

Proportionaliteit en subsidiariteit NZa dataverzameling in de ggz: een wetenschappelijke analyse

Prof. dr. Jim van Os

Hoogleraar Psychiatrische Epidemiologie en Publieke ggz

Voorzitter Divisie Hersenen

UMC Utrecht

Lid Koninklijke Nederlandse Akademie van Wetenschappen (KNAW)

Fellow King's College London

WAT IS HET NZa ZORGPRESTATIEMODEL EN WAAR IS HET OP GEBASEERD?

In Nederland werd de bestaande diagnose-behandelcombinatie (dbc)-productstructuur en de daarmee samenhangende bekostigingssystematiek in 2022 vervangen door een nieuwe productstructuur volgens het 'Engelse model' of 'care cluster model'. De handleiding Zorgvraagtypering van de NZa, versie juli 2021 ¹, presenteert deze nieuwe aanpak voor de bekostiging van de curatieve geestelijke gezondheidszorg (ggz) als het *zorgprestatiemodel*. Een belangrijk onderdeel van dit model is de *zorgvraagtypering*, die vanaf 2024 zal worden gebruikt om de inzet van zorg bij verschillende groepen patiënten te helpen bepalen.

De totstandkoming van het zorgprestatiemodel is een initiatief van het Ministerie van Volksgezondheid, Welzijn en Sport (VWS). De uitvoering van dit plan werd in handen gegeven van de Nederlandse Zorgautoriteit (NZa). In oktober 2015 bevestigden alle relevante stakeholders binnen de geestelijke gezondheidszorg en forensische zorg hun betrokkenheid bij dit initiatief. Hoewel professionals deelnamen aan de besluitvorming, vond dit proces voornamelijk plaats op bestuurlijk en beleidsniveau, waardoor het minder werd opgemerkt op de werkvloer. De NZa zocht tijdens het proces geen contact met in de materie gespecialiseerde wetenschappers aan de UMC's ter toetsing van de wetenschappelijke fundamenteën – en daarmee de praktische haalbaarheid – van het zorgprestatiemodel ².

Het zorgprestatiemodel is gebaseerd op het Engelse 'care cluster' model – een model om patiënten in de ggz, die variabel en onvoorspelbaar zijn in het beloop van symptomen, zorgbehoeften en zorggebruik – zowel binnen een en dezelfde patiënt als tussen patiënten ³ – in te delen in standaardgroepen. In het Engelse model vormen de clusters een set van 20 groepen, die elk, volgens de hypothese, een vergelijkbare mate van zorgbehoeften en (toekomstige) zorgintensiteit zouden hebben. Het model zou daarmee voor het eerst in de geschiedenis van de ggz een gestandaardiseerd systeem introduceren voor het classificeren, behandelen en budgetteren van patiënten.

De standaardisatiehypothese onder het Engelse 'care cluster' model is ambitieus, omdat 50 jaar intensief wetenschappelijk onderzoek niet heeft kunnen aantonen dat onder de heterogene casemix van patiënten in de ggz 'standaardgroepen' schuilgaan, die met een eenvoudige korte vragenlijst – of een andere test – aangetoond zouden kunnen worden ⁴. Wetenschappelijk gezien bestaan er geen standaardgroepen in de ggz, omdat die nooit zijn aangetoond en, naar vermoed wordt, ook niet bestaan ⁵. En hoewel er veel variabelen zijn aan te wijzen die een statistisch significante relatie hebben met zorggebruik ⁶, vertaalt statistische significantie in de ggz zich typisch niet in een bruikbare voorspellende waarde – een bekend misverstand ^{2,7,8}. Deze wetenschappelijke kernbevinding lijkt in conflict te zijn met de aanname van de NZa dat patiënten in de ggz betrouwbaar en valide kunnen worden gereduceerd tot standaardgroepen met vergelijkbare zorgbehoeften en zorgintensiteit, en dat dit met een eenvoudige korte vragenlijst, in de ruis van de klinische praktijk van alledag, kan worden vastgesteld.

WERD HET ZORGPRESTATIEMODEL GEMOTIVEERD EN ONDERLEGD INGEVOERD ?

