

Leren en verbeteren op basis van uitkomstinformatie



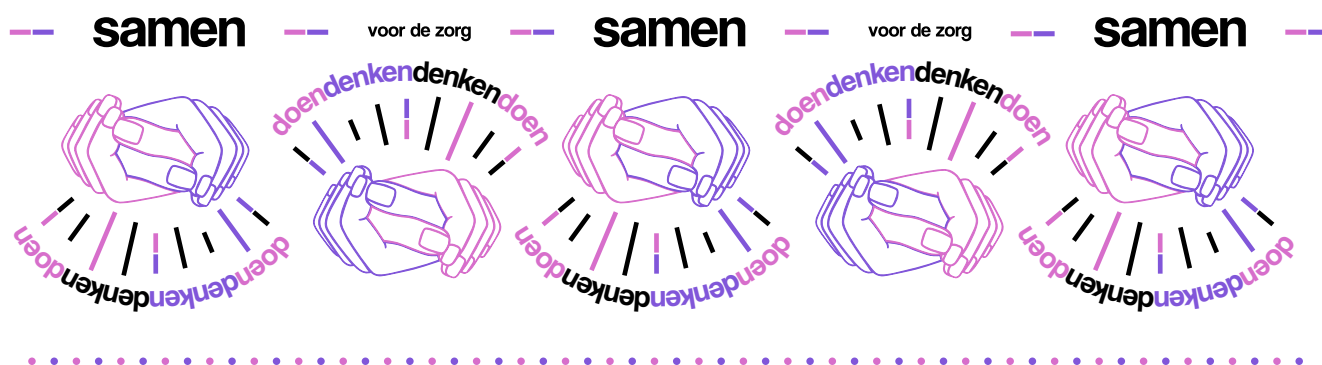
Linnean

Inhoud

1. Inleiding	5
2. Definities	6
a. Wat is een kwaliteitsindicator?	6
b. Wat is een uitkomstindicator?	6
c. Wat zijn PROs en PROMs?	6
d. Wat zijn klinische uitkomsten?	6
e. Wat is een procesindicator?	7
f. Wat is een structuurindicator?	7
g. Wat is validiteit?	7
h. Wat is betrouwbaarheid?	8
i. Wat bedoelen we met 'leren en verbeteren'?	8
J. Waar kan ik meer lezen over de basisprincipes van kwaliteit en waardegedreven zorg?	8
3. Welke factoren beïnvloeden de validiteit van uitkomstinformatie bij leren en verbeteren?	9
a Casemix	9
a.1 Wat is de 'casemix' van een zorginstelling?	9
a.2 Waarom is casemix belangrijk?	9
a.3 Wat is casemix-correctie?	9
a.4 Voor welke casemixvariabelen corrigeer je een uitkomst?	9
a.5 Voor welke casemixvariabelen moet je een uitkomst niet corrigeren?	10
a.6 Corrigeer je voor dezelfde casemix variabelen bij verschillende uitkomsten?	10
a.7 Hoe kan men uitkomsten corrigeren voor casemix?	10
a.8 Hoe kan men de resultaten van een casemix gecorrigeerde uitkomst visualiseren?	11
a.9 Voor hoeveel casemixvariabelen kan men een uitkomst corrigeren?	12
a.10 Hoe om te gaan met missende waarden?	12
a.11 Lost casemixcorrectie confounding (verstoring) door casemixvariabelen op?	13
a.12 Waar kan ik meer lezen over casemix?	13
b Registratiebias	14
b.1 Wat is registratiebias?	14
b.2 Hoe ontstaat en voorkomt men registratiebias bij PROMs?	14
b.3 Hoe ontstaat en voorkom je registratiebias bij klinische uitkomstmetingen?	15
b.4 Gestandaardiseerde terminologieën en coderingen	16

Inhoud

b.5	Welke landelijke initiatieven helpen registratiebias voorkomen?	16
b.6	Waar kan ik meer lezen over registratiebias?	17
c	Selectiebias	17
c.1	Wat is Selectiebias	18
c.2	Hoe ontstaat selectiebias bij PROMs?	18
c.3	Hoe voorkom je selectiebias bij PROMs?	18
c.4	Hoe ontstaat selectiebias bij klinische uitkomsten?	18
c.5	Hoe voorkom je selectiebias bij klinische uitkomsten?	19
c.6	Waar kan ik meer lezen over selectiebias	19
d	Relatie proces/structuur	19
d.1	Belang van relatie tussen proces/structuur variabelen - en uitkomstindicatoren	19
d.2	Invloed nabijheid van het proces op de uitkomst	20
d.3	Patiëntniveau-data nodig voor de relatie tussen proces en uitkomst	20
d.4	Meer lezen over de relatie tussen proces en uitkomst?	20
4	Welke factoren beïnvloeden de betrouwbaarheid van uitkomstinformatie bij leren en verbeteren?	21
a	Meetfouten	21
a.1	Wat is een meetfout?	21
a.2	Hoe beïnvloeden meetfouten de interpretatie van uitkomstinformatie?	21
a.3	Hoe ontstaan en voorkom je meetfouten bij klinische uitkomsten?	21
a.4	Hoe ontstaan en voorkom je meetfouten bij PROMs?	22
a.5	Waar kan ik meer lezen over meetfouten?	23
b	Steekproefonzekerheid	23
b.1	Wat is 'steekproefonzekerheid'?	23
b.2	Hoe beïnvloedt de steekproefgrootte de steekproefonzekerheid?	23
b.3	Wat zijn betrouwbaarheidsintervallen?	23
b.4	Hoe gebruik je deze bij het rapporteren van uitkomstinformatie?	24
b.5	Hoe interpreteer je betrouwbaarheidsintervallen bij kwaliteitsverbetering?	24
b.6	Waar kan ik meer lezen over steekproefonzekerheid?	24



1. Inleiding

Dagelijks wordt veel informatie verzameld over de uitkomsten van patiënten die medisch specialistische zorg ondergaan. Steeds vaker wordt deze uitkomstinformatie, in combinatie met relevante proces- en structuurinformatie, gebruikt om inzicht te krijgen in kwaliteit van zorg. Deze informatie kan enerzijds worden gebruikt om kwaliteit binnen de eigen organisatie te evalueren, en anderzijds om zorgaanbieders met elkaar te vergelijken.

Deze vergelijkingen stellen ziekenhuizen, zorgverzekeraars, patiëntenorganisaties en overheidsinstanties in staat om inzicht te krijgen in kwaliteitsverschillen en mogelijkheden voor verbetering.

Om te kunnen leren en verbeteren op basis van uitkomstinformatie is het belangrijk om te begrijpen hoe men deze informatie op een correcte en verantwoorde manier kan gebruiken. In dit zakboekje geven we:

- een overzicht van factoren die de validiteit en betrouwbaarheid van uitkomstinformatie kunnen aantasten;
- handvatten voor professionals om deze invloeden te herkennen en te verminderen.

Het zakboekje is bedoeld voor zorgprofessionals die werken met uitkomstinformatie, maar ook voor beleidsmakers, programmaleiders en partijen die (van plan zijn om) systemen en infrastructuren ontwikkelen voor het verzamelen van deze informatie, zoals kwaliteitsregistraties. Het doel is om zowel bij het gebruik als bij de inrichting van gegevensverzameling bewuste keuzes te kunnen maken. Dit met aandacht voor de methodologische valkuilen en randvoorwaarden die nodig zijn voor betrouwbare en betekenisvolle inzichten.

