclassificatie ontwikkeld. Het doel van de risicoclassificatie is het vooraf inschatten van het risico van mensgebonden Investigator Initiated studies. De risicoclassificatie inventariseert risico's op drie gebieden, namelijk risico's voor de patiëntveiligheid, maatschappelijke en overige risico's. Een studie wordt op basis van diverse criteria geëvalueerd naar hoog, matig of laag risico. Op het vooraf ingeschatte risico wordt vervolgens de mate en intensiteit van monitoring gebaseerd. In 2015 viel 77% van de studies in de categorie laag risico, 18% in de categorie matig en 5% in de hoog risico categorie. Opvallend in dit verhaal was dat deze monitoring wordt betaald door de Raad van bestuur van het Maastricht UMC.

De tweede spreker was Marleen Kemper, ziekenhuisapotheker en klinisch farmacoloog in de apotheek van het AMC. Sinds 2008 is zij hoofd van het Kenniscentrum Geneesmiddelenonderzoek. Met meer dan

250 lopende studies per jaar heeft het team van (ziekenhuis)apothekers en assistentes veel kennis verworven met het faciliteren en adviseren van geneesmiddelenonderzoek. In deze tweede lezing werd stilgestaan bij de volgende vraagstukken: Wat houdt kwaliteit van studiemedicatie in? Welke wet- en regelgeving zijn hierbij belangrijk? Wanneer is er sprake van studiemedicatie? Wat is de rol van een ziekenhuisapotheker? Er werd kort ingegaan op de belangrijkste wet- en regelgeving: de geneesmiddelenwet, GCP, GMP en GDP. Aan de hand van voorbeelden uit de praktijk werd duidelijk waar een onderzoeker aan moet denken bij het opzetten en uitvoeren van geneesmiddelenonderzoek en hoe de wet- en regelgeving toegepast dient te worden.

De derde spreker, Petra Westveer, is Adviseur onderzoek en kwaliteitsfunctionaris bij het Clinical Trial Center van het Erasmus MC Kanker Instituut. Zij is meer dan vijftien jaar werkzaam in klinisch geneesmiddelenonderzoek, onder andere als data manager, trial manager en clinical research associate. In haar lezing ging zij in op de Trial Master File (TMF). Het papierwerk van een studie op orde houden is niemands favoriete bezigheid. Toch is de TMF een onmisbaar onderdeel voor het doen van goed onderzoek. In deze presentatie werden de volgende vragen beantwoord: Wat is de TMF? Waarom heb je een TMF nodig? Hoe houd je de TMF gemakkelijk op orde? Opvallend is dat de bewaartermijn toch zo'n 15-25 jaar is geworden, veel langer dan de meeste mensen zich nu realiseren. Dit stelt aparte eisen aan bewaren en opslag, waarbij goede archivering noodzakelijk is.

Na de koffie vertelde Inge Huitinga, directeur van de Nederlandse Hersenbank, over het opzetten en de ontwikkeling van deze biobank, die in 1985 door Dick Swaab werd opgezet. Van huis uit is Inge Huitinga medisch bioloog en immunoloog. Binnen het Nederlandse Herseninstituut leidt zij de Huitinga groep, die zich met name bezighoudt met onderzoek naar Multiple Sclerose. Zij liet zien dat het huidige onderzoek naar hersenziekten in het algemeen met name geschiedt met behulp van MRI, plasma, single nucleotide polymorphism (SNPs), diermodellen en post mortem hu- man weefsel. Dit laatste gebeurt bij de Hersenbank op zeer elegante en gestructureerde wijze. Door 24-uurs diensten is tijdig uitsnijden van hersenweefsel

mogelijk. De Hersenbank vervult soms ook een rol in de diagnostiek van slecht begrepen klinische beelden, maar het hersenweefsel is met name beschikbaar voor onderzoeksdoeleinden, ook voor externe onderzoekers van over de hele wereld.

Hierna was het de beurt aan Hanneke den Breeijen, coördinator research support van het Trial Coordination Center aan UMC Groningen. Eerder was zij werkzaam als datamanager bij UMC Utrecht, TNO en het Erasmus MC. Zij vertelde over de ontwikkeling van een digitale handwijzer in het UMC Groningen, de TOOLBOX, die onderzoekers stap voor stap door alle praktische zaken rondom onderzoek loodst. Deze TOOLBOX is niet verplicht in het UMC Groningen, maar steeds meer onderzoekers maken er gebruik van.

