

# HoNOS-normgegevens voor patiënten met EPA in acute opname, langdurige klinische zorg en FACT-zorg

E. de Beurs, J. Peen, U. Nabitz, M.B. de Koning

- Achtergrond** De *Health of the Nation Outcome Scale* (HoNOS) is ruim 20 jaar in gebruik in Nederland. Normen voor de HoNOS zijn echter sinds 2004 niet meer geactualiseerd.
- Doel** Het opstellen van grootstedelijke HoNOS-normscores voor patiënten met een ernstige psychiatrische aandoening (EPA) in drie behandelsettings.
- Methode** Geanonimiseerde gegevens van een omvangrijk cohort (n = 7566) uit de jaren 2012–2021 van patiënten met EPA in acute klinische zorg (n = 4309), in langdurige klinische zorg (n = 822) en zorg bij *flexibele assertive community treatment* (FACT; n = 2435) werden gebruikt om normen op te stellen voor de (totale) indexscore, het aantal klinisch betekenisvolle en het aantal ernstige probleemgebieden.
- Resultaten** Er waren significante verschillen in HoNOS-scores tussen de drie zorgvormen: patiënten in acute klinische zorg scoorden hoger dan patiënten in langdurige klinische zorg, en zij weer hoger dan degenen in FACT-zorg. Patiënten van 18 tot 40 jaar scoorden lager dan degenen boven de 40. Juridische status bleek niet betekenisvol van invloed op de HoNOS-score. Normtabellen met ruwe scores, genormaliseerde T-scores en percentielscores werden opgesteld voor geslacht en leeftijd.
- Conclusie** De nieuwe normgegevens stellen klinici in staat om bij individuele patiënten de ernst van de psychische aandoening en het niveau van functioneren in te schatten in vergelijking met passende referentiegroepen.

De *Health of the Nation Outcome Scale* (HoNOS) is een beoordelingsschaal voor de psychische gezondheid en het functioneren van mensen met psychische klachten. De HoNOS werd eind vorige eeuw in Engeland ontwikkeld om de gezondheidstoestand en het sociaal functioneren van psychiatrische patiënten op een gestructureerde manier vast te stellen.<sup>1</sup> De HoNOS wordt vooral gebruikt bij patiënten met een ernstige psychiatrische aandoening (EPA). Hieronder schaaft men patiënten met een langdurige psychiatrische stoornis, met ernstige beperkingen in het functioneren, die gevolg en oorzaak zijn van de stoornis. Om het behandelplan te realiseren is gecoördineerde zorg geïndiceerd van professionele hulpverleners in zorgnetwerken.<sup>2</sup> Schattingen van de prevalentie van EPA lopen uiteen: de lifetimeprevalentie varieert van 1 tot 3% van de bevolking en dat zou neerkomen op 200.000 tot 500.000 patiënten in Nederland.<sup>3</sup> De HoNOS wordt veel toegepast en het gebruik kreeg recentelijk een nieuwe impuls met de ontwikkeling van de HoNOS+ ten behoeve van het zorgprestatiemodel.<sup>4</sup> In Nederland is echter relatief weinig onderzoek naar

het instrument gedaan. De HoNOS is vertaald naar het Nederlands met voorlopige normgegevens.<sup>5</sup> Er is positief gerapporteerd over gebruik van de HoNOS als instrument voor *routine outcome monitoring* (ROM) in de ggz.<sup>6,7</sup> Er is verslag gedaan van behandeluitkomsten van ACT-teams in Rotterdam met behulp van de HoNOS.<sup>8,9</sup> Ook is de toepassing van de HoNOS in de ouderenpsychiatrie en bij patiënten met een lichte verstandelijke beperking onderzocht.<sup>10,11</sup> Bij de acute opnamekliniek van Arkin in Amsterdam is onderzoek gedaan naar het behandel-effect met de HoNOS.<sup>12</sup> Er zijn uitkomstcriteria ontwikkeld op basis van Italiaanse en Nederlandse HoNOS-gegevens.<sup>13</sup> Verschillende manieren om therapie-uitkomsten te operationaliseren bij patiënten met EPA zijn vergeleken op basis van de HoNOS.<sup>14</sup> De HoNOS-scores zijn geschikt gemaakt voor economische analyses.<sup>15</sup> Recent is verslag gedaan van de toepassing van de HoNOS in de ambulante forensische ggz (J.M. Jansen, U.W. Nabitz, T. van der Pol, e.a. schriftelijke mededeling december 2025).<sup>16</sup> Voor deze doelgroep bleek de bruikbaarheid van de HoNOS beperkt.

Wat op het moment vooral ontbreekt, zijn betere Nederlandse normgegevens om HoNOS-scores te interpreteren. Er zijn Engelse, Australische, Duitse, Zwitserse en Noorse normen, maar voor Nederland zijn slechts beperkte normgegevens voorhanden.<sup>17-23</sup> De enige ons bekende publicatie met Nederlandse normgegevens betrof een groep patiënten van beperkte omvang (n = 559) die in behandeling waren bij diverse Riaggevestigingen in de regio Rotterdam.<sup>5</sup>

In dit artikel leveren wij normgegevens voor drie relevante en welomschreven groepen patiënten met EPA die in een grootstedelijke context in zorg komen. Daarbij bieden we normgegevens voor specifieke subgroepen patiënten. Hiermee kunnen klinici bij individuele patiënten de ernst van de psychiatrische aandoening en het niveau van functioneren inschatten in vergelijking met een passende referentiegroep.