De Notitie beleid: Visie op doorontwikkeling dbc-systematiek ggz⁹ identificeert verschillende beperkingen van de huidige dbc-productstructuur. Het benoemt onder meer het gebrek aan medische

herkenbaarheid, onrealistische tijdsduren van behandelingen, te lange doorlooptijden, ongeschiktheid voor zorgsturing en hoge administratieve lasten. Daarnaast worden de dbc-declaraties als onbegrijpelijk voor patiënten beschouwd.

Met de invoering van de nieuwe productstructuur, zou er beter rekening worden gehouden met de zorgvraagwaarde van patiënten en zouden nieuw gedefinieerde behandelingspakketten worden aangeboden. De nieuwe productstructuur zou ook moeten beantwoorden aan de zorgen van verzekeraars over de zogenaamde risicoverevening. Volgens het plan van aanpak dat de NZa in 2016 formuleerde¹⁰, zou een vernieuwing geïnspireerd door het Engelse model deze tekortkomingen moeten aanpakken. Er werd dus geen rekening gehouden met de mogelijkheid om de bestaande dbc-productstructuur te verfijnen.

Er lijkt echter geen expliciete *positieve* (wetenschappelijke) argumentatie te zijn voor de invoering van de nieuwe productstructuur. In een nota van de inhoudelijke werkgroep wordt enkel verwezen naar niet nader gespecificeerde 'waardevolle elementen in het Engelse model'¹¹. Er wordt ook verwezen naar een vooronderzoek en een conferentie die in mei 2015 werd gehouden, maar er zijn geen verslagen van deze evenementen beschikbaar.

Relevant is dat hoewel het Engelse 'care cluster' model inderdaad in Engeland werd ontwikkeld om 'payment by result' mogelijk te maken¹², de NHS er vanaf 2014 achter kwam dat er problemen waren in het datasysteem en de algoritmische benadering van de clusterkostprijsberekening en de voorspelling van zorgactiviteit¹³. Voor het functioneren van het betalingssysteem zouden de clusters voldoende homogeen moeten zijn qua gebruik van financiële resources. Onderzoek echter toonde aan dat er grote variatie was tussen aanbieders met betrekking tot kosten binnen clusters, waardoor het care cluster model niet bruikbaar bleek als basis voor een nationaal betalings- en prijssysteem in de ggz binnen de NHS^{14 15,16}. Het 'Engelse Model' werd in Engeland dan ook nooit breed ingevoerd omdat het niet geschikt werd bevonden voor het doel dat de NZa er in Nederland nochtans mee hoopt te bereiken¹⁷.

HET INSTRUMENT ONDER DE ZORGVRAAGTYPERING: DE HoNOS

De Health of the Nation Outcome Scales (HoNOS) is een instrument dat is ontworpen om de gezondheidstoestand en het functioneren van mensen met ernstige psychische aandoeningen te meten. Het is *niet* ontworpen om 'standaardgroepen' te identificeren die voorspellend zijn voor zorgintensiteit en zorgkosten.

De HoNOS bestaat uit 12 items die elk een bepaald aspect van de geestelijke gezondheid of het sociale functioneren beoordelen, zoals gedragsproblemen, suïcidaliteit, agressie, beperkingen in het dagelijks functioneren, symptomen van psychische aandoeningen en problemen met relaties en sociale omstandigheden. De gegevens zijn zeer daarmee zeer privacygevoelig.

Hoewel de HoNOS vaak is gebruikt in verschillende landen als klinisch instrument (i.e. niet om standaardgroepen te identificeren of kosten te voorspellen), is er kritiek op dit instrument. Een aantal punten van kritiek zijn:

Betrouwbaarheid en validiteit: Veel studies hebben vraagtekens gezet bij de betrouwbaarheid en validiteit van de HoNOS. Dit betekent dat er zorgen zijn over de consistentie van de scores tussen verschillende beoordelaars (betrouwbaarheid) en de mate waarin de schaal daadwerkelijk meet wat het bedoelt te meten (validiteit) ¹⁸.

Brooks en collega's concludeerden: *The HoNOS has at best moderate interrater reliabilities. Further, the validity of the HoNOS is under question, that is, it does not correlate with a major measure of mental health symptoms, nor with a major measure of health status* ¹⁸.

Gigantesco en collega's concludeerden: *Although the HoNOS has many advantages in its brevity, it may lack sufficient discriminatory ability for certain patient groups. Further, it correlates only moderately with measures of disability and physical health status* ¹⁹.

