

Is de ggz-inkoop door zorgverzekeraars ondoelmatig? Een analyse van 2011-2016 Vektis data.

Jim van Os

Hoogleraar Psychiatrische Epidemiologie en Publieke GGZ

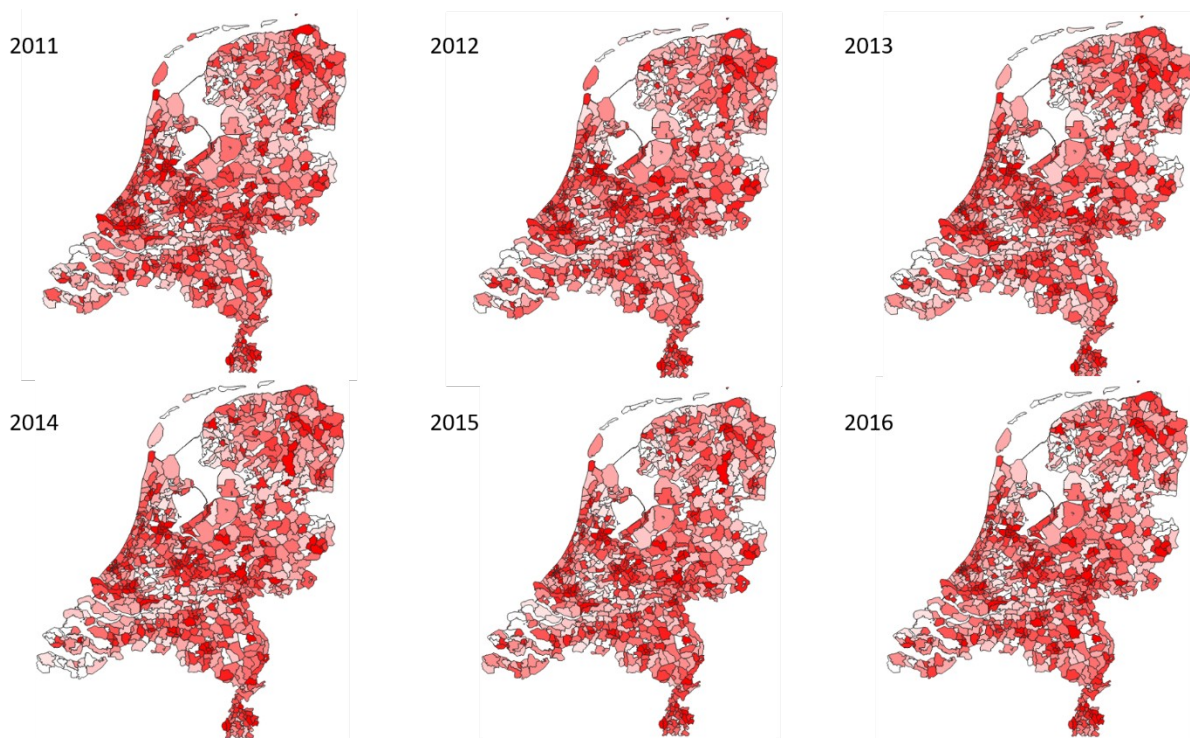
Voorzitter Divisie Hersenen

UMC Utrecht

Volgens cijfers van Vektis maken momenteel 1.1 miljoen volwassenen gebruik van de ggz, ongeveer twee keer zoveel als voor de introductie van de marktwerking. Er is geen mechanisme dat zorgt voor een goede selectie bij de ingang van de ggz, zodanig dat de mensen met de meeste zorgbehoeften ook de meest intensieve zorg krijgen. Ook is er geen mogelijkheid tot toezicht op een juiste match tussen het specifieke ggz-aanbod in een regio en de specifieke zorgbehoefte van de populatie.

In dit artikel werd getracht de gevolgen hiervan in kaart te brengen middels een analyse van variatie in ggz-consumptie (dus ggz-inkoop door zorgverzekeraars) op regioniveau, gebruik makend van de openbare Vektis dataset, gebaseerd op declaraties over de periode 2011-2016. Vektis data werden hiertoe met behulp van geocoding-software geprojecteerd op de kaart van Nederland.

