



Universiteit
Leiden
The Netherlands

Towards improving quality of the evidence base for medical decision-making

Jansen, M.S.

Citation

Jansen, M. S. (2026, January 21). *Towards improving quality of the evidence base for medical decision-making*. Retrieved from <https://hdl.handle.net/1887/4289977>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4289977>

Note: To cite this publication please use the final published version (if applicable).

Dutch summary

Nederlandse samenvatting

Nederlandse samenvatting

Wetenschappelijk onderzoek vormt een hoeksteen van de moderne gezondheidszorg. De resultaten van medisch onderzoek dienen als belangrijke basis voor dagelijkse beslissingen die in de klinische praktijk worden genomen.^{1,2} Hoewel de medische wetenschap de afgelopen decennia een enorme groei heeft doorgemaakt,³ bestaan er zorgen over de betrouwbaarheid, reproduceerbaarheid en waarde van het wetenschappelijk bewijs. Zo wordt geschat dat een aanzienlijk deel van de bevindingen uit medisch onderzoek niet klopt, en dat tot wel 85% van de investeringen in biomedisch onderzoek verloren gaat in diverse fasen van de onderzoeksketen.^{4,5} De oorzaken hiervan zijn talrijk en complex. Dit proefschrift richt zich op een deel van deze knelpunten, met als uiteindelijk doel om de kwaliteit van wetenschappelijk bewijs voor de gezondheidszorg te verbeteren.

Het proefschrift bestaat uit twee delen. **Deel I**, *‘Challenges related to generation and dissemination of evidence’*, behandelt knelpunten in het ontwerp en uitvoering van studies, en in de verspreiding van onderzoeksresultaten. **Deel II**, *‘Challenges related to translation of evidence in decision-making’*, onderzoekt de rol van en behoefte aan real-world evidence (RWE, wetenschappelijk bewijs verkregen uit routinematig verzamelde zorgdata, zoals elektronische patiëntendossiers en declaratiegegevens) bij de besluitvorming over markttoelating en vergoeding van medische producten. Hieronder worden de belangrijkste bevindingen uit het proefschrift samengevat.

Samenvatting van de belangrijkste bevindingen

Deel I van dit proefschrift begint met **Hoofdstuk 2**, waarin het belang van degelijke steekproefberekeningen en veelvoorkomende valkuilen werd besproken. Steekproefberekeningen vormen een essentieel onderdeel van het studieontwerp en bepalen hoeveel deelnemers nodig zijn om een betrouwbaar resultaat te verkrijgen. Een te kleine steekproef kan leiden tot onnauwkeurige en onduidelijke resultaten, terwijl een te grote steekproef kan resulteren in verspilling van middelen en onnodige blootstelling van deelnemers aan minder effectieve behandelingen of mogelijke bijwerkingen. Overmatig optimisme bij het kiezen van parameters voor de steekproefberekening (ook wel ‘optimism bias’) kan leiden tot te kleine studies met een verhoogde kans op nietszeggende uitkomsten.⁶ Zelfs als een effect wordt waargenomen, bestaat er een grote kans dat dit effect is overschat, wat een misleidend beeld kan geven.^{7,8} **Hoofdstuk 3** bouwt hierop voort, waarin werd onderzocht in hoeverre dit optimisme voorkomt in onderzoeksprotocollen van gerandomiseerde klinische interventiestudies (RCTs) ingediend bij een medisch-ethische toetsingscommissie (METC), en welke rol de METC speelde bij het beoordelen van steekproefberekeningen en het potentieel beperken van deze bias. Daarbij werden de opmerkingen van de commissie over de steekproefberekeningen geanalyseerd, en werden de veronderstelde effectgroottes uit de protocollen vergeleken met de daadwerkelijk geobserveerde effecten in de gepubliceerde resultaten van dezelfde studies. We vonden dat de veronderstelde effectgroottes in steekproefberekeningen vaak te optimistisch waren: in circa 80% van de studies bleek het waargenomen effect kleiner dan aangenomen in de oorspronkelijke berekening. Daarnaast bleek dat ongeveer in de helft van de gevallen de steekproefberekening niet volledig werd gerapporteerd, waarbij gekozen parameters en daarbij behorende onderbouwingen ontbraken. Desondanks bleek dat slechts ongeveer één op de vijf studies opmerkingen van de commissie ontving over de steekproefberekening, en dat deze opmerkingen meestal betrekking hadden op rekenfouten in plaats van op de aannemelijkheid van de gekozen effectgrootte. In de enkele gevallen waarin de steekproefberekening tijdens de ethische toetsing werd aangepast, leidde dit slechts in één geval tot een relevante verlaging van de veronderstelde effectgrootte. Opvallend was dat door de industrie gesponsorde studies nauwelijks overschatting vertoonden, in tegenstelling tot door academische onderzoekers geïnitieerde studies.