2. Definities

a. Wat is een kwaliteitsindicator?

Een kwaliteitsindicator is een gestandaardiseerd instrument om de kwaliteit van zorg te meten. Kwaliteitsindicatoren zijn onder te verdelen in uitkomst-, proces- en structuurindicatoren. In deze sectie lichten we deze drie typen indicatoren afzonderlijk toe.

.....

b. Wat is een uitkomstindicator

Een uitkomstindicator is een kwaliteitsindicator die het effect van de geleverde zorg op de gezondheid van de patiënt meet. Mortaliteit (het percentage patiënten dat overlijdt binnen een bepaalde periode na een medische behandeling, opname of diagnose) en kwaliteit van leven zijn voorbeelden van uitkomsten die men kan gebruiken als uitkomstindicator. Denk bij mortaliteit aan het percentage patiënten dat binnen 30 dagen na een operatie aan een primaire dikke darmtumor overlijdt. Uitkomstindicatoren meet men vaak op patiëntniveau.

.....

c. Wat zijn PROs en PROMs?

Patiënt-gerapporteerde uitkomsten (Patient Reported Outcomes ofwel PROs) zijn aspecten van gezondheid die de patiënt zelf rapporteert. PROs worden bijna altijd in kaart gebracht met vragenlijsten (Patient Reported Outcome Measures ofwel PROMs) die meten hoe patiënten hun eigen gezondheid ervaren. PROMs kunnen verschillende aspecten van gezondheid meten, zoals pijn en fysiek functioneren en kan men inzetten als uitkomstindicator.

.....

d. Wat zijn klinische uitkomsten?

Een klinische uitkomst is een uitkomst die een zorgverlener of zorgorganisatie meet of rapporteert. Binnen dit begrip vallen bijvoorbeeld biomarkers, administratieve gegevens (zoals heropnames, ligduur of complicaties) en door zorgverleners gerapporteerde metingen (zoals bloeddruk, longfunctie of hartfunctie). Klinische uitkomsten kan men gebruiken als uitkomstindicator.

.....

e. Wat is een procesindicator?

Een procesindicator is een kwaliteitsindicator die meet in welke mate men zorghandelingen volgens geldende standaarden en richtlijnen uitvoert. Een voorbeeld hiervan is de wachttijd vanaf diagnose tot operatie. Procesindicatoren verzamelt men op patiëntniveau.

.....

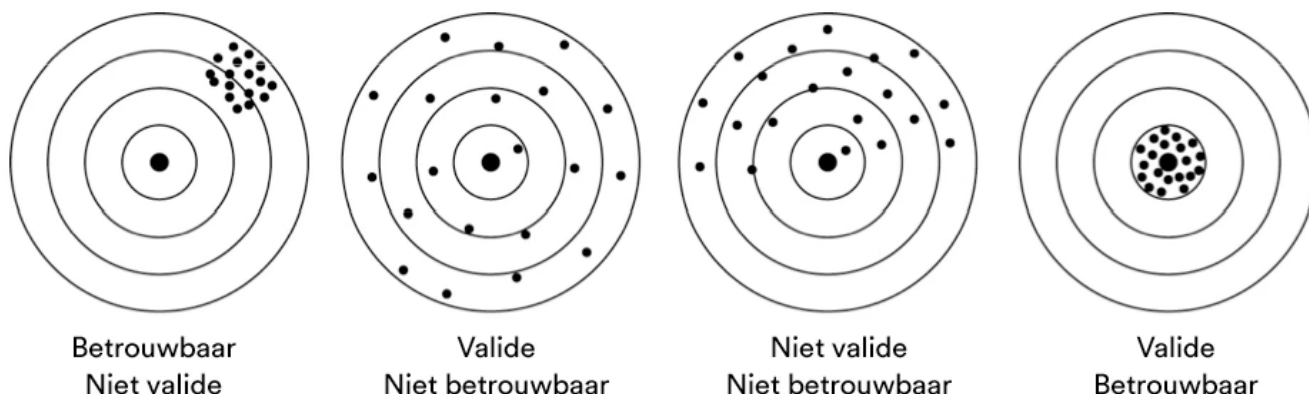
f. Wat is een structuurindicator?

Een structuurindicator is een kwaliteitsindicator die de organisatorische voorwaarden en randvoorwaarden meet die nodig zijn om kwalitatief goede zorg te kunnen leveren. Voorbeelden hiervan zijn onder andere; ziekenhuisvolume (aantal operaties/behandelingen per aandoening), of de samenstelling van een behandelteam. In het laatste geval kan dit voor bijvoorbeeld borstkankerchirurgie betekenen welke specialisten onderdeel uitmaken van het behandelteam (bijv. medisch oncoloog, chirurg, plastisch chirurg). Structuurindicatoren meet men vaak op instellingsniveau.

.....

g. Wat is validiteit?

Validiteit is de mate waarin een uitkomstindicator vrij is van systematische meetfouten of bias. In andere woorden, validiteit geeft aan in hoeverre een indicator daadwerkelijk meet wat deze beoogt te meten. In onderstaande afbeelding zie je dit terug in de afstand van de metingen (stippen) tot het doel, de roos in het midden. Hoe dichters de metingen bij het midden liggen, hoe meer valide de meting, ongeacht of de metingen gegroepeerd of verspreid zijn. De aanwezigheid van systematische meetfouten of bias zorgt ervoor dat men de werkelijke waarde van een uitkomst systematisch over- of onderschat.



h. Wat is betrouwbaarheid?

Betrouwbaarheid is de mate waarin een uitkomstindicator vrij is van willekeurige meetfouten. Een betrouwbare indicator produceert consistente resultaten bij herhaalde metingen onder vergelijkbare omstandigheden. Wanneer bijvoorbeeld een collega dezelfde meting uitvoert bij dezelfde patiënt, zou dit moeten leiden tot hetzelfde resultaat. In de afbeelding zie je betrouwbaarheid terug in hoe dicht de metingen (stippen) bij elkaar liggen: hoe kleiner de spreiding tussen de metingen, hoe betrouwbaarder de indicator. Ongeacht of ze het midden van het doel raken.

.....

i. Wat bedoelen we met ‘leren en verbeteren’?

Leren en verbeteren is elk systematisch proces waarbij zorgprofessionals en zorgorganisaties, op basis van onder andere inzichten uit data over kwaliteit van zorg, de geleverde zorg continu optimaliseren.

.....

j. Waar kan ik meer lezen over de basisprincipes van kwaliteit en waardegedreven zorg?

- Donabedian, A. (1988). The quality of care: How can it be assessed? JAMA, 260(12), 1743 –1748.
 - Porter, M. E. (2010). What is value in health care? New England Journal of Medicine, 363(26), 2477–2481.
 - Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H.C. W. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties: A clarification of its content. BMC Medical Research Methodology, 10, 22.
-

3. Welke factoren beïnvloeden de validiteit van uitkomstinformatie bij leren en verbeteren?

a. Casemix

a.1 Wat is de 'casemix' van een zorginstelling?

