

**Van:** [redacted] <[redacted]@autoriteitpersoonsgegevens.nl>

**Verzonden:** 21-11-2022 10:05

**Aan:** [redacted] <[redacted]@autoriteitpersoonsgegevens.nl>

**Onderwerp:** FW: Reactie advies over zorgvraagtypering

fyi

**Van:** [redacted] <[redacted]@nza.nl>

**Verzonden:** donderdag 17 november 2022 15:46

**Aan:** [redacted] <[redacted]@autoriteitpersoonsgegevens.nl>

**CC:** [redacted] <[redacted]@autoriteitpersoonsgegevens.nl>; [redacted] <[redacted]@autoriteitpersoonsgegevens.nl>; [redacted] <[redacted]@autoriteitpersoonsgegevens.nl>; [redacted] <[redacted]@nza.nl>; [redacted] <[redacted]@nza.nl>; [redacted] <[redacted]@nza.nl>; [redacted] <[redacted]@nza.nl>; [redacted] <[redacted]@nza.nl>

**Onderwerp:** Reactie advies over zorgvraagtypering

Beste [redacted]

Zoals vorige week afgesproken hebben wij een aantal aspecten in het document over de gegevensaanlevering zorgvraagtypering extra onderbouwd.

In de bijlage het document, met de aanvullingen geel gearceerd.

Ik vervang [redacted] tijdens zijn vakantie, vandaar deze mail vanuit mij.

Met vriendelijke groet,

[redacted]

Beleidsmedewerker | BGGFZ

06 – [redacted]

Aanwezig op maandag t/m vrijdag



**Nederlandse Zorgautoriteit**

[info@nza.nl](mailto:info@nza.nl) | [www.nza.nl](http://www.nza.nl)

[LinkedIn](#) | [Twitter](#)

**Disclaimer NZa**

Dit bericht kan informatie bevatten die niet voor u is bestemd. Indien u niet de geadresseerde bent of dit bericht per abuis aan u is toegezonden, wordt u vriendelijk verzocht dat aan de afzender te melden en het bericht te verwijderen. De NZa aanvaardt geen aansprakelijkheid voor schade, van welke aard ook, die verband houdt met risico's verbonden aan het elektronisch verzenden van berichten.



## Memo

### Aan

Autoriteit Persoonsgegevens

### Onderwerp

Reactie op advies over gegevensaanlevering zorgvraagtypering

### Datum

17 november 2022

# 1 Aanleiding

De AP heeft advies uitgebracht over de Regeling geestelijke gezondheidszorg (ggz) en forensische zorg (de Regeling), specifiek op het punt van aanlevering aan de NZa van HoNOS+ gegevens ten dienste van de ontwikkeling van het stelsel van zorgvraagtypering. Samenvattend is de AP van oordeel dat de doelstellingen van deze aanlevering onvoldoende welbepaald zijn, dat mede als gevolg daarvan de noodzaak niet kan worden vastgesteld en dat mogelijke alternatieven onvoldoende gemotiveerd zijn verworpen. De Regeling en toelichting kunnen daarom naar het oordeel van de AP niet een grondslag bieden voor rechtmatige verwerking van de HoNOS+ gegevens.

De NZa heeft privacy hoog in het vaandel staan en hecht er daarom aan te reageren op de punten zoals door de AP benoemd in haar brief. De reactie is in dit document uitgewerkt. De eerdere documenten, zoals aan de AP aangeleverd op 19 september 2022 worden bekend verondersteld. De brief van de AP heeft de NZa doen realiseren dat de gegevensuitvraag zorgvraagtypering beter in de tijd moet worden gedefinieerd, snel in volume moet worden teruggebracht en dat de NZa zich op het gebied van het gebruik van de gegevens beter moet binden. Daarom worden, in lijn met adviezen uit de voornoemde brief, een aantal wijzigingen doorgevoerd in de (toelichting bij de) Regeling (zie hoofdstuk 5). Specifiek zullen we een horizonbepaling, doelbinding en de bewaartermijn van de gegevens opnemen in de toelichting bij de Regeling.