Golay and colleagues concluded that *the suggested factor structures were not supported by the data. Predictive validity of the 12 items was weak..... HoNOS may be more adapted to the evaluation of patients case-mix rather than to the individual level and concepts such as predictive validity may be more appropriate than internal validity to guide its use* ²⁰.

Onduidelijke beoordelingscriteria: Er is kritiek geuit op de onduidelijkheid van sommige beoordelingscriteria in de HoNOS, wat kan leiden tot inconsistenties in de beoordelingen.

Speak en collega's schreven: *HoNOS should not be considered as 12 independent scales. Although the four-item scale is acceptable psychometrically, it has practical problems, as it focuses only on the social aspects of outcome measurement* ²¹.

De focus op symptomen en functioneren: Sommige critici stellen dat de HoNOS te sterk gericht is op symptomen en functioneren en niet voldoende rekening houdt met andere belangrijke aspecten van herstel, zoals empowerment en hoop.

Een intrigerend onderzoek, geleid door Keith G. Bender en gepubliceerd in het Australasian Psychiatry Journal in 2020, heeft al het onderzoek met de Health of the Nation Outcome Scale (HoNOS) als routine uitkomstmaat in de geestelijke gezondheidszorg onderzocht ²².

Bender's literatuuronderzoek omvatte een zoektocht naar de term 'HoNOS' in PubMed, Ovid MEDLINE en PsychINFO databases, resulterend in 260 onderzoeken die verder werden geanalyseerd. Uit deze analyse bleek dat geen enkele studie de kosten van het gebruik van HoNOS probeerde te beoordelen. Slechts één studie onderzocht het effect van routinematige uitkomstmeting met de HoNOS, en concludeerde dat het niet leidde tot het bieden van evidencebased zorg. Ondanks dat 97 van de onderzochte studies een activiteit in het voordeel van klinici beschreven, werd er geen enkele studie gevonden die de verstrekking van gegevens aan patiënten beschreef. Bovendien bleek dat geen enkele studie de kosten van het opzetten of gebruiken van HoNOS had onderzocht. Het onderzoek van Bender roept dus serieuze vragen op over de effectiviteit en de kosten van het gebruik van HoNOS in de geestelijke gezondheidszorg. Het belicht het gebrek aan bewijs voor de waarde van routinematige

uitkomstmeting en pleit voor meer onderzoek om de waarde ervan in de klinische praktijk vast te stellen.

Bender's onderzoek illustreert het feit dat er (i) weinig concreet bewijs is dat de brede toepassing van HoNOS als een routine uitkomstmaat gerechtvaardigd is. Verder is HoNOS onderzoek naar de vraag: wie in de ggz ontvangt welke diensten van wie, tegen welke kosten en met welk effect? – nauwelijks uitgevoerd, en daar waar het is gedaan suggereren de resultaten dat de HoNOS niet voorspellend is (zie hieronder).

De HoNOS data zijn zeer privacygevoelig. Dergelijke data kunnen alleen worden verzameld als wordt voldaan aan de beginselen van proportionaliteit en subsidiariteit. Het proportionaliteitsbeginsel houdt in dat de inbreuk op de privacy van individuen niet groter moet zijn dan strikt noodzakelijk om het beoogde doel te bereiken. Het subsidiariteitsbeginsel betekent dat er geen minder indringende maatregel beschikbaar moet zijn om hetzelfde doel te bereiken. Beide principes hebben tot doel de privacy van individuen te beschermen. Gezien de zwakke wetenschappelijke validiteit van de HoNOS kan hierover worden getwijfeld. Maar belangrijker nog voor deze toets is de vraag: kan de HoNOS inderdaad helpen bij het voorspellen van zorgintensiteit en zorgkosten, zoals gesteld door de NZa?

BEVESTIGT WETENSCHAPPELIJK ONDERZOEK DAT DE NZa ZORGVRAAGTYPERING BIJDRAAGT AAN BETERE VOORSPELLING VAN ZORGKOSTEN?

Het Engelse model werd door NZa gezien als een positieve toevoeging aan de risicoverevening in de gezondheidszorg. Risicoverevening is een systeem waarbij zorgverzekeraars een compensatie ontvangen op basis van het risicoprofiel van hun verzekerden. Zonder risicoverevening zouden verzekeraars geneigd zijn om gezonde verzekerden te selecteren en mensen met hoge gezondheidsrisico's te vermijden.