Het begrip casemix verwijst naar de kenmerken van de patiënten die bij een zorginstelling een specifieke behandeling ondergaan. Het gaat hierbij om zowel demografische (zoals leeftijd) als klinische kenmerken (zoals patiënten met meer dan één aandoening of de ernst van de ziekte)

a.2 Waarom is casemix belangrijk?

De uitkomsten die een zorginstelling bij een behandeling gemiddeld behaalt, worden niet alleen beïnvloed door de kwaliteit van de geleverde zorg, maar ook door het type patiënten dat de instelling behandelt. Mensen met ernstige klachten krijgen vaak hun behandeling in gespecialiseerde instellingen. Patiënten met ernstige klachten hebben vaak ook minder goede uitkomsten. Als men geen rekening houdt met verschillen in casemix, blijft het onduidelijk of waargenomen verschillen tussen instellingen komen door verschillen in zorgkwaliteit of de samenstelling van de patiëntengroep.

a.3 Wat is casemix-correctie?

Casemix-correctie is een statistische procedure die men toepast om de vergelijkbaarheid van zorguitkomsten van zorginstellingen te verbeteren door rekening te houden met variaties in casemix. Deze techniek streeft ernaar om de invloed van verschillen in patiëntpopulaties, zoals leeftijd en de aanwezigheid van comorbiditeiten (patiënten met meer dan één aandoening), op de resultaten te verminderen.

a.4 Voor welke casemixvariabelen corrigeer je een uitkomst?

Het is belangrijk om te corrigeren voor casemixvariabelen die 1) gerelateerd zijn aan de uitkomst en 2) variëren tussen de instellingen. Men hoeft alleen te corrigeren voor variabelen die aan beide voorwaarden voldoen; Zijn casemixvariabelen niet gerelateerd aan de uitkomst? Dan

dragen ze ook niet bij aan de uitkomstvariatie en kunnen dus geen meetafwijking veroorzaken. Hebben casemixvariabelen wel een relatie met de uitkomst maar variëren ze niet tussen de instellingen? Dan beïnvloeden ze de uitkomsten in iedere instelling in dezelfde mate, waardoor ze de vergelijkbaarheid niet verstoren.

.....

a.5 Voor welke casemixvariabelen moet je een uitkomst niet corrigeren?

Variabelen die acties weergeven die de zorgverlener onderneemt (procesvariabelen) of die de middelen weergeven die een zorgverlener ter beschikking staan (structuurvariabelen), behoren niet tot de casemix van een instelling. Het is belangrijk om dit soort variabelen niet op te nemen in een casemixcorrectiemodel. Dit omdat ze de kwaliteit van de geleverde zorg kunnen beïnvloeden en dus van belang zijn voor de uitkomstinformatie.

.....

a.6 Corrigeer je voor dezelfde casemix variabelen bij verschillende uitkomsten?

Op basis van onderzoek is bekend dat uitkomsten binnen een aandoening verschillende risicofactoren kunnen hebben. Het is daarom goed om per uitkomstindicator te kijken voor welke variabelen men moet corrigeren. Dit is vooral van belang als de indicatoren betrekking hebben op verschillende behandelingen, verschillende patiëntenpopulaties, of als men een mix van PROMs en klinische uitkomsten gebruikt.

.....

a.7 Hoe kan men uitkomsten corrigeren voor casemix?

Er zijn verschillende methoden om de gemiddelde geobserveerde scores per ziekenhuis (O_i) te corrigeren voor casemix. Toch maakt men vaak gebruik van indirecte standaardisatie via een regressiemodel. Dit gebeurt met de casemixvariabelen als onafhankelijke variabelen en de uitkomst als afhankelijke variabele. Indirecte standaardisatie verloopt via drie stappen:

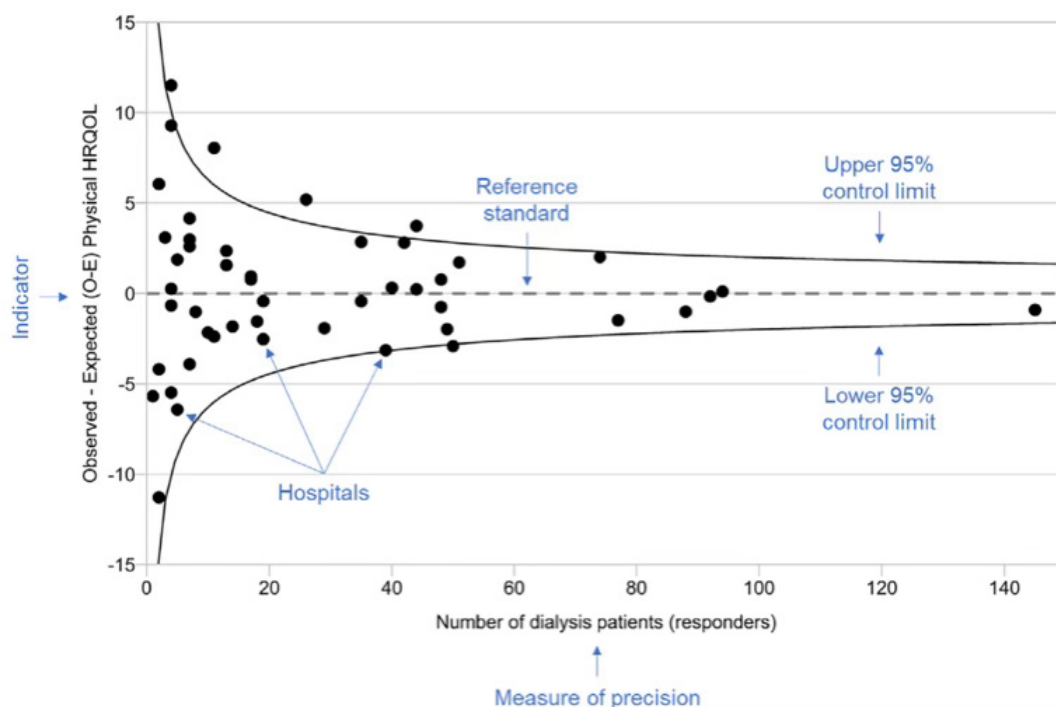
1. Het berekenen van de verwachte (voorspelde) behandeluitkomsten voor ieder ziekenhuis. Hierbij maakt men gebruik van het regressiemodel met daarin de belangrijke casemixvariabelen. Deze verwachte uitkomst geeft men vaak weer als (E_i) (en dit geeft dus de verwachte uitkomst van ziekenhuis i weer).
2. Het vergelijken van de geobserveerde (O_i) en verwachte (E_i) ziekenhuisscores.
3. Het extrapoleren van de prestatie van zorgaanbieder i naar bijvoorbeeld het landelijk gemiddelde.

Afhankelijk van het type uitkomst, kan men vervolgens een gestandaardiseerde score verkrijgen door het verschil $(O_i) - (E_i)$ te berekenen voor continue uitkomsten zoals PROM-scores, of de ratio $(O_i)/(E_i)$ voor binaire uitkomsten, zoals het percentage patiënten met een gecompliceerd beloop. Deze gestandaardiseerde scores kan men, indien gewenst, schalen rondom een referentiestandaard (R), zoals de gemiddelde indicatorscore over alle patiënten. Voor een continue uitkomstindicator zou de aanpassing voor de gestandaardiseerde score bijvoorbeeld zijn: $O_i - E_i + R$

a.8 Hoe kan men de resultaten van een casemix gecorrigeerde uitkomst visualiseren?