# 2 Verwerkingsdoeleinden

De AP heeft aangegeven dat het verwerkingsdoel 'ontwikkeling van zorgvraagtypering' onvoldoende nauwkeurig is bepaald en dat ten minste de concrete ontwikkeldoelen- en fasen zorgvuldig en nauwkeurig vastgelegd moeten zijn, zodat per ontwikkelfase duidelijk wordt in welke mate en voor welk exact doel HoNOS+ gegevens in die fase (nog) noodzakelijk zijn. Noodzaak van verwerking (subsidiariteit en proportionaliteit) is hier aan gerelateerd en zullen hieronder gezamenlijk worden behandeld. Ter toelichting strekt het volgende.

## 2.1 Inleiding

Er zijn grote problemen in de geestelijke gezondheidszorg. Er is geen financiële prikkel voor het verlenen van passende zorg aan patiënten met complexe zorgvragen, daarom staan deze patiënten vaak maanden langer op wachtlijsten dan wenselijk voordat ze in zorg genomen kunnen worden (zie rapport Algemene

rekenkamer). De Algemene Rekenkamer trekt de conclusie dat hoe complexer de problematiek is, hoe langer de wachttijd. Dat geldt zeker als een ernstige psychische aandoening gepaard gaat met andere problemen of als de patiënt verstandelijk beperkt is. Eén van de oorzaken van de te lange wachttijden is dat financiële prikkels in de bekostiging van de ggz het voor de zorgaanbieders aantrekkelijk maken om zich vooral te richten op patiënten met minder zware problemen. Dit leidt tot onnodig lijden bij de persoon met geestelijke klachten en voedt tegelijk maatschappelijke onrust over verwarde mensen op straat (zie rapport Trimbos en publicatie politie).

Met de invoering van het Zorgprestatiemodel per 1 januari 2022 zijn de eerste stappen gezet richting het oplossen van de problemen in de ggz. Voorheen was het voor zorgaanbieders financieel aantrekkelijker om patiënten met lichte psychische problemen in behandeling te nemen, dan mensen met ernstige, complexe psychische problemen. Met het Zorgprestatiemodel kunnen we passende zorg beter belonen. Zorgaanbieders die patiënten met een complexe zorgvraag behandelen, kunnen door het Zorgprestatiemodel beter betaald krijgen omdat we de behandelkosten beter kunnen verdelen. Het is echter nog niet goed mogelijk de complexe patiënten te herkennen (zie rapportage SiRM) en in te delen in vergelijkbare groepen (zie eindrapportage pilot zorgclustermodel), daarom kunnen zorgaanbieders en verzekeraars nog niet goed overleggen over passende vergoedingen voor de behandeling. Herkennen van de complexiteit van zorgvragen zal gedaan worden met een goed functionerend systeem van zorgvraagtypering. Het systeem van Zorgvraagtypering werkt op dit moment echter nog niet goed genoeg en moet daarom verder worden ontwikkeld (eigen onderzoek en eindrapportage). In de rest van dit document werken we de knelpunten, verbeteringen en route naar de verbeteringen verder uit.

Wanneer zorgvraagtypering niet verder ontwikkeld kan worden zal het niet goed geaccepteerd worden door de behandelaren, kunnen zorgaanbieders het niet gebruiken voor plannen van zorg en kunnen verzekeraars en aanbieders geen gedifferentieerde afspraken maken over passende vergoedingen bij verschillende groepen patiënten. In dat geval is er geen goed alternatief om te sturen op passende zorg, is het niet mogelijk om complexe patiënten sneller en beter vergoed in zorg te brengen en zullen de wachtlijstproblemen en gevolgen daarvan voor de patiënt en maatschappij blijven voortbestaan.