## METHODE

### Opzet

We verrichtten een studie op basis van geanonimiseerde ROM-data verzameld in het kader van de reguliere zorg. De rapportage en het gebruik van termen in dit artikel zijn in overeenstemming met de COSMIN-richtlijnen voor psychometrisch onderzoek.<sup>24</sup>

### Context

De data werden verzameld bij Mentrum, onderdeel van Arkin, een grote ggz-instelling in Amsterdam en omstreken. Mentrum is het onderdeel van Arkin waar patiënten met EPA's worden behandeld in drie zorgvormen: acute klinische zorg (AKZ), langdurige klinische zorg (LKZ) en FACT-zorg. In de twee acute klinieken van Mentrum varieert de beoogde opnameduur van dagen tot enkele maanden. LKZ wordt geboden op vier locaties, die tegenwoordig samen de 'herstelondersteunende vervolgklinieken' worden genoemd. Opnamecriteria en -duur verschillen tussen de locaties, maar overkoepelend gaat het bij LKZ om patiënten die op de acute kliniek binnen 3-6 maanden onvoldoende herstellen om met ontslag naar huis of beschermd wonen te kunnen en daarnaast patiënten die worden terugverwezen door gespecialiseerde klinieken voor klinische forensische of hoogbeveiligde zorg. De opnameduur varieert van 1 tot > 5 jaar. De ambulante teams van Mentrum werken volgens de FACT-methodiek en zijn FACT-gecertificeerd.<sup>25</sup>

### Patiënten

Patiënten werden geïnccludeerd voor wie van 1 januari 2012 tot en met 31 december 2021 een eerste HoNOS (aanvangsmeting) was ingevuld.

### Dataverzameling en medisch-ethische aspecten

Van patiënten werden de volgende data geëxtraheerd uit de ROM-gegevensset: de HoNOS-scores, sekse, leeftijd,

juridische status (wel/geen Bopz/Wvggz-maatregel) en hoofddiagnose (volgens DSM-IV- of DSM-5-classificatie). In dit onderzoek gebruikten we de HoNOS van de eerste beoordeling, bij de start van de behandeling bij het zorgtype. Een HoNOS-beoordeling werd als aanvangsmeting binnen het betreffende zorgtype beschouwd als deze binnen een week voor de start van de behandeling en bij de LKZ niet langer dan 6 maanden na de start van de behandeling had plaatsgevonden.

Het komt voor dat een patiënt na een eerdere behandeling weer opnieuw in behandeling komt bij een zorgtype. Het komt ook voor dat patiënten intern binnen Mentrum doorverwezen worden, bijvoorbeeld van acute opname naar langdurige klinische zorg en vervolgens naar FACT-zorg. Als dit het geval was, dan werd bij heropname of bij de overgang naar een andere zorgsetting opnieuw een HoNOS afgenomen. Dezelfde patiënten konden dus voorkomen in de diverse steekproeven, maar zij werden beoordeeld op telkens een ander moment in hun 'patient journey'.

We analyseerden 7566 eerste HoNOS-beoordelingen: voor AKZ: n = 4309; voor LKZ: n = 822; en voor in zorg bij FACT: n = 2435. De metingen waren afkomstig van 4806 unieke patiënten: respectievelijk n = 2804, n = 282 en n = 1720.

Een HoNOS-beoordeling werd door zorgprofessionals gedaan op basis van het intakegesprek, observaties en dossiergegevens. Clinici werden uitgebreid getraind in afname en scoring van de HoNOS.

De benodigde aanvullende data (sekse, leeftijd, juridische status, zorgtype en hoofddiagnose) werden door datamanagers met een query geëxtraheerd uit de elektronische patiëntendossiers. Hiervoor was geen dossierinzage nodig. De gegevens werden volledig geanonimiseerd. Er bestaat dus geen sleutelbestand en het is niet meer mogelijk om te achterhalen welke gegevens van welke patiënt waren. De verzamelde data zijn dermate grofmazig dat het redelijkerwijs ook niet mogelijk is om een patiënt op basis van deze data te herkennen. Het gaat dus om anonieme data, en niet om gecodeerde data. Het gebruik van anonieme data voor wetenschappelijk onderzoek is toegestaan, aangezien de Algemene verordening gegevensbescherming (AVG) en de Wet op de geneeskundige behandelingsovereenkomst (WGBO) betrekking hebben op handelingen met persoonsgegevens.<sup>26</sup>