In de figuur hieronder ziet u de gemiddelde ggz-kosten per inwoner, in de leeftijdscategorie 18-65 jaar, per 3-cijferig postcodegebied en per jaar over de periode 2011-2016. Een PC3-gebied telt mediaan ongeveer 16.000 inwoners; Nederland telt ongeveer 800 PC3-gebieden (ter vergelijking: er zijn ongeveer 4000 PC4- en ongeveer 90 PC2-gebieden).



De cijfers werden om methodologische redenen enigszins bewerkt, zoals elders beschreven (Van Os, 2019), daar zal ik u nu niet mee lastig vallen. Wat telt is dit: er zijn 10 tinten rood, en hoe roder het

gebied, hoe hoger de gemiddelde ggz-kosten per inwoner per jaar. Het verschil tussen de tinten is behoorlijk groot: in de orde van 0-100 euro per persoon voor een wit gebied, tot 260-330 euro voor het meest rode gebied. Dus het kan makkelijk tot een factor 10 schelen. Ook is belangrijk te weten dat de uitschieters eruit zijn gehaald: leeftijd-geslacht clusters in de Vektis data met kostwaarden groter dan 2 x MAD (*median absolute deviation*) boven de mediaan werden geëxcludeerd (Van Os, 2019).

Deze kosten reflecteren het ggz-inkoop beleid van de zorgverzekeraars. Dus een analyse van deze kosten is een analyse van de rationaliteit van het inkoopbeleid van deze verzekeraars.

De regionale ggz-inkoop is zeer stabiel van jaar tot jaar

Uit de figuur kunnen we direct al twee belangrijke dingen afleiden met betrekking tot de zorginkoop van ggz-zorg, te weten:

1. Er zijn forse verschillen in gemiddelde ggz-kosten voor volwassenen tussen PC3-regio's.
2. Deze verschillen zijn stabiel van jaar tot jaar.

Met andere woorden: het regionale inkoopbeleid van zorgverzekeraars is zeer stabiel. Zo stabiel, zou je kunnen zeggen, dat ze evengoed elk jaar een bepaald bedrag hadden kunnen overmaken naar elk PC3-gebied, want hoeveel elke regio krijgt relatief aan een ander staat al grotendeels vast. Lekker makkelijk – scheelt een hoop rompslomp.

Klopt de stabiele toewijzing van regionale middelen door zorgverzekeraars?

Dit is de hamvraag: klopt het eigenlijk wel dat de zorgverzekeraars jaar in jaar uit een extreem gedifferentieerd maar tegelijkertijd extreem stabiel patroon van ggz-zorginkoop hanteren, waarbij grote verschillen tussen regio's jaar in jaar uit blijven gehandhaafd? Waar zijn die verschillen eigenlijk op gebaseerd? Zijn de extreem grote verschillen in ggz-inkoop gebaseerd op verschillen in ggz-zorgbehoefte of is het een gevolg van extreme ondoelmatigheid, bijvoorbeeld door verschillen in ggz-aanbod?

Als u deze vraag aan de zorgverzekeraars stelt krijgt u niet echt een antwoord. Dat komt omdat zorgverzekeraars kijken naar individuele instellingen en hun DBC-gedrag. Ze kijken veel minder naar de hele populatie in regio's en hoeveel ggz de ene regio in vergelijking met de andere consumeert, en of dat ergens op gebaseerd is. Concurrerende zorgverzekeraars hebben alleen data van hun eigen klanten – niet van een hele regio. Ze zijn dus relatief blind voor wat er in een regio gebeurt – tenzij alle verzekerden in een regio bij dezelfde zorgverzekeraar zitten.