Bij het verder ontwikkelen op basis van gegevens uit de praktijk gelden de volgende uitgangspunten:

- Zonder goed functionerende zorgvraagtypering is er geen geloofwaardig instrument om grote maatschappelijke problemen, zoals continu stijgende zorgkosten, verlieslatende zorg voor complexe patiënten en lange wachttijden te bestrijden.
- Twee specifieke hoofddoelen moeten worden bereikt voordat zorgvraagtypering goed functioneert. Specifiek is het nodig dat 1) er een betrouwbaar algoritme gemaakt wordt om van vragenlijst tot zorgvraagtype te komen en 2) de verdeling van zorgvragen en zorginzet geijkt wordt op de zorgvraagtypes in de Nederlandse ggz-praktijk. Voor beide doeleinden zijn gegevens nodig.
- Minder gegevens verzamelen is beter, maar we weten nog te weinig over de mogelijke zorgvragen in de Nederlandse praktijk en hoe deze zich verhouden tot de zorgvraagtypes. Daarom is in 2022 en 2023 informatie nodig uit het zorgveld over de volle breedte van de zorgvragen, zodat we de zorgvraagtypering en algoritmes goed kunnen ijken.
- Vanaf 2024 is brengen we de gegevensuitvraag sterk terug. Vanaf dat moment kan worden volstaan met een monitorings-uitvraag, aangevuld met eventuele uitvragen waar specifieke hypothesegedreven informatiebehoefte leeft.



## 2.2 Algoritme zorgvraagtypering ijken

### 2.2.1 Hoofddoel

Om een tweetal redenen is het belangrijk dat er minstens één goed functionerend algoritme bestaat. Het algoritme is namelijk enerzijds een manier om vertrouwen tussen zorgaanbieder en zorgverzekeraar over het gekozen zorgvraagtype te waarborgen en anderzijds een manier om de behandelaar te adviseren over de best passende zorgvraagtypes bij de zorgvraag. Hieronder werken we beide punten verder uit.

Ten eerste is zorgvraagtypering een systeem dat gebruikt moet worden om de zorginkoop vorm te geven. Als we zorgvraagtypering vanuit die rol bekijken hebben zorgaanbieders en zorgverzekeraars tegengestelde belangen, een lichter (goedkoper) zorgvraagtype is financieel gezien gunstig voor de verzekeraar, maar nadeliger voor de aanbieder dan een complex of duur zorgvraagtype. De behandelaar is vrij om een zorgvraagtype te kiezen en er is niet genoeg vertrouwen bij de verzekeraars dat die keuze puur ingegeven zal zijn door de zorgvraag van de patiënt. Daarnaast is er onvoldoende vertrouwen bij de zorgaanbieders dat de verzekeraars in hun inkoop niet onterecht zullen sturen op de lichtere zorgvraagtypes. De behandelaar moet het zorgvraagtype kiezen dat het beste bij de zorgvraag van zijn patiënt past, het is daarbij te verwachten dat het gekozen zorgvraagtype soms anders is dan het advies van het algoritme. Een niveau van afwijken van het model zal ook door de verzekeraar verwacht worden, maar er moet een manier zijn om na te gaan dat die keuze niet altijd strategisch dezelfde kant op valt. De mogelijkheid tot verantwoording wordt gevormd door de algoritmes die we zullen maken op basis van de uitvraag zorgvraagtypering. De huidige algoritmes zijn niet goed genoeg in het leggen van de verbanden en zijn vooral niet goed in het vinden van de uitzonderingsgevallen (complexe zorgvragen) waar de belangen van verzekeraar en aanbieder juist het grootst zijn.

Ten tweede is het belangrijk dat de behandelaar die zorgvraagtypering uitvoert goed op de hoogte is van de waarschijnlijke zorgvraagtypes bij de zorgvraag van zijn patiënt. Om tot een zorgvraagtype te komen vult een behandelaar (een deel van) de HoNOS+-vragenlijst in. Eén tweetal algoritmes adviseert de behandelaar daarna over de meest waarschijnlijke zorgvraagtype(s) bij de vragenlijst. In het geval van het 'dynamische' algoritme wordt op basis van een subset van de HoNOS+-vragen één enkel zorgvraagtype het meest waarschijnlijk geacht. In het geval van het 'Britse' algoritme wordt aan de hand van de volledige vragenlijst een aantal meest waarschijnlijke zorgvraagtypes gegeven met bij elke een waarschijnlijkheidspercentage.