Hoofdstuk 4 richt zich op een ander belangrijk probleem dat bijdraagt aan onderzoeksverspilling (zogenaamde ‘research waste’): het voortijdig stopzetten van klinische studies. Net als studies met een te kleine steekproefgrootte, leveren ook voortijdig beëindigde studies vaak beperkt bruikbare kennis op. In dit hoofdstuk werden mogelijke voorspellers van voortijdige beëindiging van klinische trials onderzocht op het moment van de ethische toetsing. We vonden dat ongeveer één derde van de klinische studies voortijdig werd beëindigd, waarbij rekruteringsproblemen de meest voorkomende oorzaak vormden. Verschillende kenmerken bleken geassocieerd met voortijdige beëindiging, waaronder een multicenteropzet en als academische onderzoekers de studie initieerden. Dit de laatste bleek specifiek gerelateerd aan rekruteringsproblemen. Mogelijk beschikken academische onderzoekers over minder middelen en gespecialiseerd personeel om de werving te optimaliseren, en schatten zij de haalbaarheid van inclusie te optimistisch in door een gebrek aan ervaring of pilotdata. Opvallend was dat elementen uit de ethische toetsingsprocedure, zoals het aantal en type opmerkingen van de toetsingscommissie, voorspellend bleken voor voortijdige beëindiging van studies. Dit suggereert dat de ethische toetsing zou kunnen fungeren als een voortijdig waarschuwingssysteem om studies met een hoog risico op mislukking te identificeren, waardoor preventieve maatregelen, zoals aanvullende checks gericht op haalbaarheid of aanpassingen in het studieontwerp, mogelijk zijn.

Hoofdstuk 5 onderzocht de mate waarin onderzoeksresultaten, in het bijzonder van kleine interventiestudies (150 deelnemers of minder), worden gepubliceerd en gedeeld, een andere cruciale schakel in het voorkomen van onderzoeksverspilling. Dergelijke studies, die vaak door academische onderzoekers worden geïnitieerd en beperkte middelen hebben, zijn vaker van lagere kwaliteit en lopen mogelijk een groter risico op niet-publicatie. We vonden dat 30% van de kleine studies resultaten niet openbaar maakten, noch via een peer-reviewed tijdschrift, noch via een online trialregister. De kans om te publiceren nam bovendien sterk af vijf jaar na voltooiing van de studie. De mediane tijd tot publicatie bedroeg 3,4 jaar, wat aanzienlijk langer is dan de 2,1 jaar die werd gevonden in een recente meta-analyse van studies van uiteenlopende steekproefomvang.⁹ Opvallend was dat kleine trials die er niet in slaagden hun geplande aantal deelnemers te includeren, een verhoogd risico hadden op niet-publicatie. Deze bevindingen illustreren hoe problemen in een vroeg stadium van het onderzoeksproces, zoals bij de planning en uitvoering, gevolgen kunnen hebben voor latere fasen, waaronder de rapportage en verspreiding van resultaten. Dit onderstreept opnieuw het belang van zorgvuldig ontwerp en uitvoering van klinische studies om onderzoeksverspilling te beperken.

In **Deel II** van dit proefschrift verschuift de focus van het ontwerp en de uitvoering van klinisch onderzoek naar de toepassing van wetenschappelijk bewijs in besluitvorming. Dit deel richt zich op de groeiende rol van RWE bij besluiten over markttoelating en vergoeding van medische producten. Regulatorische instanties, zoals het Europees Geneesmiddelen Agentschap (EMA) en de Amerikaanse Food and Drug Administration (FDA), beoordelen de werkzaamheid en veiligheid van geneesmiddelen bij de beslissing over markttoelating. Gezondheidseconomische beoordelingsinstanties (‘health technology assessment’, HTA), zoals Zorginstituut Nederland, beoordelen daarentegen de relatieve effectiviteit (bijvoorbeeld de toegevoegde waarde van een middel in vergelijking met bestaande middelen), veiligheid en kosteneffectiviteit van medische producten om te bepalen of deze in aanmerking komen voor vergoeding binnen het zorgstelsel. Steeds vaker wordt RWE, naast gegevens uit traditionele RCTs, ingediend bij zowel regulatorische instanties als HTA-organisaties ter ondersteuning van deze besluiten.¹⁰ Toch is onduidelijk wanneer RWE als noodzakelijk wordt beschouwd en hoe het moet worden gewogen ten opzichte van bewijs uit traditionele RCTs in het besluitvormingsproces.¹¹ In **Hoofdstukken 6 en 7** werden, op basis van literatuuronderzoek en interviews met verschillende belanghebbenden, factoren geïdentificeerd die de behoefte aan RWE in regulatorische en HTA-besluitvorming vergroten. De noodzaak van RWE bleek af te hangen van twee pijlers: (1) de aard van de vragen die beantwoord moeten worden om besluitvorming te ondersteunen, en (2) contextuele factoren zoals de haalbaarheid en ethiek van traditionele RCTs, beperkingen van het

beschikbare bewijs, en ziekte-specifieke factoren. De contextuele factoren bleken grotendeels overeen te komen tussen besluitvorming over markttoelating en vergoeding, maar de aard van de benodigde informatie verschilde. Voor besluiten over vergoeding is doorgaans een breder spectrum aan vragen relevant, zoals over de effectiviteit van nieuwe medische producten in vergelijking met bestaande behandelingen, kosten, implementatie en kwaliteit van zorg. Doordat deze vragen doorgaans niet allen beantwoord kunnen worden met bewijs uit traditionele RCTs, is de behoefte aan RWE voor HTA-organisaties vaak groter dan voor regulatoire instanties. Deze bevindingen kunnen bijdragen aan een beter begrip van wanneer en hoe RWE het meest waardevol is in besluitvorming, en kunnen helpen om het proces van de totstandkoming van wetenschappelijk bewijs te stroomlijnen, zodat zowel regulatoire als HTA besluitvorming effectiever wordt ondersteund.