### Meetinstrument

De door ons onderzochte HoNOS is de versie die tot 1 januari 2022 overal in Nederland werd gebruikt. Deze versie bestaat uit 12 meerkeuze-items en 3 aanvullende items die bij vertaling van de HoNOS werden toegevoegd.<sup>5</sup> De indexscore op de HoNOS wordt berekend door de score op de eerste 12 items op te tellen en heeft een bereik van 0-48. De 12 oorspronkelijke HoNOS-items zijn: 1. overactief en agressief gedrag, 2. opzettelijke zelfverwonding, 3. alcohol, drugs of medicatie, 4. cognitieve problemen, 5. lichamelijke problemen, 6. halluci-

naties en waanvoorstellingen, 7. depressieve stemming, 8. overige psychiatrische symptomen, 9. sociale contacten, 10. ADL-activiteiten, 11. woonomstandigheden en 12. beroep/dagbesteding. De drie aanvullende items zijn: A1. maniforme ontremming, A2. gebrek aan motivatie voor behandeling en A3. therapietrouw voor medicatie. Met item 8 is er nog een bijzonderheid. De beoordelaar wordt bij dit item gevraagd het ernstigste klinische probleem te scoren dat niet valt onder de voorgaande items. De beoordelaar kan hierbij kiezen uit 10 categorieën: A. fobie, B. angst, C. dwangmatigheid, D. gespannenheid, E. dissociatie, F. somatisatie, G. eetproblemen, H. slaap, I. seksuele problemen, J. andere problemen. Van de gekozen symptoomgroep wordt vervolgens weer de ernst beoordeeld.

Bij de HoNOS geldt als observatieperiode de afgelopen 2 weken. Items hebben 5 responsies: 0 = geen probleem, 1 = ondergeschikt probleem dat geen therapeutische actie vereist, 2 = licht probleem, 3 = matig ernstig probleem, 4 = ernstig tot zeer ernstig probleem. Naast deze omschrijving hebben de responsies voor elk item een unieke toelichting, waarin de diverse niveaus uitgebreider omschreven worden. De score op HoNOS-items kan gedichotomiseerd worden in 0-1 (afwezig of ondergeschikt probleem) vs. 2-4 (aanwezig en klinisch relevant) of in 0-2 (afwezig tot licht probleem) vs. 3-4 (matig ernstig tot zeer ernstig). Anders gezegd: een item met een score van 2 of meer wordt als klinisch betekenisvol beschouwd; een item met een score van 3 of meer als ernstig. Een HoNOS-afname resulteert zo in 12 + 3 itemscores en drie samenvattende scores: de indexscore op de eerste 12 items (somscore; bereik 0-48) en twee keer een optelling van gedichotomiseerde itemscores (bereik 0-12).

### Statistische analyse

Met variantieanalyse (ANOVA) en een correlatieve analyse onderzochten we of HoNOS-indexscores over de tijd stabiel waren. De indexscore vergeleken we ook voor de drie zorgtypen (AKZ, LKZ en FACT-zorg) onderling met ANOVA. We geven normgegevens voor de drie groepen patiënten met EPA in de vorm van groepsgemiddelden (en standaarddeviaties), genormaliseerde T-scores en percentielrangordscores (PR-scores). Geïnteresseerde lezers vinden uitgebreide uitleg over T- en PR-scores in het onlinesupplement bij dit artikel.

We normaliseerden HoNOS-data met de RankNorm-functie beschikbaar in het RNOmni package, versie 1.2.0 van R.<sup>27</sup> Psychometrische indicatoren verkregen we met het psych package, versie 2.2.9 van R. Met t-toetsen voor onafhankelijke steekproeven of met ANOVA (bij meer dan twee groepen) gingen we na of er verschillen waren in indexscore op de HoNOS tussen de seksen, leeftijdsgroepen, groepen met verschillende juridische status, zorgtype, en zes diagnostische groepen (zie **tabel 1**). In geval van verschil in indexscore tussen subgroepen is het mogelijk om normgegevens voor subgroepen te gebruiken. Dit staat een preciezere interpretatie toe,

zoals ‘vergeleken met andere oudere cliënten die in FACT-zorg komen, heeft deze cliënt een hoge indexscore’. Gedetailleerde normgegevens (per zorgtype en per leeftijdsgroep) hebben we opgenomen in het online-supplement.

## RESULTATEN

### Normatieve cohorten

We analyseerden HoNOS-data bij drie groepen patiënten: patiënten in AKZ, LKZ en ambulante FACT-zorg.

**Tabel 1** laat de omvang van de cohorten en enkele achtergrondkenmerken van de groepen zien.

Allereerst onderzochten we of de HoNOS-scores stabiel waren over de tijd. De gemiddelde HoNOS-indexscore per jaar in de drie klinieken is weergegeven in **online-tabel S-3**. Volgens een ANOVA was er een significante afname van HoNOS-score over de jaren bij de AKZ ( $F(9,4297) = 9,63$ ;  $p < 0,001$ ;  $r = -0,10$ ) en bij de FACT ( $F(9,2422) = 6,20$ ;  $p < 0,001$ ;  $r = -0,09$ ), maar niet bij de LKZ ( $F(9,807) = 1,23$ ;  $p < 0,272$ ;  $r = -0,06$ ). Pearsons  $r$  geeft de correlatie weer tussen jaar van beoordeling en HoNOS-indexscore. De afname was echter gering (2,6 punten over 10 jaar voor de gehele groep;  $d = 0,36$ , een klein effect).

Vervolgens bepaalden we de indexscore, het aantal klinische en het aantal ernstige probleemgebieden voor alle 7566 beoordelingen. De gemiddelde indexscore was 14,28 (SD: 7,32), het aantal als klinisch relevant beoordeelde items was  $M: 4,60$  (SD: 2,64) en het aantal ernstige probleemgebieden was  $M: 2,34$  (SD: 2,14).