Deze algoritmes zijn nog niet goed genoeg. Het advies van de algoritmes wijkt vaker dan nodig af van de keuze van de behandelaar (eigen onderzoek), de algoritmes zijn vaak zeer stellig in de keuze, terwijl de behandelaar vaak afwijkt (eigen onderzoek), het 'dynamische' algoritme wijst nog niet vaak genoeg naar het juiste meest waarschijnlijke zorgvraagtype (signalen uit het veld, onderzoek SiRM, eigen onderzoek pilots en simulaties) en in veel gevallen worden nog meer (bijzondere) persoonsgegevens uitgevraagd dan misschien nodig is met het dynamische algoritme (onderzoek SiRM en eigen onderzoek op pilotgegevens en simulaties). Om de algoritmes te verbeteren is het nodig gegevens op te vragen over de zorgvragen (HoNOS+-vragenlijst en bijbehorende elementen, zoals zorgvraagtype). Op basis van de gegevens kunnen drie benodigde subdoelen in het ijken van de algoritmes worden gehaald. Na behalen van deze mijlpalen zal de NZa een veel kleinere groep zorgvraagtyperingen nodig hebben voor het verder onderzoeken en verbeteren van de algoritmes. Op dat moment gaan twee andere subdoelen een rol spelen (4 en 5). Subdoelen 1 t/m 4 hieronder hebben betrekking op de algoritmes, subdoel 5 is breder en gaat ook over de ijking van de zorgvraagtypes.



Voor het behalen van deze doeleinden zijn alle HoNOS+-vragenlijsten die gedurende 1,5 jaar zijn geregistreerd nodig met bijbehorende behandelinzet:

1. Verbeteren en actualiseren van het 'dynamische' algoritme (beslisboom) dat het mogelijk moet maken met (veel) minder vragen tot een zorgvraagtype te komen.
2. Het vervangen van het 'Britse algoritme voor een beter functionerend en beter gekalibreerd alternatieve classificeerder, vermoedelijk in de vorm van een zogeheten *random forest* model.
3. Inzicht in de 'gebieden' van de zorgvraagruimte (combinaties van scores op de HoNOS+) waar groepen patiënten voorkomen met complexe problematiek.

Voor het behalen van deze doeleinden kunnen we volstaan met een subset:

4. Geijkt houden van de algoritmes op de langere termijn. Zowel op de behandelinzet als op elkaar.
5. Onderzoeken en verbeteren van zorgvraagtypering naar aanleiding van bijzonderheden.

Zonder goed werkende algoritmes zal enerzijds de acceptatie van zorgvraagtypering bij de behandelaren niet goed zijn en is het risico groot dat zorgvraagtypering niet serieus gebruikt zal worden in de behandelkamer. Anderzijds kunnen verzekeraars de zorgvraagtypes niet goed op waarde schatten zonder goed algoritme. Als zorgvraagtypering niet goed wordt ingevuld en de verzekeraars de betrouwbaarheid van de zorgvraagtypes niet kunnen inschatten verliest zorgvraagtypering zijn nut in de zorginkoop. Er is in dat geval geen goede manier om ervoor te zorgen dat patiënten met een complexe zorgvraag de zorg krijgen die ze nodig hebben tegen een reële vergoeding en er is dan geen goed alternatief om de wachtlijstproblemen te verminderen.