De verevening in de geestelijke gezondheidszorg (ggz) functioneert niet goed. Het budget voor ggz-kosten wordt door de minister van VWS bepaald en verdeeld over alle volwassen inwoners. Echter, de toewijzing aan subgroepen op basis van voorspellende kenmerken is niet nauwkeurig genoeg. De subgroepen bevatten zowel mensen met hoge kosten als mensen met lage of geen kosten, wat resulteert in tekorten voor degenen met hoge kosten en overschotten voor degenen met lage kosten. Onderzoek toont aan dat verzekeraars een tekort hebben op elke ggz-gebruiker, vooral op degenen met zeer hoge kosten. De verdeling van ggz-gebruikers en de kosten varieert sterk tussen verzekeraars, wat leidt tot ongelijke financiële lasten. Dit creëert een risico voor verzekeraars die een goede reputatie hebben op het gebied van ggz-zorginkoop, omdat ze aantrekkelijk zijn voor ggz-gebruikers waarop ze financieel tekortkomen. De aparte financiering van de ggz en de onvoorspelbaarheid van het ggz-gebruik zorgen voor een sterke prikkel tot risicoselectie, wat ongewenst is. Hoewel bepaalde voorspellers en gedwongen behandelingen de verevening enigszins kunnen verbeteren, blijft de onvoorspelbaarheid van het ggz-gebruik een belemmering. Compensaties achteraf bieden ook slechts beperkte hulp²³.

Zoals eerder beschreven heeft onderzoek laten zien dat het baseren van het bekostigingssysteem in de ggz op het Engelse care cluster model niet kan helpen, vooral vanwege de uitdagingen bij het homogeen

en precies onderscheiden van de kosten van patiënten binnen deze clusters^{13 14 15,16}. Bijvoorbeeld, in Engeland werd een poging gedaan om te zien of de classificatie in 20 clusters kon worden gerepliceerd in een database met bijna twee miljoen cluster episoden¹⁵. Deze hypothese bleek niet juist: de overlapping tussen de clusters van het model en de empirische clusters was beduidend klein. Bovendien was de homogeniteit van de kosten (zowel onder patiënten binnen dezelfde cluster als onder zorgverleners voor dezelfde cluster) te onstabiel om als basis voor een betalingssysteem te dienen.

Een andere studie, uitgevoerd in het VK in het boekjaar 2012-2013, betrof een historisch cohort van 1.343 patiënten met 'algemene geestelijke gezondheidsproblemen', geclassificeerd volgens ICD-10 codes tussen F32-48. De studie onderzocht de verbanden tussen individuele HoNOS-items en de totale HoNOS-score bij aanvang, en de kosten voor geestelijke gezondheidszorg na één jaar¹⁶. Wat de onderzoekers ontdekten, was verrassend: na correctie voor verschillende factoren bleken de meeste HoNOS-items – elf van de twaalf – niet geassocieerd met kosten. Alleen het item 'zelfverwonding' toonde een significante relatie met kosten, met een toegenomen kans op hogere kosten. Verder toonde de totale HoNOS-score geen verband met kosten, alhoewel deze wel verantwoordelijk leek voor een bescheiden deel van de incidentie van hogere kosten.

In een studie uit 1998 van Goldney, Fisher en Walmsley, gepubliceerd in de *Australian and New Zealand Journal of Psychiatry*, werd het nut van de Health of the Nation Outcome Scales (HoNOS) onderzocht bij het voorspellen van de verblijfsduur van patiënten in psychiatrische ziekenhuizen²⁴. De studie vond plaats in zes particuliere psychiatrische ziekenhuizen in vier Australische staten, waarbij de HoNOS zowel bij opname als ontslag werd afgenomen. Hoewel er significante verschillen in totaalscore en alle subschalen tussen opname en ontslag waren, bleek er geen significant verband te zijn tussen de totale HoNOS-scores en de verblijfsduur. Dit was het geval voor zowel individuele ziekenhuizen als voor specifieke diagnosegerelateerde groepen. Bovendien waren de gevonden associaties tussen de verblijfsduur van alle patiënten en individuele HoNOS-subschalen erg zwak, wat hun praktische klinische relevantie in twijfel trekt. Uit deze bevindingen komt naar voren dat, hoewel de HoNOS enige waarde heeft als een gemakkelijk toepasbaar en begrijpelijk meetinstrument voor klinici, het onwaarschijnlijk is dat het nuttig is bij het voorspellen van de verblijfsduur of bij het aanbieden van een 'gate-keeping'-dienst bij de besluitvorming over de toewijzing van middelen voor individuele patiënten. Daarmee ondersteunt deze studie de bredere discussie over de beperkte voorspellende waarde van de HoNOS met betrekking tot de intensiteit van ggz-zorggebruik.