In **online-tabel S-4** staan de gemiddelde itemscore en de indexscore op de HoNOS voor elke groep, evenals het resultaat van toetsing van verschillen met ANOVA.

Er was overlap van patiënten in de drie groepen, waardoor de gegevens van de drie groepen niet geheel onafhankelijk waren, hetgeen van invloed was op de toetsingsresultaten van de vergelijking van zorggroepen. We herhaalden daarom de analyses voor de 4806 unieke patiënten, door per patiënt alleen de eerste HoNOS-beoordeling, ongeacht het zorgtype, te selecteren. Het algemene patroon van resultaten en de significante verschillen tussen de drie klinische groepen bleven hierbij in stand.

De drie groepen verschilden op alle items, met uitzondering van item 5 (lichamelijke problemen). De groepen verschilden ook in gemiddelde indexscore ( $F(2,7553) = 567,98$ ;  $p < 0,001$ ;  $\eta^2 = 0,131$ ): de AKZ-groep scoorde gemiddeld 4,1 punten hoger dan de LKZ-groep, een gemiddeld verschil (Cohens  $d = 0,60$ ; 95%-BI: 0,53-0,68) en 5,6 punten hoger dan de FACT-groep, een aanzienlijk verschil (Cohens  $d = 0,83$ ; 95%-BI: 0,78-0,88). De LKZ-groep scoorde gemiddeld 1,5 punt hoger dan de FACT-groep, een gering verschil (Cohens  $d = 0,22$ ; 95%-BI: 0,14-0,30). Ook geven we in **online-tabel S-4** het gemiddelde aantal items weer dat als ‘klinisch relevant’ en als ‘matig tot ernstig probleem’ werd beoordeeld. De AKZ-groep was er het slechtst aan toe met gemiddeld ruim 5 aanwezige

**Tabel 1. Steekproefbeschrijving**

|                               | AKZ-groep<br>(n = 4309) |             | LKZ-groep<br>(n = 822) |             | FACT-groep<br>(n = 2435) |             |
|-------------------------------|-------------------------|-------------|------------------------|-------------|--------------------------|-------------|
|                               | N                       | (%)         | N                      | (%)         | N                        | (%)         |
| Geslacht: mannen              | 2603                    | (60,3)      | 507                    | (62,1)      | 1443                     | (59,3)      |
| <b>Juridische status</b>      |                         |             |                        |             |                          |             |
| Vrijwillig                    | 3379                    | (78,4)      | 578                    | (70,3)      | 2120                     | (87,1)      |
| Maatregel                     | 930                     | (21,6)      | 244                    | (29,7)      | 315                      | (12,9)      |
| <b>Primaire diagnoses</b>     |                         |             |                        |             |                          |             |
| Psychotische stoornis         | 2349                    | (54,5)      | 590                    | (71,8)      | 1524                     | (62,5)      |
| Bipolaire stoornis            | 606                     | (14,1)      | 69                     | (8,4)       | 210                      | (8,6)       |
| Middelengebruik               | 566                     | (13,1)      | 32                     | (3,9)       | 77                       | (3,2)       |
| Persoonlijkheidsstoornis      | 242                     | (5,6)       | 50                     | (6,1)       | 211                      | (8,7)       |
| Depressieve stoornis          | 250                     | (5,8)       | 32                     | (3,9)       | 110                      | (4,5)       |
| Andere stoornis               | 294                     | (6,8)       | 49                     | (6,0)       | 303                      | (12,4)      |
| <b>Leeftijdsgroepen, in j</b> |                         |             |                        |             |                          |             |
| 17-30                         | 1033                    | (24,0)      | 171                    | (20,8)      | 412                      | (16,9)      |
| 31-40                         | 1016                    | (23,6)      | 211                    | (25,7)      | 603                      | (24,8)      |
| 41-50                         | 1066                    | (24,7)      | 189                    | (23,0)      | 663                      | (27,2)      |
| 51-60                         | 825                     | (19,1)      | 177                    | (21,5)      | 542                      | (22,3)      |
| 61+                           | 372                     | (8,6)       | 74                     | (9,0)       | 215                      | (8,8)       |
|                               | <b>M</b>                | <b>(SD)</b> | <b>M</b>               | <b>(SD)</b> | <b>M</b>                 | <b>(SD)</b> |
| Leeftijd, in j                | 41,45                   | (12,97)     | 42,24                  | (12,83)     | 43,19                    | (11,96)     |

AKZ: acute klinische zorg; LKZ: langdurig klinische zorg; n: aantal; %: percentage van de subgroep; M: gemiddelde, SD: standaarddeviatie.

en bijna 3 ernstige problemen, gevolgd door de LKZ-groep (4 aanwezige en bijna 2 ernstige problemen) en de FACT-groep (3,6 aanwezige en 1,5 ernstige problemen). In de laatste drie kolommen worden de F-waarden en de resultaten van Tukeys *honestly significant difference* (HSD) paarsgewijze vergelijkingen weergegeven. Tukeys HSD wees uit dat alle drie de groepen van elkaar verschilden. Hetzelfde gold voor het aantal items dat als 'klinisch relevant' of als 'matig tot ernstig' werd gescoord. Er was dus voldoende reden om aparte normtabellen op te stellen voor de drie groepen.