## 2.2.2 Subdoelen

Zonder goed werkende algoritmes kan zorgvraagtypering zijn de doelen niet waar maken. Om van de zorgvraag van een patiënt tot een zorgvraagtype te komen moeten twee stappen worden gezet. Allereerst moet de zorgvraag systematisch in kaart worden gebracht. Tijdens een voortraject sinds 2015 met verkenningen, pilots en discussies met inhoudelijke experts in werkgroepen en panels hebben we verkend welke methode het meest geschikt is voor het meten van de zorgvraag. De HoNOS+-vragenlijst bleek het meest geschikte middel, omdat het met een relatief klein aantal vragen niet alleen klachten en symptomen, maar ook het sociaal functioneren en de leefsituatie van de patiënt in kaart brengt. Als tweede moet de behandelaar een advies krijgen over de meest waarschijnlijke zorgvraagtypes bij de ingevulde zorgvraag. Dit advies moet ook gebruikt kunnen worden om de link tussen groepen zorgvragen en zorgvraagtypes te verantwoorden bij eventuele materiële controles. Voor het adviseren over het meest waarschijnlijke zorgvraagtype hebben we een systeem gekopieerd dat in Groot-Brittannië is ontwikkeld. Dit is een model (algoritme) dat de behandelaar adviseert over welke zorgvraagtypes waarschijnlijk zijn bij de opgegeven zorgvraag. Er is veel winst te behalen in het aanpassen van het algoritme op de Nederlandse situatie. Die verbeteringen zijn nodig om het model geaccepteerd te krijgen en om de zorgvraagtypering goed genoeg te krijgen zodat het bruikbaar is in de zorginkoop. Om de verbeteringen tot stand te brengen heeft de NZa gegevens nodig. De verbeteringen zullen lopen via de onderstaande subdoelen.

Subdoel 1. Dynamisch algoritme Het is namelijk waarschijnlijk dat niet alle vragen over klachten en symptomen voor alle patiënten relevant zijn. Het is dus mogelijk dat er met de volledige HoNOS+ voor sommige patiënten meer gegevens worden geregistreerd dan strikt noodzakelijk om een waarschijnlijk

zorgvraagtypes aan te wijzen. We weten alleen nog niet voor welke patiënten welke gegevens belangrijk zijn. Dit onderwerp hebben we in het najaar van 2020 met adviesbureau SiRM verkend (zie onderzoek SiRM). De insteek was dat de administratieve belasting bij behandelaren omlaag kan als niet alle vragen beantwoord hoeven te worden. Dit heeft ook een positief gevolg voor de privacy van de patiënt. Er worden dan immers minder gegevens over zijn gezondheid geregistreerd door de aanbieder en verzonden aan de NZa. De uitwerking heeft als basis dat alleen relevante vragen gesteld worden, dit betekent dat één vraag wordt gesteld en vervolgens relevante vragen worden bepaald aan de hand van de ingevulde scores in de vorm van een beslisboom. In veel gevallen hoeven niet alle 19 HoNOS+-vragen gescoord te worden, maar kan een behandelaar voor grote groepen patiënten al na 5 of 6 vragen een zorgvraagtype kiezen (eigen onderzoek). De beslisboom kan alleen goed worden geschat als we goed weten wat relevante combinaties van scores zijn en welke vragen daaruit de meest relevante vervolgvragen en welk zorgvraagtype uiteindelijk wordt gekozen. Daarvoor is goed zicht nodig op de voorkomende zorgvragen en de invloed van de verschillende onderdelen van de zorgvraag op de behandelinzet.

Subdoel 2. Vervangen 'Britse' algoritme Daarnaast is het huidige 'Britse' algoritme geschat op een steekproef van zorgvragen van Britse patiënten en blijkt geen goede weerspiegeling te zijn van de keuze die Nederlandse behandelaren maken (eigen onderzoek). Tijdens een pilot hebben we HoNOS+-gegevens opgehaald waarmee we het model in de Nederlandse praktijk hebben getest. Het blijkt dat het model in iets meer dan 50% van de gevallen hetzelfde zorgvraagtype aanwijst als dat de behandelaar kiest. Een vergelijkbare afwijking blijkt ook te bestaan als het model aangeeft 90% zeker te weten wat het meest waarschijnlijke zorgvraagtype is. We horen van behandelaren dat het als erg onprettig wordt ervaren om af te wijken van een algoritme dat zo stellig is, het zorgt voor twijfel en staat acceptatie van zorgvraagtypering in de weg. Om deze reden hebben we geprobeerd op de (te beperkte) gegevens uit de pilots alternatieve modellen te schatten. De verwachting van de alternatieve modellen is wat meer in lijn met de keuze van de behandelaar en geeft vooral een realistischere schatting van de zekerheid. Dit zal goed zijn voor de acceptatie van zorgvraagtypering door de behandelaren en minder risico geven op ongewenste beïnvloeding van de behandelaar door een 'overmoedig' algoritme. We weten dus dat het model dat de zorgvraag en zorgvraagtype aan elkaar verbindt preciezer geschat kan worden met gegevens uit de Nederlandse praktijk. Er zijn vooralsnog alleen nog niet voldoende gegevens beschikbaar om die verbinding te kunnen leggen.