Al deze studies tonen aan dat het care cluster model gebaseerd op de HoNOS niet valide is. Hoewel dit een ernstig probleem vormt, is het niet geheel onverwacht. De clusters zijn hybride structuren waarin symptomen en vereiste zorgniveaus zijn samengevoegd. Dit kan worden beschouwd als een fundamentele vergissing in het design. De clusters bestaan uit groepen klachten of problemen en niet, zoals het zou moeten, uit zorgproducten met een bepaalde kostenintensiteit. Dit is dan ook de reden dat het *Royal College of Psychiatrists* het care cluster model niet ondersteunt en het *Payment by Results* systeem, gebaseerd op de care clusters, nooit werd ingevoerd in Engeland¹⁷.

De ggz zal dus in problemen komen als het care cluster model door de NZa gaat worden gebruikt. Stel bijvoorbeeld voor dat twee patiënten dezelfde typering in eenzelfde care cluster, maar verschillende behandelbehoeften en complexiteit van hun aandoeningen. Patiënt A heeft misschien een milde vorm

van de zorgvraagtypering en reageert goed op kortdurende behandeling, terwijl patiënt B een ernstigere vorm heeft en intensieve langdurige behandeling vereist. Ondanks dat ze beide binnen dezelfde zorgvraagtypering VALLEN, zullen de kosten voor de behandeling van patiënt B aanzienlijk hoger zijn dan die van patiënt A. Dit illustreert de variabiliteit van de kosten binnen hetzelfde care cluster. Ook zijn er altijd regionale verschillen in de kosten per care cluster. Verschillende regio's kunnen unieke uitdagingen hebben, zoals verschillen in demografie, beschikbaarheid van zorgfaciliteiten en sociaaleconomische factoren. Nederlands onderzoek heeft aangetoond dat deze factoren een grote invloed hebben op de behandelkosten binnen een bepaald cluster in verschillende regio's ²⁵⁻²⁷.

Onderzoek van de NZa zelf bevestigt dat het Engelse care cluster niet geschikt is voor zorgvraagtypering en zorginkoop. In hun analyses op basis van data uit heel Nederland bleek dat zorgvraagtypering met de HoNOS slechts 24% van de variantie in de kosten kan verklaren. Dit percentage is veel te laag om op een transparante en rechtvaardige manier de kosten en zorginloop vast te stellen. Een bekostigingssysteem dat voornamelijk gebaseerd is op dit model zou leiden tot onnauwkeurige en oneerlijke vergoedingen en risicoverevening, omdat het geen rekening houdt met de daadwerkelijke variatie in kosten en behoeften binnen de clusters.

Om een meer rechtvaardig en nauwkeurig bekostigingssysteem te creëren, is het van belang om rekening te houden met de complexiteit en individuele behoeften van patiënten, evenals met regionale verschillen in kosten. Een holistische benadering die rekening houdt met individuele kenmerken, behoeften en behandelresultaten is noodzakelijk om een eerlijke en effectieve bekostiging in de ggz te waarborgen ²⁸.

Dat ook de NZa met de zorgvraagtypering niet verder komt dan 24% voorspelde variantie van de kosten is in overeenkomst met de staat van de wetenschap in de ggz. Over een periode van 50 jaar onderzoek is bekend geworden dat het beloop van psychisch aandoeningen slechts in beperkte mate (ongeveer 25%) voorspelbaar is op basis van psychopathologische, functionele, cognitieve of biologische data. Dit is typisch voor complexe systemen waar onze mentale variatie onder valt. Meer data gaan de analyses van de NZa niet preciezer maken – de beperking zit immers niet in de hoeveelheid data, maar in het stochastische element dat complexe systemen eigen is ³.