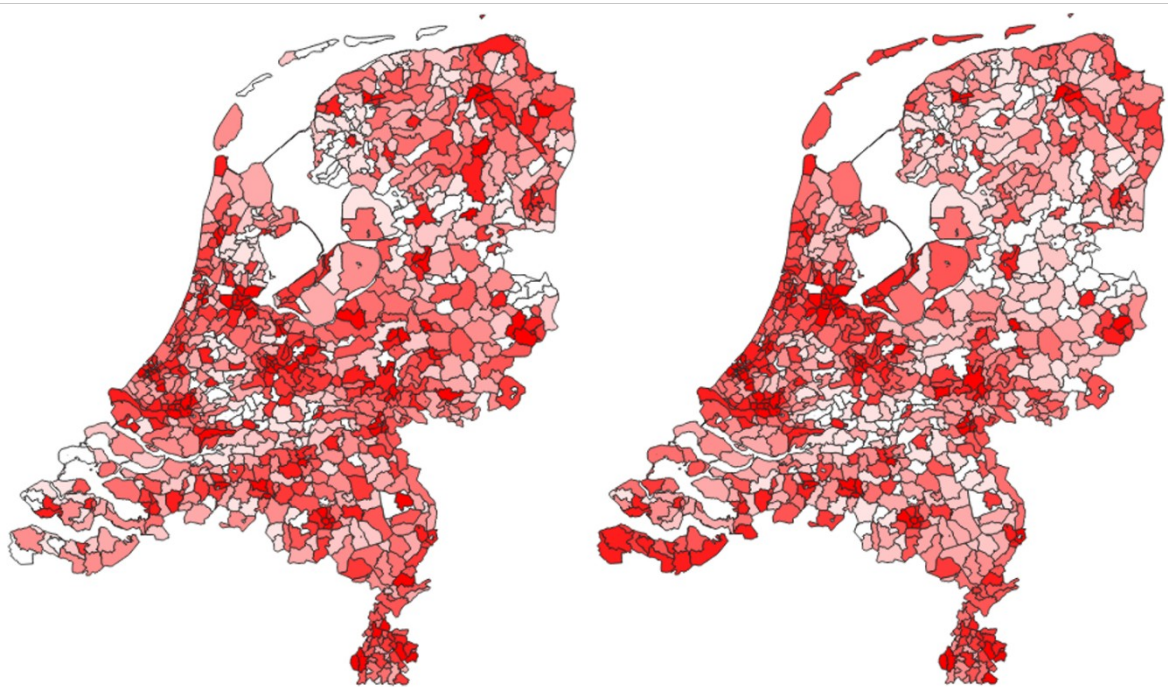
Kunnen de regionale verschillen in ggz-inkoop worden verklaard door demografische verschillen?

Nu zullen de epidemiologisch gevormden onder u willen toetsen of de extreme verschillen in ggz-zorginkoop volumes te maken hebben met demografische indicatoren van ggz-problematiek. Het is immers bekend dat ggz-problematiek sterk samenhangt met jongere leeftijd en een hogere mate van stedelijkheid. Als je hiervoor corrigeert zullen de verschillen – statistisch uitgedrukt als het PC3-gebied *random effect* –tussen de regio's dus wellicht afnemen. Dit hebben wij onderzocht in de data, maar wat blijkt: correctie voor leeftijd, geslacht en stedelijkheid verklaart hoogstens 15% van het PC3-gebied *random effect* (Van Os, 2019).

Wat meer verklarende impact heeft op de verschillen in ggz-inkoop is de sociaaleconomische score van de PC3-gebieden. Deze sociaaleconomische score werd berekend op PC3-niveau op basis van een principale component analyse van de variabelen: percentage vrouwen, percentage allochtonen, percentage Westerse allochtonen, percentage niet-Westerse allochtonen, percentage eenpersoonshuishoudens, percentage meerpersoonshuishoudens met kinderen, percentage meerpersoonshuishoudens zonder kinderen en gemiddelde huishoudgrootte. Het is natuurlijk bekend dat sociaaleconomische factoren een belangrijke impact hebben op de populatie ggz-zorgbehoefte. En dit blijkt ook in Nederland het geval: De sociaaleconomische score verklaart ongeveer 55% van de regionale verschillen in ggz-inkoop. De sociaaleconomische score van het PC3-gebied lijkt dus een belangrijke zorgbehoefte-indicator. Nou is 55% niet niets, maar het laat nog 45% van de extreme regionale verschillen in ggz-kosten onverklaard – en dat gaat om een hoop geld. Dit is goed te zien in de figuur hieronder: de PC3-verdeling van de rode tinten links (kosten) lijkt op die van rechts (sociaaleconomische score) maar er zijn ook belangrijke verschillen.