Subdoel 3. Inschatten ruimte zorgvragen De twee bovenstaande verbeteringen hebben pas echt impact als duidelijk is welke combinaties van scores veel voorkomen en welke combinaties veel impact hebben op de verdere zorginzet. Om daar achter te komen moeten we richting 'verzadiging' gaan van de ruimte waarin relevante zorgvraagtyperingen voorkomen, we moeten dus van elk van de relevante 'hoekjes' in de zorgvraagtyperingsruimte enkele voorbeelden gezien hebben. Die verzadiging kunnen we meten door te bekijken hoe groot het percentage unieke vragenlijsten is op het totaal. Bij een laag percentage unieke vragenlijsten is sprake van verzadiging, terwijl bij een hoog percentage meer vragenlijsten nodig zijn om de ruimte aan zorgvragen goed in te kunnen schatten.

In de pilot met het zorgclustermodel hebben we bijna 20.000 vragenlijsten en bijbehorende zorginzet verzameld. Het blijkt dat 79% van die vragenlijsten een unieke combinatie van antwoorden bevat. Vervolgens blijkt dat de gemiddelde kosten van een zorgtraject met een unieke vragenlijst ongeveer vier keer zo hoog zijn als de kosten voor een traject met een niet-unieke vragenlijst. Dit duidt erop dat complexe zorgvragen (hoge kosten) vaak ook de niet veel voorkomende zorgvragen zijn. Binnen de pilot lijken vragenlijsten dus niet veel op elkaar. Daaruit volgt dat we met de huidige 20.000 vragenlijsten nog niet voldoende zicht hebben op de 'ruimte' van scores waarin zorgvragen zich bevinden en dus ook geen goed zicht hebben op waar de complexe zorgvragen zich bevinden. Als alternatief hebben we met simulaties, gebaseerd op de vragenlijsten van de pilots, geprobeerd met het aanpassen van de scores op



één of twee vragen (8 miljoen totaal) de relevante ruimte verder te verkennen. Het blijkt echter dat dat nog steeds 75% unieke vragenlijsten oplevert en de ruimte dus ook met simulaties niet snel verzadigd is, vermoedelijk omdat we nog niet goed weten welk deel van de ruimte relevant is. Complexe patiënten blijken soms op onverwachte plaatsen voor te komen, zoals onder behandeling bij een vrijgevestigde psycholoog of in monodisciplinaire setting. Daarom hebben we gedurende de eerste fase van het onderzoek gegevens van veel patiënten nodig. We weten nog te weinig en zullen dus moeten verkennen in welke vorm de zorgvragen bestaan en welke (combinatie van) factoren wijst op complexe zorg. Daarvoor is een combinatie van vragenlijst en behandelinzet vereist.