Een veelgebruikte methode om casemix gecorrigeerde indicatorscores te presenteren is de funnel plot. In de afbeelding is zo'n funnel plot weergegeven. Op de X-as plaatst men een maat voor volume, in dit voorbeeld is dit het aantal dialysepatiënten. De Y-as geeft de gestandaardiseerde indicatorscore weer, in dit geval is dat de $(O_i) - (E_i)$ score van "physical health related quality of life". In de plot geeft men een referentiestandaard weer (zie afbeelding). In dit geval is dit op de nullijn, want als de geobserveerde score (O_i) precies de verwachte score (E_i) is, is het verschil 0.

Aan beide kanten van deze referentielijn voegt men controlelimieten toe die 95% betrouwbaarheidsintervallen vertegenwoordigen. Instellingen buiten deze limieten presteren opvallend anders dan de referentiestandaard. Dat zou kunnen duiden op buitengewoon goede of slechte kwaliteit van zorg. De funnel plot heet zo omdat de spreiding afneemt naarmate de steekproefgrootte toeneemt, dat lijkt op een trechtervorm ("funnel"). Een funnel plot kan men ook maken voor een ongecorrigeerde uitkomst. In dat geval relateert men de geobserveerde scores vaak aan de gemiddelde indicatorscore over de ziekenhuizen.



a.9 Voor hoeveel casemixvariabelen kan men een uitkomst corrigeren?

Het aantal variabelen dat men in een casemixcorrectiemodel voor een binaire uitkomst kan opnemen, wordt in praktische zin vooral beperkt door de 'event rate'. De event rate verwijst naar hoeveel patiënten de uitkomst hebben (bijvoorbeeld het percentage patiënten in een ziekenhuis dat complicaties na een operatie ervaart). Men kan ook te veel casemixvariabelen toevoegen ten opzichte van het aantal events. Dan neemt het risico toe dat de regressiecoëfficiënten de ruis in de dataset weerspiegelen, in plaats van de daadwerkelijke verbanden in de populatie (overfitting). Een veel aangehouden vuistregel om dit te voorkomen is om minimaal 10 events per casemixvariabele aan te houden. In recente wetenschappelijke literatuur benadrukt men echter dat ook andere factoren een rol spelen, waardoor men een formele 'poweranalyse' kan overwegen. Voor continue uitkomsten, zoals PROM-scores, houdt men vaak 10 patiënten per casemixvariabele aan als vuistregel.

a.10 Hoe om te gaan met missende waarden?

Een eenvoudige maar vaak ongewenste benadering is om alle patiënten met ontbrekende waarden te verwijderen (listwise deletion). Tenzij de missende waarden om volstrekt willekeurige redenen ontbreken (MCAR, missing completely at random), zorgt deze benadering namelijk voor bias. Zelfs als er sprake is van MCAR leidt de toepassing van deze methode tot verlies van onderscheidend vermogen (power). Om bias door ongeobserveerde variabelen te voorkomen kan men missende waarden imputeren. Dit houdt in dat men gebruik maakt van een regressiemodel om de missende waarden te voorspellen. Omdat het voorspellen van ontbrekende waarden extra onzekerheid met zich meebrengt, voert men deze zogenaamde imputatie meerdere keren uit.

Een veelgebruikte richtlijn is om het aantal imputaties af te stemmen op het percentage missende data. Bij bijvoorbeeld 20% aan ontbrekende waarden adviseert men om minste 20 keer te imputeren. Vervolgens kan men casemix-regressiemodellen apart toepassen op elk van deze geïmputeerde datasets (in dit voorbeeld dus 20 keer). Waarna men de resultaten volgens de regels van Rubin combineert tot één set uitkomsten. In het poolen houdt men rekening met de extra onzekerheid door de geschatte standaardfout aan te passen. Veel moderne statistische software biedt mogelijkheden om imputaties te kunnen uitvoeren. Ook is het mogelijk om de extra onzekerheid, door missende waarden, op een verantwoorde manier mee te nemen in de analyses.

a.11 Lost casemixcorrectie confounding (verstoring) door casemixvariabelen op?

Een belangrijke beperking van casemixcorrectie is dat men alleen kan corrigeren voor de variabelen die daadwerkelijk beschikbaar zijn en zijn opgenomen in het model. Dit betekent dat ongemeten of onbekende confounders nog steeds een vertekend beeld kunnen geven van de resultaten. Zelfs als men betrouwbare verschillen waarneemt tussen instellingen, is het daarom nooit met zekerheid te zeggen of deze verschillen in uitkomsten worden veroorzaakt door variaties in kwaliteit van geleverde zorg of door residuele confounding.

a.12 Waar kan ik meer lezen over casemix?

- Iezzoni, L. I. (Ed.). (2003). Risk Adjustment for Measuring Health Care Outcomes (Vol. 3). Health Administration Press, Chicago.
- Amini, M., Eijkenaar, F., Lingsma, H. F., den Hartog, S. J., Olthuis, S. G. H., Martens, J., van der Worp, B., van Zwam, W., van der Hoorn, A., Roosendaal, S. D., Roozenbeek, B., Dippel, D., & van Leeuwen, N.; MR CLEAN Registry Investigators. (2023). Validity of Early Outcomes as Indicators for Comparing Hospitals on Quality of Stroke Care. *Journal of the American Heart Association*, 12(8), e027647. <https://doi.org/10.1161/JAHA.122.027647>
- Fischer, C., Lingsma, H. F., van Leersum, N., Tollenaar, R. A., Wouters, M. W., & Steyerberg, E. W. (2015). Comparing Colon Cancer Outcomes: The Impact of Low Hospital Case Volume and Case-Mix Adjustment. *European Journal of Surgical Oncology*, 41(8), 1045-1053. <https://doi.org/10.1016/j.eso.2015.04.009>
- Oemrawsingh, A., van Leeuwen, N., Venema, E., Limburg, M., de Leeuw, F. E., Wijffels, M. P., de Groot, A. J., Hilken, P. H. E., Hazelzet, J. A., Dippel, D. W. J., Bakker, C. H., Voogdt-Pruis, H. R., & Lingsma, H. F. (2019). Value-Based Healthcare in Ischemic Stroke Care: Case-Mix Adjustment Models for Clinical and Patient-Reported Outcomes. *BMC Medical Research Methodology*, 19(1), 229. <https://doi.org/10.1186/s12874-019-0864-z>
- Nuttall D, Parkin D, Devlin N (2015) Inter-provider comparison of patient-reported outcomes: developing an adjustment to account for differences in patient case mix. *Health Econ* 24:41–54
- Deel 4 Methodologische uitdagingen van het boek *Behandeluitkomsten: Bron voor kwaliteitsbeleid in de GGZ*. De Beurs, Barendregt & Warmerdam (2017).
- Ding YY. Risk adjustment: towards achieving meaningful comparison of health outcomes in the real world. *Ann Acad Med Singap*. 2009 Jun;38(6):552-7. PMID: 19565108.
- Bouter, L. M., van Dongen, M. C. J. M., & Zielhuis, G. A. (2005). Epidemiologisch onderzoek: opzet en interpretatie. Bohn Stafleu Van Loghum.
- van Buuren, S. (2018). *Flexible Imputation of Missing Data*, Second Edition (2nd ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/9780429492259>

b. Registratiebias

b.1 Wat is registratiebias?