PC3 kostenverdeling 2011-2016

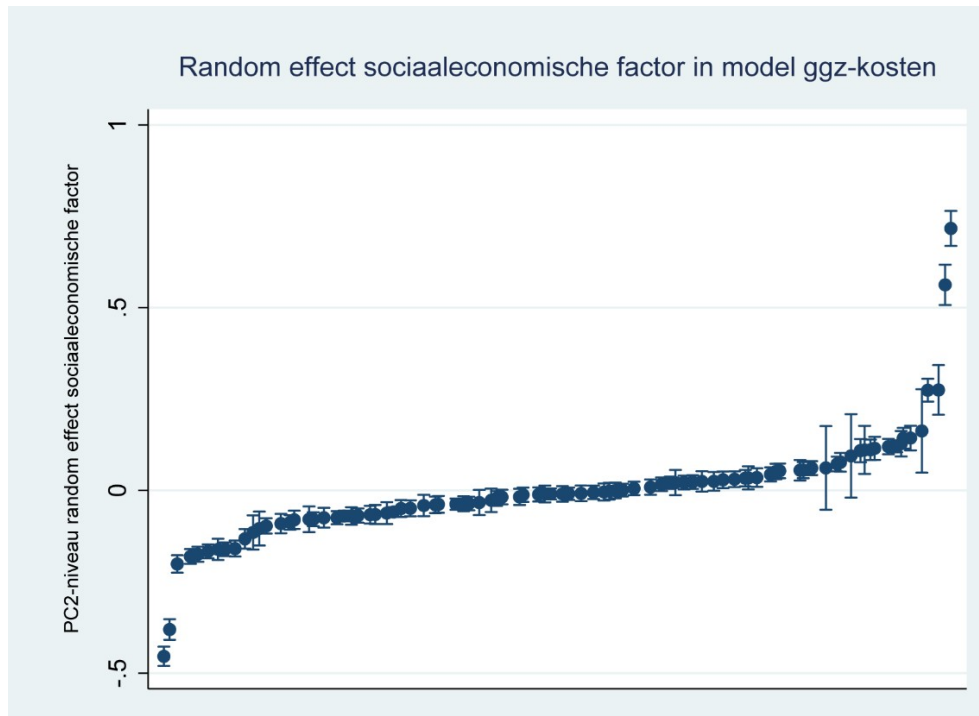
PC3 sociaaleconomische score



Zorgverzekeraars hanteren geen benchmark van ggz-inkoop in relatie tot ggz-zorgbehoefte

Nu we weten dat de sociaaleconomische score van een PC3-gebied een belangrijke voorspeller is van ggz-zorgbehoefte, kunnen we kijken in hoeverre de zorgverzekeraars dit hanteren als een *benchmark* bij de zorginkoop. Je hoopt immers dat zorgverzekeraars hun inkoopbeleid op dezelfde wijze zullen laten afhangen van de populatie ggz-zorgbehoefte, als aangegeven door de sociaaleconomische score. Hiertoe hebben we een analyse gedaan van de mate van variatie waarin de PC3-sociaaleconomische score binnen PC2-gebieden samenhangt met de PC3-zorgkosten over de periode 2011-2016 in de Vektis data (Van Os, 2019).

Het resultaat ziet u hieronder: er blijkt een groot verschil in de mate waarin de regionale zorgkosten (zorginkoop) samenhangt met de regionale sociaaleconomische score. Zorgverzekeraars vertonen dus zeer verschillend gedrag in het afstemmen van de ggz-inkoop op de ggz-zorgbehoefte van de populatie in kwestie. Er is geen benchmark. Oneerbiedig gezegd: Iedereen doet maar wat.



Conclusie: noodzaak voor doelmatigheidscontrole bij zorgverzekeraars?

De conclusie van onze analyse is dat de enorme en kostbare complexiteit van ggz-zorginkoop door zorgverzekeraars in feite niet nodig is omdat PC3-regio's jaar in jaar uit stabiel verschillend worden bedield door de zorgverzekeraars.

Waarom die verschillen er zijn is niet goed bekend: 45% van de verschillen in regionale ggz-zorginkoop kan niet worden verklaard door bekende populatie indicatoren van ggz-zorgbehoefte. Bijna de helft van de regionale verschillen in ggz-inkoop is dus mogelijk irrationeel.