Met t-testen onderzochten we sekseverschillen. Alleen bij de AKZ-groep vonden we een sekseverschil (mannen scoorden enigszins hoger) met een klein effect (Cohens d was 0,13, 0,16 en 0,13 voor respectievelijk verschil in indexscore, aantal klinisch relevante en aantal ernstige problemen).

Ook vergeleken we 5 leeftijdsgroepen (18-29, 30-39, 40-49, 50-59 en 60+). Uit deze resultaten bleek dat het effect van leeftijd bij de AKZ-groep gering was, maar bij de andere twee zorgtypen (LKZ en FACT-zorg) hadden de jongste twee leeftijdsgroepen (18-39) lagere scores dan

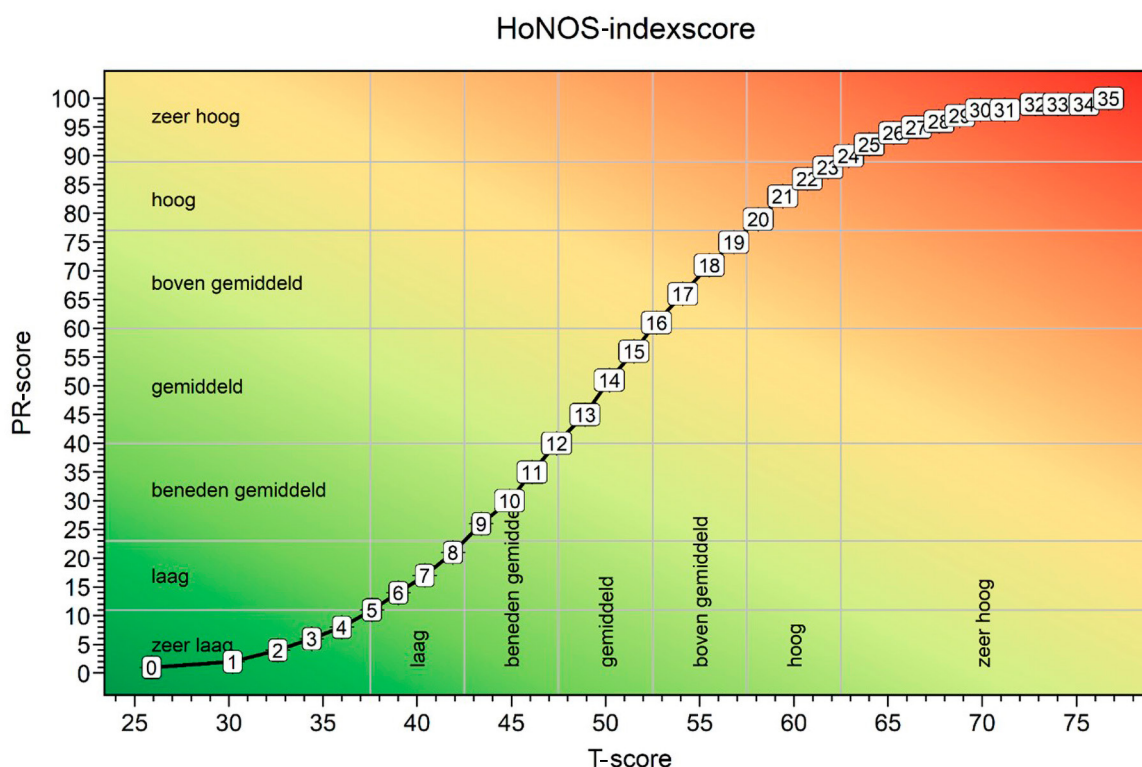
de leeftijdsgroepen 40-60+. Dit was aanleiding om aparte normen op te stellen voor jongere en oudere patiënten.

In **online tabel S-5** geven we de gemiddelde indexscores en het aantal aanwezige en ernstige problemen weer voor de drie zorgtypen, uitgesplitst voor patiënten jonger dan 40 en 40+.

Bij de AKZ-groep gingen we ook na of er een verschil in score was tussen patiënten met en zonder justitiële maatregel. Volgens een t-test (tweezijdig) was er een significant verschil tussen beide groepen:  $t(4303) = 2,75$ ;  $p < 0,01$ ; Cohens  $d = 0,10$  (95%-BI: 0,03-0,17). Het verschil was echter klein. De gemiddelde score bij patiënten met een justitiële maatregel ( $n = 930$ ) was 17,09 (SD: 6,80) en bij vrijwillig opgenomen patiënten ( $n = 3379$ ) 16,39 (SD: 6,86). Patiënten met een justitiële maatregel scoorden dus 0,7 punt hoger dan vrijwillig opgenomen patiënten. Dit verschil rechtvaardigde geen aparte normen. Vergelijkbare bevindingen deden we voor het aantal aanwezige en het aantal ernstige problemen.