Na de pauze sprak Annemieke Schellevis, die na een gedegen kennismaking met het monitoren van klinisch onderzoek in een grote Amerikaanse CRO, overstapte naar de wereld van UMC's. Sinds 2003 werkt Annemieke in het Erasmus MC en zij ontwikkelde in 2014 de interne WMO-audit, die inzicht geeft in de compliance van klinisch WMO-plichtig onderzoek waarvan Erasmus MC de verrichter is. Zij richtte zich in haar presentatie op hoe je er als onderzoeker voor zorgt je ervoor dat je onderzoek goed verloopt én dat je kan aantonen dat je volgens de regels werkt en dus 'compliant' bent aan de WMO en GCP vanaf het begin van je studie. Hierbij werd uitgelegd wat de bedoeling is van monitoring en werd uitgebreid aandacht besteed aan het verschil tussen monitoring en auditing van onderzoek. Met voorbeelden uit de audit-praktijk werd het nut en de noodzaak van monitoring en auditing geïllustreerd. Al vrij associërend kwam zij met de zaal tot de volgende zaken die hiervoor nodig zijn: visie/doel, afspraken/wetten, beleid, facilitering, bewaking en bijsturing.

De daarop volgende spreker, Bert Smit, is stafadviseur van de Directie voor Patientcare op het Erasmus MC en daar tevens werkzaam als senior consultant Neonatologie. Hij vertelde over een Kepra studie waar hij bij betrokken was en over IGZ bezoek. Hierna vertelde hij vooral over het nut van een proefinspectie, zoals hij die regelmatig doet op het Erasmus MC.

Adam Cohen sloot de dag af. Hij is bestuurder van het Centre for Human Drug Research in Leiden

en hoogleraar klinische farmacologie bij het LUMC. Tussen 1999 en 2011 was hij vicevoorzitter van de CCMO en voorzitter van de implementatiecommissie van de EU-regels voor klinische trials. In zijn betoog werd benadrukt dat integriteit moeilijk kan worden gedefinieerd, en daardoor moeilijk kan worden gestimuleerd. Veelal blijft het bij goede voornemens in het beste geval, en vrij ongecontroleerde repressie in het slechtste geval. Dikwijls concentreert integriteit zich op het vinden en dan elimineren van wat wordt gezien als rotte appels in de mand. Er is zelden introspectie ten aanzien van de rol van het systeem in geval van integriteitsschending. Het belangrijkste doel van deze lezing was het stimuleren van die (organisatorische) introspectie in tegenstelling tot het gedachteloos overnemen van regels. Dat gebeurde via het bespreken van een aantal zeer spraakmakende gevallen, waar Adam Cohen op meer of mindere afstand bij betrokken was.

Al met al een zeer gemêleerde dag waar zowel onderzoekers als apothekers en klinici op afgekomen zijn. Dit geeft aan dat er vanuit allerlei verschillende disciplines behoefte is aan meer kennis en discussie over het (vergroten van) kwaliteit van klinisch onderzoek. De doelstelling is gehaald, aangezien er kennis m.b.t. kwaliteit van klinisch onderzoek is gedeeld en hierdoor is verspreid. Ook kwam tijdens de middag een discussie op gang over wat kwaliteit van onderzoek nu precies inhoudt. Echter, dit symposium leek eerder een aanzet tot kennisverbreding dan een sluitende bevrediging van de behoefte van onderzoekers aan meer kennis en discussie over kwaliteit van onderzoek. Opzet is dan ook om deze dag met enige regelmaat (elke paar jaar) te herhalen om in deze behoefte te voorzien.

Dr. Roos van Westrhenen (MD, PhD), psychiater & klinisch farmacoloog, Capri hartrevalidatie / Afdeling psychiatrie Erasmus MC en professor dr. Teun van Gelder (MD, PhD), internist & klinisch farmacoloog, afdeling Apotheek Erasmus MC.

Aankondiging

Een ontleding van oude en nieuwe concepten

Zaterdag 5 november 2016; 10.15 – 16.30 uur
Leids Universitair Medisch Centrum, Leiden

De begrippen ziekte en gezondheid lijken vanzelfsprekend en duidelijk, maar zijn dat allerminst. Hoewel artsen, paramedici, onderzoekers en beleidsmakers dagelijks met deze begrippen werken, blijkt het niet simpel om ze eenduidig te definiëren en af te bakenen. Toch is het belangrijk te weten wat we precies bedoelen met gezondheid en ziekte.

Met onder andere Machteld Huber, Marianne Boenink, Elselijn Kingma, Berend Verhoeff, Jacqueline Kool en Luc Bonneux.

Aanmelden kan tót 28 oktober
via de website:

www.filosofieengeneeskunde.nl/congres

Kosten voor het congres bedragen:

- VFG-lid: **gratis**
- Niet-lid normaal: **€80**
- Niet-lid student/promovendus: **€30**

Accreditatie:

5 punten voor huisartsen, specialisten ouderengeneeskunde, artsen voor verstandelijk gehandicapten, medisch specialisten en sociaal geneeskundigen.

Meer informatie

www.filosofieengeneeskunde.nl of

info@filosofieengeneeskunde.nl