Registratiebias is een vertekening bij het vergelijken van ziekenhuisprestaties. Deze ontstaat door inconsistenties of variaties in de manier waarop men uitkomstinformatie verzamelt of registreert. De aanwezigheid van registratiebias zorgt ervoor dat men de prestaties van bepaalde ziekenhuizen systematisch over- of onderschat.

.....

b.2 Hoe ontstaat en voorkomt men registratiebias bij PROMs?

Lokale variaties in wijze van afname

PROMs kan men op verschillende manieren afnemen. Dit kan via papieren of digitale vragenlijsten en op diverse locaties en momenten in het zorgtraject. Thuis of in het ziekenhuis, voorafgaand aan de behandeling, tijdens de zorg of na afloop. Onderzoek toont aan dat verschillende wijzen van afname kunnen leiden tot verschillende antwoorden van patiënten. De mate waarin dit tot merkbare vertekening van de PROM-scores leidt, varieert tussen aandoeningen en afnamemethoden. Om registratiebias te voorkomen, is het daarom aan te bevelen de afnamemethode te standaardiseren en consistent toe te passen in alle deelnemende centra.

Hulp van anderen bij invullen vragenlijsten

Soms is het toegestaan dat een zorgverlener of naaste patiënten bij het invullen van een PROM ondersteunen. Of dat deze de hele vragenlijst namens de patiënt invult. Onderzoek toont aan dat hulp van anderen vaak leidt tot vertekening. Vooral als iemand anders de hele vragenlijst invult. Dit kan leiden tot twee vormen van bias: bias op het niveau van individuele items en bias op de totaalscores.

Voor de eerste vorm van bias kan men scoringsprocedures ontwikkelen die hier rekening mee houden. Het tweede type bias beïnvloedt de scores op de hele PROM. Dit kan men daarom beter aanpakken door in de analyse een effect op te nemen voor de invuller. Het is daarom belangrijk om te registreren wie de PROM invult, zodat men deze informatie kan meenemen in de analyse. Zo kan men de vertekening minimaliseren

Gebruik van vragenlijsten in meerdere talen

voorbeeld omdat een deel van de patiënten de Nederlandse taal slecht beheerst. Het is dan belangrijk om waar mogelijk gebruik te maken van officiële, vertaalde versies die zijn ontwikkeld met behulp van adequate cross-culturele adaptatieprotocollen. Deze protocollen

zijn erop gericht om de vragenlijsten zo vergelijkbaar mogelijk te houden met de oorspronkelijke versie van de vragenlijst. Zelfs met deze protocollen is het echter niet gegarandeerd dat de scores daadwerkelijk vergelijkbaar zijn tussen verschillende versies. Daarom is het belangrijk om te kijken naar studies die de cross-culturele validiteit van de vragenlijsten hebben onderzocht. Indien nodig kan men gebruik maken van scoringsprocedures. Deze kan men corrigeren voor culturele bias die wordt veroorzaakt door items waarvoor geen cultureel soortgelijke vertaling beschikbaar is.

Gebruik van verschillende versies van PROMs

Er bestaan vaak meerdere versies van veelgebruikte PROMs. Het komt regelmatig voor dat gebruikers zich niet bewust zijn van de versie van de PROM die ze gebruiken. Omdat versies vaak verschillen qua exacte definitie van de items, aantal items, antwoordopties en scoringsprocedures, kan men de scores tussen verschillende versies niet zomaar vergelijken. Het is daarom essentieel om in elke zorginstelling goed te registreren welke versie men gebruikt. En ervoor te zorgen dat iedereen in alle deelnemende centra dezelfde versie inzet.

Gebruik van verschillende itemcoderingen en/of scoringsregels

Voor sommige PROMs bestaan er meerdere, al dan niet gevalideerde, scoringsprocedures. Daarnaast kan de codering van vragen variëren tussen databronnen. Het gebruik van verschillende itemcoderingen en scoringsregels kan ervoor zorgen dat patiënten verschillende scores krijgen, zelfs als hun antwoorden identiek zijn. Bij het samenvoegen van PROMscores uit meerdere bronnen is het van groot belang om te verifiëren dat elke partij dezelfde itemcodering en scoringsregels volgt van de gebruikte PROM en om data waar nodig te hercoderen.

b.3 Hoe ontstaat en voorkom je registratiebias bij klinische uitkomstmetingen?

Registratie in vrije tekstvelden

Veel uitkomstinformatie leidt men af uit gegevens die men in de dagelijkse zorg ongestructureerd vastlegt in vrije tekstvelden van het Elektronisch Patiëntendossier (EPD). Later moeten anderen, zoals datamanagers of verpleegkundigen, deze data structureren. Daarna kan men deze data aggregeren (samenvoegen) over centra en gebruiken in kwantitatieve analyses. Dit proces kan leiden tot interpretatieverschillen en inconsistenties. Wanneer verschillende mensen deze taak uitvoeren zonder uitgelijnde procedures, kan dat registratiebias tot gevolg hebben. Om deze bias te verminderen, is het belangrijk om gebruik te maken van gestandaardiseerde terminologieën en coderingen, zoals SNOMED CT en LOINC (zie kadertekst), voor directe gestructureerde vastlegging. Als dit niet mogelijk is, moeten datamanagers volgens geharmoniseerde procedures werken om consistentie in de gegevensverzameling te verzekeren. Reguliere kalibratie en audits kunnen hierbij helpen.

Gebruik van lokale standaarden voor gestructureerde vastlegging

Soms legt men uitkomstinformatie weliswaar gestructureerd vast, maar wel volgens lokale coderingssystemen. Dit maakt het moeilijk om data op een uniforme manier te aggregeren en te analyseren. Lokale coderingssystemen kunnen leiden tot inconsistenties en verminderde vergelijkbaarheid van gegevens tussen verschillende instellingen. Om dit probleem op te lossen, is het belangrijk om gebruik te maken van informatiestandaarden zoals SNOMED CT en LOINC.

.....

b.4 Gestandaardiseerde terminologieën en coderingen

SNOMED CT (Systematized Nomenclature of Medicine – Clinical Terms) is een uitgebreide, hiërarchisch opgebouwde verzameling medische termen. Deze standaard maakt het mogelijk om medische concepten, zoals diagnoses, symptomen en verrichtingen, op een gestructureerde en herkenbare manier vast te leggen. Zo kan men bijvoorbeeld “pijn op de borst” of “diabetes mellitus type 2” altijd op dezelfde manier registreren, ongeacht wie het invoert of in welk ziekenhuis het gebeurt.

LOINC (Logical Observation Identifiers Names and Codes) richt zich op het standaardiseren van meetwaarden en observaties, zoals laboratoriumuitslagen of vragenlijstitems. Denk bijvoorbeeld aan bloeddrukmetingen, hemoglobinewaarden of antwoorden op PROs (Patient Reported Outcomes). Door gebruik te maken van LOINC-codes kan men deze gegevens eenduidig interpreteren en vergelijken, ook als ze uit verschillende bronnen komen.

.....

b.5 Welke landelijke initiatieven helpen registratiebias voorkomen?