Conclusie

Dit proefschrift onderzocht verschillende knelpunten binnen de medische wetenschap die de kwaliteit van bewijs voor besluitvorming in de gezondheidszorg nadelig beïnvloeden. Daarbij werd gekeken naar meerdere fasen van de onderzoeksketen; van de opzet en uitvoering van studies tot de vertaling van onderzoeksresultaten in besluitvormingsprocessen. Er werd gevonden dat behandelingseffecten in steekproefgrootteberekeningen vaak worden overschat, dat voortijdige beëindiging van studies – met name als gevolg van wervingsproblemen – veel voorkomt, en dat onderzoeksresultaten van kleine studies regelmatig ongepubliceerd blijven. Aangezien METCs studieprotocollen beoordelen voordat een onderzoek van start gaat, hebben zij de mogelijkheid om de uitkomst en kwaliteit van studies wezenlijk te beïnvloeden. De bevindingen in dit proefschrift suggereren dat METCs een actievere rol zouden kunnen spelen bij het toetsen van de aannemelijkheid van veronderstelde effectgroottes in steekproefgrootteberekeningen en het stimuleren van de verspreiding van onderzoeksresultaten. Daarnaast lijken kenmerken van het toetsingsproces voorspellend te zijn voor voortijdige beëindiging van studies, wat de mogelijkheid biedt om onderzoeken met een hoog risico op mislukking al in een vroeg stadium te identificeren. Voortbouwend op de bevindingen in dit proefschrift, kunnen toekomstige studies zich richten op de ontwikkeling van voorspellende modellen voor voortijdige studiebeëindiging, waarmee gerichte preventieve maatregelen kunnen worden genomen om dit risico te verkleinen. Tot slot werden in dit proefschrift factoren geïdentificeerd die de wens en noodzaak voor RWE bepalen voor de besluitvorming over markttoelating en vergoeding van medische producten. Deze inzichten kunnen helpen om de bewijsvereisten van besluitvormers te verduidelijken en kunnen dienen als een eerste stap richting een raamwerk voor systematische beoordeling van heterogene vormen van wetenschappelijk bewijs.

Referenties

1. Evidence-Based Medicine Working G. Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA*. Nov 4 1992;268(17):2420-5. doi:10.1001/jama.1992.03490170092032
2. Dickersin K, Straus SE, Bero LA. Evidence based medicine: increasing, not dictating, choice. *BMJ*. Jan 6 2007;334 Suppl 1:s10. doi:10.1136/bmj.39062.639444.94
3. Bornmann L, Haunschild R, Mutz R. Growth rates of modern science: a latent piecewise growth curve approach to model publication numbers from established and new literature databases. *Humanities and Social Sciences Communications*. 2021/10/07 2021;8(1):224. doi:10.1057/s41599-021-00903-w
4. Ioannidis JP. Why most published research findings are false. *PLoS Med*. Aug 2005;2(8):e124. doi:10.1371/journal.pmed.0020124
5. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet*. Jul 4 2009;374(9683):86-9. doi:10.1016/S0140-6736(09)60329-9
6. Djulbegovic B, Kumar A, Magazin A, et al. Optimism bias leads to inconclusive results-an empirical study. *J Clin Epidemiol*. Jun 2011;64(6):583-93. doi:10.1016/j.jclinepi.2010.09.007
7. Ioannidis JP. Why most discovered true associations are inflated. *Epidemiology*. Sep 2008;19(5):640-8. doi:10.1097/EDE.0b013e31818131e7
8. van Zwet E, Gelman A, Greenland S, Imbens G, Schwab S, Goodman SN. A New Look at P Values for Randomized Clinical Trials. *NEJM Evid*. Jan 2024;3(1):EVIDoa2300003. doi:10.1056/EVIDoa2300003
9. Showell MG, Cole S, Clarke MJ, DeVito NJ, Farquhar C, Jordan V. Time to publication for results of clinical trials. *Cochrane Database Syst Rev*. Nov 27 2024;11(11):MR000011. doi:10.1002/14651858.MR000011.pub3
10. Eichler HG, Pignatti F, Schwarzer-Daum B, et al. Randomized Controlled Trials Versus Real World Evidence: Neither Magic Nor Myth. *Clin Pharmacol Ther*. May 2021;109(5):1212-1218. doi:10.1002/cpt.2083
11. Jaksa A, Wu J, Jonsson P, Eichler HG, Vitito S, Gatto NM. Organized structure of real-world evidence best practices: moving from fragmented recommendations to comprehensive guidance. *J Comp Eff Res*. Jun 2021;10(9):711-731. doi:10.2217/cer-2020-0228