Er daar waar de regionale verschillen wel verklaard kunnen worden op basis van een sociaaleconomische score, is een benchmark van hoe zorgverzekeraars hiermee omgaan bij de inkoop van ggz-zorg ver te zoeken.

Maar misschien de belangrijkste conclusie is: hoe kan het dat er evidente discrepanties lijken te zijn tussen zorginkoop en zorgbehoefte op populatieniveau? In Engeland worden dit soort type populatiegerichte analyses door *Public Health England* en academische centra al heel lang gebruikt als stuurinformatie om ggz-zorgaanbod en ggz-zorgbehoefte op populatieniveau op elkaar af te stemmen. Maar niet in Nederland. Hier lijken we in de ban van de idee dat de zorg moet worden aangestuurd als een consumentenbedrijf. De cijfers waar we op willen sturen in de ggz zijn pre-post symptoomscores, die wetenschappelijk bewezen niet betrouwbaar, niet valide, niet relevant en niet privacy-compatibel zijn (Hafkenscheid and van Os, 2013, Hafkenscheid and Van Os, 2014b,

Hafkenscheid and Van Os, 2014a, Hafkenscheid and van Os, 2016, Hafkenscheid and van Os, 2018, van Os *et al.*, 2014, van Os *et al.*, 2017, van Os *et al.*, 2018, Van Os and Delespaul, 2018, van Os *et al.*, 2012). Of de miljarden die we uitgeven aan ggz überhaupt in verhouding staan tot de populatiezorgbehoefte kan ons blijkbaar niet boeien, ook al ligt het bewijs voor het oprapen in openbare databases. Antiwetenschap als onderlegger van gezondheidszorgbeleid.

Bibliografie

- Hafkenscheid A, Van Os J** (2013). [Current routine outcome monitoring undermines valid quality measurement]. *Tijdschr Psychiatr* **55**, 179-81.
- Hafkenscheid A, Van Os J** (2014a). Naar een deugdelijke ROM. *Maandblad Geestelijke Volksgezondheid* **69**, 20-28.
- Hafkenscheid A, Van Os J** (2014b). ROM van geïndividualiseerde behandeldoelen. *Psychopraktijk* **6**, 29-32.
- Hafkenscheid A, Van Os J** (2016). [ROM measurements in mental health care: users need to be aware of the problems and pitfalls]. *Tijdschr Psychiatr* **58**, 388-96.
- Hafkenscheid A, Van Os J** (2018). Twee misvattingen over ROM. *De Psycholoog in de pers*.
- Van Os J** (2019). Loopt de inkoop van ggz-zorg uit de pas met de populatiezorgbehoefte? Onverklaarde regiovariatie van ggz-kosten en van associaties tussen ggz-output en populatiezorgbehoefte. *submitted*.
- Van Os J, Berkelaar J, Hafkenscheid EaA** (2017). Benchmarken: doodlopende weg onder het mom van 'ROM'. *Tijdschr Psychiatr* **59**, 247-250.
- Van Os J, Delespaul P** (2018). Een Valide Kwaliteitskader Voor De GGZ: Van Benchmark-ROM Aan De Achterkant Naar Regionale Regie en Co-creatie Aan De Voorkant. *Tijdschrift voor Psychiatrie* **60**, 96-104.
- Van Os J, Delespaul P, Barge D, Bakker RP** (2014). Testing an mHealth momentary assessment Routine Outcome Monitoring application: a focus on restoration of daily life positive mood states. *PLoS One* **9**, e115254.
- Van Os J, Guloksuz S, Vijn TW, Hafkenscheid A, Delespaul P** (2018). The diagnosis-evidence-based group-level symptom-reduction model as the organizing principle for mental health care: time for change? *World Psychiatry in press*.
- Van Os J, Kahn R, Denys D, Schoevers RA, Beekman AT, Hoogendijk WJ, Van Hemert AM, Hodiament PP, Scheepers F, Delespaul PA, Leentjens AF** (2012). [Behavioural standard or coercive measure? Some considerations regarding the special issue on ROM]. *Tijdschr Psychiatr* **54**, 245-53.