Er zijn meerdere nationale programma's die focussen op het verminderen van registratiebias in het Nederlandse zorgsysteem. Het programma Uitkomstgerichte Zorg en Registratie aan de bron zijn voorbeelden van dit soort programma's waarvan men de protocollen kan volgen om registratiebias te verminderen.

.....

b.6 Waar kan ik meer lezen over registratiebias?

- Arts, D. G., De Keizer, N. F., & Scheffer, G. J. (2002). Defining and improving data quality in medical registries: A literature review, case study, and generic framework. *Journal of the American Medical Informatics Association*, 9(6), 600–611.
- Kahn, M. G., Raebel, M. A., Glanz, J. M., Riedlinger, K., & Steiner, J. F. (2012). A pragmatic framework for single-site and multisite data quality assessment in electronic health recordbased clinical research. *Medical Care*, 50(Suppl), S21–S29.
- Weiskopf, N. G., & Weng, C. (2013). Methods and dimensions of electronic health record data quality assessment: A systematic review of the literature. *Journal of the American Medical Informatics Association*, 20(1), 129–140.
- Haux, R. (2006). Health information systems – past, present, future. *International Journal of Medical Informatics*, 75(3–4), 268–281.
- Stausberg, J., Breitschdel, L., & Knaup-Gregori, P. (2006). Quality of routine health data: A review of data accuracy, data completeness and data linkage. *Gesundheitswesen*, 68(4), 249–256.
- Benchimol, E. I., et al. (2015). The REporting of studies Conducted using Observational Routinely collected Data (RECORD) Statement. *PLoS Medicine*, 12(10), e1001885.
- Black, N. (2013). Patient reported outcome measures could help transform healthcare. *BMJ*, 346, f167.
- Gwaltney, C. J., Shields, A. L., & Shiffman, S. (2008). Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: A meta-analytic review. *Value in Health*, 11(2), 322–333.
- Muehlhausen, W., Doll, H., Quadri, N., Fordham, B., O'Donohoe, P., Dogar, N., & Wild, D. J. (2015). Equivalence of electronic and paper administration of patient-reported outcome measures: A systematic review and meta-analysis of studies conducted and published between 2007 and 2013. *Journal of Medical Internet Research*, 17(5), e113.
- Efficace, F., Bottomley, A., Coens, C., Van Steen, K., & Ringash, J. (2008). Does proxy reporting affect the quality of patient-reported outcome data in cancer clinical trials? A meta-analysis. *Journal of Clinical Oncology*, 26(29), 4796–4801.
- Wild, D., Grove, A., Martin, M., Eremenco, S., McElroy, S., Verjee-Lorenz, A., & Erikson, P. (2005). Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient Reported Outcomes (PRO) Measures: Report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value in Health*, 8(2), 94–104.

c. Selectiebias

c.1 Wat is selectiebias?

Selectiebias is een vertekening die ontstaat wanneer de groep patiënten waarvoor uitkomstinformatie beschikbaar is niet representatief is voor de gehele populatie die men behandelt. Als bepaalde groepen patiënten systematisch wel of juist niet deelnemen, kunnen de uitkomsten vertekend zijn. Hierdoor kunnen de geobserveerde PROM-scores de werkelijke ervaringen van alle patiënten overschatten of onderschatten.

c.2 Hoe ontstaat selectiebias bij PROMs?

Selectiebias ontstaat met name door een selectieve en/of lage respons, waarbij vooral bepaalde patiëntgroepen (bijvoorbeeld jongere, gezondere of digitaal vaardige patiënten) wel meedoen, en andere groepen niet. Wanneer dit responspercentage bovendien verschilt tussen instellingen, kan dit leiden tot oneerlijke vergelijkingen op het gebied van kwaliteit van zorg. Zo lijkt de ene instelling mogelijk beter te presteren dan de andere. Terwijl men dit verschil ook zou kunnen verklaren door vast te stellen wie er wel of niet vragenlijsten heeft ingevuld.

.....

c.3 Hoe voorkom je selectiebias bij PROMs?

Om selectiebias te voorkomen is het belangrijk om drempels voor deelname te verlagen, bijvoorbeeld door ondersteuning te bieden bij het invullen, vragenlijsten aan te bieden in meerdere talen en zowel digitaal als op papier. Ook is het essentieel om responspercentages per centrum te monitoren en bij voorkeur te streven naar gelijke en voldoende hoge deelname. Hiermee kan de vergelijkbaarheid van de PROM-scores tussen de centra verbeteren.

.....

c.4 Hoe ontstaat selectiebias bij klinische uitkomsten?

Selectiebias bij klinische uitkomsten kan ontstaan wanneer de patiëntengroep waarvan men de uitkomsten analyseert, niet representatief is voor de gehele groep patiënten die de zorg heeft ontvangen. Selectie effecten kunnen op verschillende manieren ontstaan en worden voorkomen:

Selectie-effecten door inclusiecriteria

Soms worden vooraf bepaalde inclusie- en exclusiecriteria gehanteerd, bijvoorbeeld alleen patiënten met een specifieke diagnose of die een volledig behandeltraject hebben doorlopen. Dit is op zich legitiem, maar kan leiden tot selectieve uitkomsten als deze criteria niet goed aansluiten bij de doelpopulatie, of als ze verschillen tussen instellingen. Het is daarom belangrijk om de gekozen criteria duidelijk te definiëren, transparant te rapporteren en te toetsen of ze passend zijn. Dit met het oog op het doel van de analyse en de populatie waarover men uitspraken wil doen.

Onbedoelde selectieve exclusie

Het kan ook voorkomen dat kenmerken van het initiatief ertoe leiden dat bepaalde groepen buitensporig buiten beeld raken of blijven. Een bekend voorbeeld is dat patiënten met beperkte taalvaardigheid of lage gezondheidsvaardigheden minder vaak meedoen aan vragenlijstmetingen. Een ander bekend voorbeeld is dat de complicatiegraad bij ernstige aandoeningen onderschat wordt als mensen uitvallen of overlijden voordat alle complicaties

zijn opgetreden. Ook bij klinische uitkomstmetingen kan dit probleem zich voordoen. Zo wordt de complicatiegraad bij ernstige aandoeningen soms onderschat doordat patiënten uitvallen of overlijden voordat alle complicaties zich hebben kunnen manifesteren of geregistreerd zijn.

c.5 Hoe voorkom je selectiebias bij klinische uitkomsten?