Ook bij de FACT-groep vonden we verschillen op alle drie de HoNOS-scores tussen patiënten met en zonder justitiële maatregel, maar nu in omgekeerde richting: de

**Figuur 1. HoNOS-indexscores in 7 scoreklassen**



gemiddelde indexscore bij de groep met een maatregel ( $n = 315$ ) was 9,85 (SD: 6,95) en bij de vrijwillig opgenomen groep ( $n = 2120$ ) 11,04 (SD: 6,69). Bij de LKZ-groep was er geen verschil tussen vrijwillig en met een justitiële maatregel opgenomen patiënten.

In **onlinetabel S-6** geven we de gemiddelde en standaarddeviaties weer voor diverse diagnostische groepen. Er was een verschil tussen de groepen in indexscore:  $F(5,1586,36) = 13,73$ ;  $p < 0,001$ ;  $\eta^2 = 0,008$ . Van de 15 mogelijke paarsgewijze vergelijkingen van de gemiddelde indexscore kwam er bij 6 een significant verschil naar voren: de score bij de bipolaire stoornis was lager dan bij de andere 5 diagnostische groepen en de score bij de groep met een stoornis in middelengebruik was hoger dan bij de psychotische stoornis. De effectgrootte van de paarsgewijze verschillen was eerder klein dan gemiddeld (bereik van  $d$ : 0,12-0,38). Bij het aantal als klinisch beoordeelde items waren er 5 significante verschillen: de groep met een bipolaire stoornis had een lagere score dan de andere 5 groepen. Ten slotte waren er bij de vergelijking van het aantal als ernstig beoordeelde items 4 significante verschillen. De vergelijking bipolaire stoornis met persoonlijkheidsstoornis was hier niet langer significant. We zagen in deze resultaten geen aanleiding om aparte normtabellen voor diagnostische groepen op te stellen, maar we waarschuwen de gebruiker dat bij bipolaire stoornissen gemiddeld iets lagere scores en bij patiënten met een stoornis in middelengebruik iets hogere scores voorkomen.

Met item 8 van de HoNOS kan de ernst van symptomen (anders dan psychose en depressie) worden beoordeeld.

Hierbij kan men aangeven welk symptoom of klacht men beoordeeld heeft. **Onlinetabel S-7** geeft voor de drie groepen de frequentie en het gemiddelde  $M$  (en SD) weer. In de tabel staan  $M$  (en SD) van problemen die vaker dan 15 keer werden gescoord. Dit item bleek vaak niet ingevuld of er was niet aangegeven welke type probleem was beoordeeld. Onder de beoordelingen waarbij het symptoom wel was aangegeven, werden angst, gespannenheid en slaapproblemen het meest gekozen en fobie werd het minst gekozen. Afgaande op de gemiddelde itemscore werden 'angst' en 'somatiserend' als het ernstigst beoordeeld, 'dissociatief' als het minst ernstig.

### Normgegevens

We onderzochten de relatie tussen ruwe scores en genormaliseerde T-scores en kozen daarbij voor toepassing van curvilineaire omzettingsformules. Curvilineaire formules corrigeren de enigszins scheve verdeling van de ruwe HoNOS-scores.<sup>28</sup> Nadere toelichting bij de formules en de afleiding ervan zijn te vinden in het onlinesupplement (tekst bij **figuur S-1**, **S-2** en **S-3**). Verder kozen we voor de curvilineaire formule gebaseerd op de rankit-benadering.  $T_{\text{Rankit}}$ -scores hebben iets meer spreiding en benaderen SD = 10 beter dan  $T_{\text{IRT}}$ , zoals **onlinetabel S-1** laat zien.

De berekening van normgegevens voor de drie groepen cliënten, uitgesplitst in twee leeftijdsgroepen, leverde 8 tabellen op. Deze normtabellen hebben we opgenomen in de aanvullende gegevens (**onlinetabel N-1** tot **N-2-3**). In **figuur 1** hebben we ruwe scores weergegeven als blokjes op de kromme voor de relatie tussen T- en

PR-scores; de corresponderende T- en PR-scores kunnen op de x-as en y-as worden afgelezen. Tevens zijn met horizontale en verticale lijnen de afgrenzingen tussen scoreklassen weergegeven. Uit **figuur 1** valt zo bijvoorbeeld af te lezen dat een indexscore 16-19 'bovengemiddeld' is.

## DISCUSSIE

Uit de scores op de HoNOS bij aanvang van de behandeling bleken er verschillen te zijn in item- en totaalscores tussen de drie hoofdgroepen patiënten: AKZ-, LKZ- en FACT-groep. Als we inzoomen op scores van HoNOS-items, dan zien we bij patiënten uit de AKZ-groep vooral hoge scores op hallucinaties/wanen, gebrek aan sociale contacten en gebrek aan motivatie en lagere scores op zelfverwonding en lichamelijke problemen. Bij patiënten uit de LKZ-groep zien we hoge scores op gebrek aan sociale contacten, motivatie en lagere scores op 'zelfverwonding' en 'maniforme ontremming'. Bij de FACT-groep waren scores op de items 'gebrek aan sociale contacten' en 'overige symptomen' verhoogd en scores op 'woonomstandigheden' en 'maniforme ontremming' lager dan bij de andere groepen. De scores van de FACT-groep komen het meest overeen met een normgroep die voornamelijk uit ambulante behandelde patiënten bestond.<sup>5</sup>

Subgroepen met specifieke stoornissen verschilden ook op de indexscore en hier zien we grofweg hetzelfde patroon bij de drie groepen patiënten. Patiënten met als hoofddiagnose een stoornis in middelengebruik hebben de hoogste scores, patiënten met een bipolaire stoornis de laagste en patiënten met psychotische, depressieve en persoonlijkheidsstoornissen scoren daar tussenin. Verder zien we bij deze vergelijking verhoogde scores op de items die vooral relevantie hebben voor de stoornis (bijvoorbeeld 'hallucinaties en wanen' bij psychotische stoornissen, 'middelen' bij middelengebruik, 'depressie' bij stemmings- en bipolaire stoornissen), wat de validiteit van de HoNOS ondersteunt.