Subdoelen 4 en 5, waarvoor een subset van de vragenlijsten volstaan, werken we verder uit in hoofdstuk 3 onder fase 2 en 3. In dat hoofdstuk geven we ook aan welke steekproeven nodig zijn en hoe we die bepalen.

| Doel                           | Subdoel                                | Benodigde gegevens per zorgvraag  |
|--------------------------------|----------------------------------------|-----------------------------------|
| Algoritme                      | 1. Dynamisch algoritme                 | Zorgvraagtypering                 |
| Algoritme                      | 2. Vervangen Brits algoritme           | Zorgvraagtypering                 |
| Algoritme                      | 3. Inschatten 'ruimte' zorgvragen      | Zorgvraagtypering + behandelinzet |
| Algoritme                      | 4. Geijkt houden algoritmes            | Zorgvraagtypering + behandelinzet |
| Algoritme/IJken zorgvraagtypes | 5. Onderzoek/verbeteren bijzonderheden | Zorgvraagtypering + behandelinzet |
| IJken zorgvraagtype            |                                        | Zorgvraagtypering + behandelinzet |

## 2.3 Verdeling zorgvraagtyperingen ijken

Om te dienen voor de zorginkoop moet de indeling in zorgvraagtypes voldoende ruimte bieden aan alle patiëntengroepen in de Nederlandse ggz. Tegelijkertijd moeten de zorgvraagtypes voldoende specifiek zijn dat de behandelprofielen van groepen patiënten binnen een bepaald zorgvraagtype voldoende op elkaar lijken. Daarnaast moet er een hanteerbaar aantal zorgvraagtypes bestaan. Er moeten immers in de zorginkoop passende afspraken gemaakt kunnen worden over een reële vergoeding bij bepaalde groepen patiënten. Daarom is het belangrijk dat de indeling in zorgvraagtypes voldoende recht doet aan de diversiteit aan zorgvragen en behandelinzetten. In de huidige vorm is dat nog niet het geval.

Vanuit de gegevens die we met de pilots hebben verzameld weten we dat de groepen zorgvraagtypes meer homogeen zijn dan de diagnosehoofdgroepen die we tot nu toe hadden. Maar we weten ook dat er in de huidige vorm veel spreiding is van zorgkosten binnen de zorgvraagtypes (eindevaluatie pilot), dit betekent mogelijk dat de 'grenzen' tussen zorgvraagtypes niet overal op de goede plaats liggen. Daarnaast ontvangen we veel signalen van ggz-behandelaren dat niet alle zorgvragen goed te vangen zijn in de zorgvraagtypes, dit kan er op duiden dat er zorgvraagtypes moeten worden toegevoegd. Tot slot zou zorgvraagtypering er vanuit het oogpunt van zorginkoop voor moeten zorgen dat binnen een zorgvraagtype een min of meer homogene groep patiënten voorkomt, maar het lijkt er op dat de meest complexe zorgvragen vermengd met minder complexe gevallen in bepaalde zorgvraagtypes vallen (onderzoek SiRM). Dat kan betekenen dat bepaalde zorgvraagtypes gesplitst moeten worden.

Zoals we ook bij het herijken van het algoritme noemden, weten we nog te weinig van de mogelijke zorgvragen in de Nederlandse ggz om een redelijke steekproef te bepalen. Er is een gigantische hoeveelheid mogelijke combinaties van scores mogelijk (bijna 20.000.000.000.000, ofwel meer dan één miljoen unieke mogelijke vragenlijsten voor elke Nederlander, zie ook sectie 3.1 onder ad 1.) en we weten nog niet wat de relevante ingevulde vragenlijsten onder die mogelijkheden zijn (zie ook sectie 2.2.2 onder subdoel 3). Daarom zullen we in eerste instantie alle zorgvragen nodig hebben, ook al verwachten we dat



we daarmee de mogelijkheden nog lang niet volledig gedekt hebben. We zullen vragenlijsten gebruiken om de grenzen tussen zorgvraagtypes vast te stellen en behandelinzet om te evalueren of de voorgestelde grens zal leiden tot een meer homogene verdeling van kosten binnen de zorgvraagtypes. Met deze combinatie is snel een optimale verdeling van zorgvragen en behandelinzet over zorgvraagtypes tot stand te brengen. We verwachten dat we na de eerste ijking van de zorgvraagtypes toe kunnen met een fractie van de zorgvragen en behandelinzet om te monitoren of de grenzen nog op de juiste plaats liggen.