Om selectieve exclusie te voorkomen is het belangrijk om vooraf na te denken over en te anticiperen op barrières voor patiënten om deel te (blijven) nemen. Vervolgens is het zaak om de metingen en andere procedures daar zo veel mogelijk op af te stemmen. Bij vragenlijstmetingen kan men denken aan toegankelijke vragenlijsten en bij klinische uitkomstmetingen kunnen definitie en het meetmoment van de uitkomst helpen om selectieve exclusie te beperken. Zo kan het verkorten van de follow-up periode (bijvoorbeeld van zes naar drie maanden) of het kiezen van uitkomsten die zich eerder voordoen, ertoe bijdragen dat minder patiënten wegvallen vóór de meting plaatsvindt. In sommige gevallen kan het zelfs zinvoller zijn om een andere uitkomstmaat te kiezen. Bijvoorbeeld als te voorzien is dat een groot deel van de doelgroep de beoogde meting niet zal halen.

c.6 Waar kan ik meer lezen over selectiebias

- Rothman, K. J., Greenland, S., & Lash, T. L. (2008). *Modern Epidemiology* (3rd ed.). Lippincott Williams & Wilkins.
- Sackett, D. L. (1979). Bias in analytic research. *Journal of Chronic Diseases*, 32(1-2), 51-63.
- Hernán, M. A., Hernández-Díaz, S., & Robins, J. M. (2004). A structural approach to selection bias. *Epidemiology*, 15(5), 615-625.
- Delgado-Rodriguez, M., & Llorca, J. (2004). Bias. *Journal of Epidemiology & Community Health*, 58(8), 635-641.
- Kelfve, S., Thorslund, M., & Lennartsson, C. (2013). Living conditions in old age: A comparison of survey and register data on the oldest old in Sweden. *European Journal of Ageing*, 10(2), 109-120.
- Groves, R. M. (2006). Nonresponse rates and nonresponse bias in household surveys *Public Opinion Quarterly*, 70(5), 646-675.

d. Relatie proces/structuur

d.1 Belang van relatie tussen proces/structuur variabelen - en uitkomstindicatoren

Uitkomstinformatie is alleen bruikbaar voor het verbeteren van zorg als uitkomsten beïnvloedbaar zijn. Dit geldt voor zorgverleners (proces) of de middelen waarover zij beschikken (structuur). Zonder deze relatie geeft een uitkomst geen handvatten voor

verbetering. Uitkomsten die uitsluitend afhangen van patiëntkenmerken geven dus geen informatie over de kwaliteit van de geleverde zorg.

d.2 Invloed nabijheid van het proces op de uitkomst

Hoe directer een proces leidt tot een uitkomst, hoe beter de relatie te meten is. Bij complexe zorgprocessen met meerdere stappen of bij uitkomsten die pas na lange tijd zichtbaar worden, is dit lastiger. Zo is het effect van een antibioticakuur op een infectie snel meetbaar. De impact van leefstijladvies op hart- en vaatziekten echter pas na jaren. Hoe meer tussenliggende factoren en tijdsvertraging, hoe moeilijker het is om het effect van een specifiek proces vast te stellen. Voor het vaststellen van de relatie tussen proces en uitkomst is het dus makkelijker om te kiezen voor ‘simpele’ processen, die waarschijnlijk een snel meetbaar resultaat opleveren. Vervolgens is het belangrijk om een zorgvuldige follow-up periode te kiezen.

d.3 Patiëntniveau-data nodig voor de relatie tussen proces en uitkomst

Voor het vaststellen van de relatie tussen proces en uitkomst is het belangrijk dat men gebruik maakt van data op patiëntniveau. De klinische studies waarop procesindicatoren zijn gebaseerd, zijn namelijk uitgevoerd op basis van individuele patiëntgegevens. Als men data alleen op ziekenhuis- of groepsniveau analyseert, kan dit leiden tot aggregatiebias. Dit betekent dat conclusies die zijn getrokken op basis van groepsdata niet altijd gelden voor individuele patiënten. Hierdoor kan men verkeerde verbanden leggen tussen zorgprocessen en uitkomsten.

d.4 Meer lezen over de relatie tussen proces en uitkomst?

- Finney JW, Humphreys K, Kivlahan DR, Harris AH. Excellent Patient Care Processes in Poor Hospitals? Why Hospital-Level and Patient-Level Care Quality-Outcome Relationships Can Differ. J Gen Intern Med. 2016 Apr;31 Suppl 1(Suppl 1):74-7. doi: 10.1007/s11606-015-3564-3. PMID: 26951280; PMCID: PMC4803671.
- Parast L, Doyle B, Damberg CL, Shetty K, Ganz DA, Wenger NS, Shekelle PG. Challenges in assessing the process-outcome link in practice. J Gen Intern Med. 2015 Mar;30(3):359-64. doi: 10.1007/s11606-014-3150-0. Epub 2015 Jan 7. PMID: 25564435; PMCID: PMC4351283.

4. Welke factoren beïnvloeden de betrouwbaarheid van uitkomstinformatie bij leren en verbeteren?

a. Meetfouten

a.1 Wat is een meetfout?

Een meetfout is een willekeurige afwijking tussen de gemeten waarde en de (vaak onbekende) werkelijke waarde voor een individuele observatie. Iedere meting heeft een zekere mate van meetfout, ongeacht de verschillen tussen zorgaanbieders te detecteren.

a.2 Hoe beïnvloeden meetfouten de interpretatie van uitkomstinformatie?

Meetfouten hebben op individueel niveau een vertekenende invloed en verhogen op groepsniveau de variabiliteit van de metingen. Dit vergroot de onzekerheid over de werkelijke waarde van individuele metingen en het groepsgemiddelde. De aanwezigheid van meetfouten maakt het moeilijker om bestaande verschillen tussen zorgaanbieders te detecteren.

a.3 Hoe ontstaan en voorkom je meetfouten bij klinische uitkomsten?

Handmatige invoer

Menselijke (invoer)fouten kunnen meetfouten veroorzaken. Probeer daarom het verzamelen en verwerken van gegevens zoveel mogelijk te automatiseren. Maak waar mogelijk gebruik van gegevens die al in het EPD worden geregistreerd en die men zoveel mogelijk automatisch vanuit medische apparatuur of andere systemen invoert. Vermijd het gebruik van secundaire systemen waarbij men gegevens handmatig moet overtypen. Beter is het om te zorgen voor een automatische koppeling tussen het EPD en andere systemen. Hierdoor kan men gegevens direct en zonder menselijke tussenkomst overnemen.

Enkelvoudige metingen

Omdat meetfouten willekeurig optreden, neemt hun invloed af bij het gebruik van herhaalde metingen. Voer, wanneer de uitkomst zich hiervoor leent en het praktisch realiseerbaar is, herhaalde metingen uit. Gebruik het gemiddelde van deze metingen als meetwaarde. Bijvoorbeeld bij het meten van bloeddruk, waarbij meerdere metingen op verschillende momenten een betrouwbaarder beeld geven dan een eenmalige meting.

Onbetrouwbare meetinstrumenten

De mate waarin men consistente metingen verkrijgt, varieert vaak tussen verschillende instrumenten voor dezelfde klinische uitkomst. Daarom is het belangrijk om, waar mogelijk, gebruik te maken van instrumenten waarvan de betrouwbaarheid voldoende is aangetoond in de betreffende toepassing en populatie. Maak bij beschikbaarheid van meerdere, uitwisselbare instrumenten gebruik van de meest betrouwbare instrumenten. Een voorbeeld hiervan is het meten van lichaamstemperatuur; een rectale meting is betrouwbaarder dan een meting in het oor. Deze wordt minder beïnvloed door omgevingsfactoren en geeft een stabielere kernlichaamstemperatuur weer.

Voor klinische meetinstrumenten die zijn gebaseerd op het oordeel van een zorgverlener is het van belang dat metingen door verschillende beoordelaars consistent zijn. Deze eigenschap, ook bekend als interbeoordelaarsbetrouwbaarheid, kan men voor continue maten kwantificeren met de Intraclass Correlatiecoëfficiënt (ICC). De ICC meet de mate van overeenstemming tussen metingen en varieert van 0 tot 1. Hogere waarden wijzen hierbij op een hogere betrouwbaarheid. Instrumenten met een $ICC > 0,75$ beoordeelt men vaak als voldoende betrouwbaar voor gebruik op groepsniveau, terwijl men een $ICC > 0,90$ als zeer goed beoordeelt. Voor categorische data, kan men de interbeoordelaarsbetrouwbaarheid evalueren met de Cohen's Kappa-coëfficiënt. Kappa-waarden variëren ook van 0 tot 1 en waarden boven 0,75 kan men als goed beschouwen en waarden boven 0,90 als zeer goed.