### Ruwe scores, T-scores of PR-scores

Ruwe scores geven maar beperkte informatie over hoe een individu scoort ten opzichte van anderen. Gezien de aanzienlijke verschillen tussen subgroepen patiënten qua zorgtype in ruwe HoNOS-totaalscore zijn de T-scores en PR-scores informatief omdat ze de hoogte van het testresultaat relatief aan de vergelijkingsgroep uitdrukken. T-scores en percentielscores geven zo een duidelijk beeld van waar een individu staat ten opzichte van de normgroep. We krijgen met deze scores een beter beeld van de uitzonderlijkheid van een testresultaat en de relatieve positie van de beoordeelde patiënt ten opzichte van de referentiegroep dan de ruwe score biedt. Voor de keuze tussen T-scores en PR-scores geldt dat – psychometrisch gezien – T-scores de voorkeur verdienen boven percentielscores, omdat ze een intervalschaal

hebben en een pre-postverschilscore (DT) uitdrukken van wat in de behandeling bereikt is. PR-scores hebben deze intervalschaal niet en zijn ordinaal; in verhouding tot T-scores is de percentielschaal rond het gemiddelde uitgerekt met ongeveer  $PR = 4 \cdot T$  en aan de uiteinden juist samengedrukt (zie **figuur 1**). Een verschuiving in PR-score heeft daardoor niet overal op de schaal dezelfde betekenis. Dit impliceert dat je PR-scores niet van elkaar kunt aftrekken.

Daar staat tegenover dat percentielscores eenvoudiger te interpreteren en te communiceren zijn dan T-scores: het percentiel geeft weer welk percentage uit de referentiegroep een lagere score heeft. In de communicatie met patiënten is dat makkelijker uit te leggen dan dat de testuitslag uitzonderlijk is, omdat die 'x standaarddeviaties verwijderd is van het gemiddelde van vergelijkbare patiënten'. Echter, bij veelvuldig gebruik van T-scores zullen zowel behandelaars als cliënten wennen aan de schaal, net zoals dat voor IQ-scores heeft gewerkt. Gebruik van een uniforme meetschaal draagt bij aan een beter begrip van testuitslagen en faciliteert de communicatie tussen behandelaar en patiënt over ROM-uitslagen.<sup>29</sup>

### Sterke kanten en beperkingen aan het onderzoek

De gepresenteerde normen zijn gebaseerd op 10 jaar dataverzameling met de HoNOS en ze zijn representatief voor patiënten van de drie typen zorg binnen Mentrum (onderdeel van Arkin). De normen zijn met state-of-the-art statistiek vastgesteld en gebaseerd op omvangrijke groepen patiënten, wat het mogelijk maakte om in te zoomen op subgroepen patiënten.

Een beperking is dat de door ons onderzochte patiënten niet representatief zijn voor alle Nederlandse patiënten met EPA. De samenstelling van de Amsterdamse bevolking wijkt af van de bevolkingssamenstelling van heel Nederland, onder andere vanwege een hoger percentage mensen met een migratieachtergrond. De prevalentie van ernstige psychische aandoeningen is verhoogd onder deze bevolkingsgroep, waar vele factoren mogelijk een rol in spelen.

Verder is in veel studies een sterke relatie aangetoond tussen verstedelijking en EPA. Het risico op een psychotische stoornis kan wel twee keer zo hoog zijn in de grote stad, ook wanneer gecorrigeerd is voor andere risicofactoren. Naast een hogere prevalentie zijn er ook aanwijzingen dat patiënten met EPA slechter af zijn in de grote stad en ernstigere klachten hebben vanwege een groter sociaal isolement. Het is niet onderzocht of HoNOS-scores van de Nederlandse stedelijke populatie patiënten met EPA daadwerkelijk verhoogd zijn in vergelijking met die in rurale gebieden. Een Australische vergelijking van behandeluitkomsten liet zien dat stedelijke patiënten weliswaar minder baat hebben van behandeling, maar dat hun score bij aanvang van de behandeling gelijk was aan die van patiënten van het platteland. We weten dus niet of de door ons gepresenteerde normen

ook van toepassing zijn voor patiënten uit meer rurale gebieden, want mogelijk zijn scores op de HoNOS daar gemiddeld wat lager.