CONCLUSIE

De principes van proportionaliteit en subsidiariteit zijn belangrijk in gegevensbescherming en privacywetgeving, zoals de Algemene Verordening Gegevensbescherming (AVG) in de EU.

Het proportionaliteitsbeginsel houdt in dat de inbreuk op de privacy van individuen niet groter moet zijn dan strikt noodzakelijk om het beoogde doel te bereiken. Het subsidiariteitsbeginsel betekent dat er geen minder indringende maatregel beschikbaar moet zijn om hetzelfde doel te bereiken. Beide principes hebben tot doel de privacy van individuen te beschermen.

De Nederlandse Zorgautoriteit (NZa) kan daarom de gevraagde gegevens van GGZ-patiënten op basis van de Health of the Nation Outcome Scales (HoNOS) niet vragen als het beoogde doel (het voorspellen van zorgintensiteit en zorgkosten) wetenschappelijk niet kan worden gerealiseerd. Het zou een

disproportionele inbreuk op de privacy van de patiënten zijn, vooral gezien de gevoelige aard van de HoNOS-data.

Als er ook geen aanwijzingen zijn dat deze specifieke gegevens essentieel zijn voor het behalen van het doel, dan kan het verzamelen ervan ook in strijd zijn met het subsidiariteitsbeginsel. Het is altijd nodig om naar alternatieve, minder invasieve methoden te zoeken om hetzelfde doel te bereiken.

Daarnaast vereist de AVG dat de verwerking van persoonsgegevens noodzakelijk en relevant is voor het beoogde doel. Als de HoNOS-gegevens het beoogde doel niet kunnen bereiken, dan zou het verzamelen ervan ook in strijd kunnen zijn met deze eisen.

Kortom, het is van groot belang dat de privacy van individuen wordt gerespecteerd en dat de verzameling en verwerking van persoonlijke gegevens zorgvuldig worden afgewogen tegen de potentiële voordelen. In het geval van de gevraagde data voor de zorgvraagtypering staan, wetenschappelijk gezien, de voordelen in geen verhouding tot de nadelen.

REFERENTIES

1. Handleiding Zorgvraagtypering Versie juli 2021. Downloaded at chrome-extension: <https://www.zorgprestatie.nl/content/uploads/2021/07/Handleiding-Zorgvraagtypering-juli-2021.pdf>.
2. van Os J. Zorgvraagzwaartemodell 1.0: naar een model van random zorgtoewijzing? *Tijdschrift voor Psychiatrie*. 2014;56.
3. Scheepers F. *Mensen zijn ingewikkeld. Een pleidooi voor acceptatie van de werkelijkheid en het loslaten van modellen*. Amsterdam: Arbeiderspers; 2020.
4. Kapur S, Phillips AG, Insel TR. Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Mol Psychiatry*. 2012;17(12):1174-1179.
5. Kendler KS, Zachar P, Craver C. What kinds of things are psychiatric disorders? *Psychological Medicine*. 2010;1-8.
6. Twomey CD, Baldwin DS, Hopfe M, Cieza A. A systematic review of the predictors of health service utilisation by adults with mental disorders in the UK. *BMJ Open*. 2015;5(7):e007575.
7. Bremer-Hoeve S, de Groot E, van Wel EB, Rutten BPF, van Bronswijk SC. [Zorgvraagtypering met de HoNOS+: kijken in een glazen bol?]. *Tijdschr Psychiatr*. 2022;64(7):416-4147.
8. van Os J, Reininghaus U. The Clinical Epidemiology of Schizophrenia. In: Kaplan B, Saddock J, Saddock V, Ruiz P, eds. *Kaplan & Saddock's Comprehensive Textbook of Psychiatry, 11th Edition*. London: Lippincott Williams & Wilkins 2021.
9. Notitie beleid: Visie op doorontwikkeling DBC-systematiek GGZ. 2015. <https://werkenmetdbcs.nza.nl/documenten-ggz/visiedocumenten/7977-bijlage-3-2-notitie-beleid-visie-doorontwikkeling-productstructuur-ggz/file?lang=nl>.
10. Nederlandse Zorgautoriteit. Plan van aanpak. Doorontwikkeling productstructuur GGZ / FZ. De stappen naar een nieuw model (2016-2019). Utrecht: NZa; 2016.
11. Werkgroep Inhoud. Verkenning doorontwikkeling Productstructuur DB(B)C GGZ/FZ. Het Engelse bekostigingssysteem uitgedaagd. (Concept) Advies werkgroep Inhoud. 2015.
12. Mason, A., & Goddard, M. (2009) Payment by results in mental health: A review of the international literature and an economic assessment of the approach in the English NHS. CHE Research Paper 50. Centre for Health Economics, University of York. Available from: http://www.york.ac.uk/media/che/documents/papers/researchpapers/rp50_pay... [last accessed 31 August 2012].
13. CAPITA. The quality of Mental Health care cluster costing and activity data in the NHS Payment by Results data assurance framework September 2014 (downloaded at chrome-extension:https://www.chks.co.uk/userfiles/files/The_quality_of_Mental_Health_care_cluster_costing_and_activity_data_in_the_NHS.pdf).
14. Jacobs R, Chalkley M, Bohnke JR, Clark M, Moran V, Aragon MJ. Measuring the Activity of Mental Health Services in England: Variation in Categorising Activity for Payment Purposes. *Adm Policy Ment Health*. 2019;46(6):847-857.
15. Jacobs R, Chalkley M, Aragón MJ, Böhnke JR, Clark M, Moran V, e.a. Funding of mental health services: do available data support episodic payment ? CHE Research Paper 137. York: CHE; 2016.
16. Twomey C, Prina AM, Baldwin DS, et al. Utility of the Health of the Nation Outcome Scales (HoNOS) in Predicting Mental Health Service Costs for Patients with Common Mental Health Problems: Historical Cohort Study. *PLoS One*. 2016;11(11):e0167103.
17. Broekman TG, Schippers GM. [Why Dutch mental health care practitioners should have second thoughts about adopting the English cluster model]. *Tijdschr Psychiatr*. 2017;59(11):702-709.