Ook voor instrumenten waarbij geen oordeel van meerdere beoordelaars nodig is, kan men de consistentie van opeenvolgende metingen onder vergelijkbare omstandigheden eveneens onderzoeken met de ICC of Kappa.

a.4 Hoe ontstaan en voorkom je meetfouten bij PROMs?

Onbetrouwbare meetinstrumenten

Net als bij klinische meetinstrumenten varieert de betrouwbaarheid van PROM-scores. De test-hertestbetrouwbaarheid kan men beoordelen door dezelfde vragenlijst op verschillende tijdstippen aan dezelfde groep patiënten voor te leggen en de consistentie van de scores te vergelijken. Daarnaast kan men de interne consistentie van PROMs, die bestaan uit meerdere vragen per schaal, bepalen met betrouwbaarheidscoëfficiënten. Een waarde van 0,7 of hoger beoordeelt men doorgaans als acceptabel voor betrouwbaarheidscoëfficiënten. Terwijl men waarden van 0,8 of hoger als goed waardeert en 0,9 of hoger als ideaal voor zeer hoge betrouwbaarheid.

Gebruik van PROMs die patiënten niet begrijpen

Het gebruik van PROMs die patiënten niet begrijpen is een belangrijke oorzaak van onbetrouwbare scores. Het is daarom belangrijk om te verifiëren dat patiënten de vragen in PROMs goed begrijpen. Dit kan men testen met een aangepaste versie van de

Pharos-sneltest. Deze beoordeelt snel en effectief of patiënten de vragen correct interpreteren. Door deze test bij de keuze voor PROMs te gebruiken, is de kans groter dat de verkregen PROM-scores daadwerkelijk de beoogde gezondheidsuitkomsten weergeven.

a.5 Waar kan ik meer lezen over meetfouten?

- Streiner, D. L., & Norman, G. R. (2008). Health Measurement Scales: A Practical Guide to Their Development and Use (4th ed.). Oxford University Press.
- De Vet, H. C. W., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). Measurement in Medicine: A Practical Guide. Cambridge University Press.

b. Steekproefonzekeerheid

b.1 Wat is ‘steekproefonzekeerheid’?

Steekproefonzekeerheid ontstaat wanneer je een eigenschap van een hele populatie probeert te schatten op basis van gegevens uit een steekproef van die populatie. Omdat de steekproef niet alle leden van de populatie omvat, zal de gemeten waarde van deze eigenschap een willekeurige afwijking vertonen ten opzichte van de werkelijke waarde in de hele populatie. Deze willekeurige afwijking noemt men de steekproefonzekeerheid.

b.2 Hoe beïnvloedt de steekproefgrootte de steekproefonzekeerheid?

De steekproefgrootte heeft een directe invloed op de steekproefvariabiliteit: hoe groter de steekproef, hoe kleiner de steekproefonzekeerheid. In grotere steekproeven is de variabiliteit tussen de waargenomen resultaten meestal lager, omdat ze een representatiever beeld geven van de gehele populatie. Daarentegen leiden kleinere steekproeven vaak tot grotere steekproefvariabiliteit. Dit omdat de uitkomsten meer onderhevig zijn aan toevallige fluctuaties. Daarom is het, bij het interpreteren van uitkomstinformatie, belangrijk om te weten hoe groot de steekproef was, aangezien dit de betrouwbaarheid van de conclusies beïnvloedt.

b.3 Wat zijn betrouwbaarheidsintervallen?

Betrouwbaarheidsintervallen (BI's) zijn schattingen die een bereik aangeven waarin een bepaalde waarde, zoals een gemiddelde of proportie, met een bepaalde zekerheid (meestal 95%) ligt. Ze geven de mate van onzekerheid of precisie van een schatting weer. Een

betrouwbaarheidsinterval wordt vaak gepresenteerd als een getal of percentage, gevolgd door een bereik tussen twee waarden, bijvoorbeeld “78% (95% BI: 72% - 84%)”.

b.4 Hoe gebruik je deze bij het rapporteren van uitkomstinformatie?

Een 95% betrouwbaarheidsinterval betekent dat bij het herhaald trekken van steekproeven, 95% van de zo geconstrueerde intervallen de werkelijke populatieparameter zou bevatten. Bij het rapporteren van uitkomstinformatie gebruikt men betrouwbaarheidsintervallen. Dit om aan te geven hoe nauwkeurig de schatting van een bepaalde uitkomst is. Maar ook om een idee te geven van de variabiliteit is die men kan verwachten, als men de meting in een andere steekproef zou herhalen. Een smal betrouwbaarheidsinterval duidt op een hoge precisie, terwijl een breed betrouwbaarheidsinterval aangeeft dat er meer onzekerheid is rond de geschatte waarde.

b.5 Hoe interpreteer je betrouwbaarheidsintervallen bij kwaliteitsverbetering?

Ligt een 95% betrouwbaarheidsinterval volledig boven een bepaalde referentiewaarde? Dan kunnen we met een hoge mate van zekerheid stellen dat de gemeten prestatie daadwerkelijk beter is dan die benchmark. Omgekeerd, als het betrouwbaarheidsinterval de benchmark overlapt, is het niet mogelijk om met zekerheid te bepalen of de geobserveerde verschillen betekenisvol zijn of slechts het resultaat van toevallige variatie. Door deze informatie mee te nemen in de interpretatie van resultaten, kunnen zorgverleners de effectiviteit van kwaliteitsverbeteringen beter inschatten.

b.6 Waar kan ik meer lezen over steekproefonzekerheid?

- Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L. M., & Woloshin, S. (2007). Helping Doctors and Patients Make Sense of Health Statistics. *Psychological Science in the Public Interest*, 8(2), 53-96.
- Altman, D. G., & Bland, J. M. (2005). Standard deviations and standard errors. *BMJ*, 331(7521), 903.
- Gardner, M. J., & Altman, D. G. (1986). Confidence intervals rather than P values: estimation rather than hypothesis testing. *British Medical Journal (Clinical research ed.)*, 292(6522), 746-750.

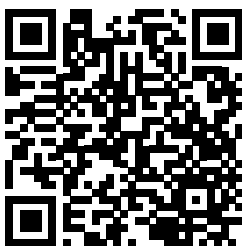
Linnean

Doendenkers voor de zorg

Betekenis 'doendenken'

doen·den·ken (werkwoord, onbepaalde wijs)

1 als passie voor de zorg wordt omgezet naar concrete daden, naar innovatieve prestaties die bepalend zijn, het goede voorbeeld geven en die iedereen en alles in beweging brengen.



.....

Sluit ook aan bij het Linnean Initiatief en **doendenk** mee! Al meer dan 1.650 mensen gingen je voor. Ga naar linnean.nl of scan de QR code



Leren en verbeteren op basis van uitkomstinformatie

Linnean