Een laatste kwestie wat betreft de door ons gepresenteerde normgegevens is de vraag in hoeverre HoNOS-scores stabiel waren over de geanalyseerde jaren 2012-2022. De indexscore bleek over de jaren heen in beperkte mate af te nemen (met 2,6 punten in 10 jaar). Als uitgangspunt voor de normering hebben we daarom de scores van de hele periode gebruikt, omdat zo subgroepen met voldoende grootte gecreëerd konden worden. Maar dat impliceert dat de normwaarden wellicht iets te hoog zijn. Met de hier vastgestelde normen worden nieuwe patiënten als iets minder disfunctioneel ingeschat dan ze in vergelijking met recentere gemiddelde HoNOS-scores zijn. Men kan daarom een correctie van 1-2 indexpunten op de ruwe HoNOS-score in gedachten houden.

## CONCLUSIE

De presentatie van geactualiseerde normen voor de HoNOS voorziet in een leemte. De oorspronkelijke Nederlandse normgegevens zijn van ruim 20 jaar geleden en afkomstig van een specifieke normgroep van 559 patiënten van Riagg-vestigingen in Rotterdam en omgeving. De normengroepen die we in dit artikel presenteren, sluiten beter aan bij de huidige klinische praktijk.

## LITERATUUR

- 1 Wing JK, Beevor AS, Curtis RH, e.a. Health of the Nation Outcome Scales (HoNOS). Research and development. Br J Psychiatry 1998; 172: 11-8.
- 2 AKWA-GGZ. GGZ Standaarden. Zorgstandaard EPA (ernstige psychische aandoeningen). Utrecht: AKWA-GGZ; 2017. [www.ggzstandaarden.nl/zorgstandaarden/ernstige-psychische-aandoeningen](http://www.ggzstandaarden.nl/zorgstandaarden/ernstige-psychische-aandoeningen)
- 3 Delespaul P, Consensusgroep EPA. Consensus over de definitie van mensen met een ernstige psychische aandoening (EPA) en hun aantal in Nederland. Tijdschr Psychiatr 2013; 55: 427-38.
- 4 Bremer-Hoeve S, de Groot E, van Wel EB, e.a. Zorgvraagtypering met de HoNOS+: kijken in een glazen bol? Tijdschr Psychiatr 2022; 64: 416-7.
- 5 Mulder C, Staring A, Loos J, e.a. De Health of the Nation Outcome Scales (HONOS) als instrument voor 'routine outcome assessment'. Tijdschr Psychiatr 2004; 46: 273-84.
- 6 Nugter MA, Teer W. Het gebruik van meetinstrumenten in de langdurende psychiatrie. In: Buwalda VJA, Nugter MA, Swinkels JA, e.a., red. Praktijkboek ROM in de ggz; een leidraad voor gebruik en implementatie van meetinstrumenten. Utrecht: De Tijdstroom; 2011. p. 87-102.
- 7 Buwalda VJA. Routine outcome monitoring in Dutch psychiatry: measurement instruments, implementation and outcome [proefschrift]. Amsterdam: Vrije Universiteit; 2013.
- 8 Kortrijk HE, Mulder CL, Roosenschoon BJ, e.a. Treatment outcome in patients receiving assertive community treatment. Community Ment Health J 2010; 46: 330-6.
- 9 Kortrijk HE, Mulder CL, van der Gaag M, e.a. Symptomatic and functional remission and its associations with quality of life in patients with psychotic disorder in assertive community treatment teams. Compr Psychiatry 2012; 53: 1174-80.
- 10 Tenneij N, Didden R, Velkamp E, e.a. Reliability and validity of the HoNOS-LD and HoNOS in a sample of individuals with mild to borderline intellectual disability and severe emotional and behavior disorders. J Ment Health Res Intellect Disabil 2009; 2: 188-200.
- 11 Veerbeek MA, Voshaar RCO, Pot AM. Psychometric properties of the Dutch version of the Health of the Nation Outcome Scales for older adults (HoNOS 65+) in daily care. Int J Nurs Stud 2013; 50: 1711-9.
- 12 Nusselder K, Nabitz U, Jalink M, e.a. Behandeleffect van een grootstedelijke acute opnamekliniek: een prospectief cohortonderzoek. Tijdschr Psychiatr 2021; 63: 614-22.
- 13 Parabiaghi A, Kortrijk HE, Mulder CL. Defining multiple criteria for meaningful outcome in routine outcome measurement using the Health of the Nation Outcome Scales. Soc Psychiatry Psychiatr Epidemiol 2014; 49: 291-305.

De overige literatuurverwijzingen zijn online te raadplegen.

## AUTEURS

**Edwin de Beurs**, senior onderzoeker, Arkin GGZ, Amsterdam en hoogleraar Klinische psychologie, Universiteit Leiden.

**Jaap Peen**, senior onderzoeker, Arkin GGZ, Amsterdam.

**Udo Nabitz**, senior onderzoeker, Arkin GGZ, Amsterdam.

**Mariken de Koning**, senior onderzoeker, Arkin GGZ, Amsterdam en afd. Psychiatrie, Amsterdam UMC.

### Correspondentie

Prof. dr. Edwin de Beurs ([edwin.de.beurs@arkin.nl](mailto:edwin.de.beurs@arkin.nl)).

Geen strijdige belangen gemeld.

Het artikel werd voor publicatie geaccepteerd op 20-1-2026.

### Citeren

Tijdschr Psychiatr. 2026;68(03):111-117