18. Brooks R. The reliability and validity of the Health of the Nation Outcome Scales: validation in relation to patient derived measures. *Aust N Z J Psychiatry*. 2000;34(3):504-511; discussion 520-501.
19. Gigantesco A, Picardi A, de Girolamo G, Morosini P. Discriminant ability and criterion validity of the HoNOS in Italian psychiatric residential facilities. *Psychopathology*. 2007;40(2):111-115.
20. Golay P, Basterrechea L, Conus P, Bonsack C. Internal and Predictive Validity of the French Health of the Nation Outcome Scales: Need for Future Directions. *PLoS One*. 2016;11(8):e0160360.
21. Speak B, Muncer S. The structure and reliability of the Health of the Nation Outcome Scales. *Australas Psychiatry*. 2015;23(1):66-68.
22. Bender KG. The meagre outcomes of HoNOS. *Australas Psychiatry*. 2020;28(2):206-209.
23. Beekman ATF, van der Lee A. [Het spook risicoselectie; ggz onverzekerbaar in huidig stelsel]. *Tijdschr Psychiatr*. 2022;64(7):411-415.
24. Goldney RD, Fisher LJ, Walmsley SH. The Health of the Nation Outcome Scales in psychiatric hospitalisation: a multicentre study examining outcome and prediction of length of stay. *Aust N Z J Psychiatry*. 1998;32(2):199-205.
25. Van Os J. Variatie in ggz-kosten op regioniveau in relatie tot indicatoren van populatiezorgbehoefte: op weg naar evidence-based zorginkoop? *Tijdschrift voor Psychiatrie*. 2019;61:617-625.
26. van Os J, Mulder W. Analysis of mental health care utilization data in comparison with medical-specialist and general practitioner health care consumption data. *Tijdschrift voor Psychiatrie*. 2021;63(1):39-47.
27. Van Os J. To what degree is the strong social gradient of mental health care accompanied by a uniformly consistent association between mental health care and social care? *Tijdschrift voor Psychiatrie*. 2022;64:366-376.
28. Van Os J. *Van Routine Outcome naar Routine Omgeving Monitoring: Een nieuwe methode om verzekeraars en ggz te helpen met focus en doelmatigheid. Rapport aangeboden aan Staatssecretaris Paul Blokhuis op 16-7-2020*. Utrecht: UMC Utrecht;2020.