## 3 Tijdspad

Het verder ontwikkelen van zorgvraagtypering kan vermoedelijk niet in één stap gebeuren. Het is waarschijnlijk dat na een initiële grote uitvraag, we snel met minder gegevens uit de voeten kunnen. Hieronder beschrijven we de keuzes en onderliggende overwegingen om zo snel mogelijk terug te schalen en hypothesegericht uitvragen te kunnen doen.

### 3.1 Fase 1: 2022-2023

Op dit moment hebben we nog te weinig kennis van de zorgvragen die voor kunnen komen en de behandelinzet die daarbij waarschijnlijk of gepast is. De zorgvragen in de ggz zijn zeer divers, we weten niet goed hoe we patiënten met een complexe zorgvraag kunnen herkennen en hebben nog geen zicht op het volledige spectrum waarop zorgvragen voor kunnen komen zoals gemeten op de HoNOS+. Daarom is het nog onmogelijk te weten hoe de beslisboom van het dynamische algoritme moet lopen, hoe we het 'Britse' algoritme kunnen aanpassen aan de Nederlandse zorg en waar de grenzen tussen de zorgvraagtypes goed kunnen liggen.

Voor het beter schatten van het algoritme zullen we een zogeheten *classifier*, een statistisch model dat gegevens indeelt in groepen schatten op de opgehaalde vragenlijsten met bijbehorende zorgvraagtypes. Na eigen onderzoek lijkt een *random probability forest* in eerste instantie de meest geschikte classifier (bijlage 1). Het veld van classificatie is nog sterk in ontwikkeling. Er is nog geen consensus over de kwantificatie van de steekproefgrootte die een acceptabel model zal opleveren. Uit de literatuur weten we dat voor het schatten van een classifier veel gegevens nodig zijn en dat de precieze hoeveelheid gegevens van veel factoren afhangt. De belangrijkste hiervan zijn, namelijk 1. het aantal parameters van het model/feature-'ruimte', 2. de correlatie tussen de parameters, 3. het aantal uitkomstklassen, 4. de verschillendheid tussen de klassen en 5. de gelijkheid binnen de klassen. Al deze factoren dragen eraan bij dat we voor de zorgvraagtypering in eerste instantie een zeer grote steekproef nodig hebben. Hieronder de uitwerking van de elementen die bijdragen aan de benodigde grootte van een steekproef.

- ad 1. We hebben 95 parameters (19 items keer 5 mogelijke scores) en een enorme 'ruimte' waarin de vragenlijsten voor kunnen komen  $5^{19}$  (~2 met 13 nullen erachter), waarvan we nog niet goed in kunnen schatten welke gebieden in die ruimte relevant zijn. Wanneer we goed zouden weten welke HoNOS+-items het meest relevant zijn kunnen we snel toe met een kleinere steekproef. Tot die tijd hebben we een grote steekproef nodig om de classifier te trainen.

*Veel parameters betekent veel combinaties van antwoorden op de vragenlijst en een groot aantal mogelijkheden om te vergelijken.*

- ad 2. Uit de wetenschappelijke literatuur weten we dat verschillende aspecten van ggz-zorgvragen zeer sterk gecorreleerd kunnen zijn. Dit geldt ook voor de HoNOS+, vooral voor correlaties tussen de historische en bijbehorende actuele items. Dit is zeer nadelig voor de

informatieve waarde per item en geeft daarmee een grote benodigde steekproef om de classifiër te trainen.

*Veel correlaties tussen variabelen betekent dat variabelen (deels) dezelfde informatie bevatten, dat betekent dat de afzonderlijke variabelen gemiddeld minder 'unieke' informatie bevatten en er dus meer gegevens nodig zijn om voldoende informatie te verkrijgen.*

- ad 3. Voor hoofdgroep X (1 t/m 8) en Y (10 t/m 17) zijn er acht mogelijke zorgvraagtypes en voor hoofdgroep Z zijn er vier zorgvraagtypes (18 t/m 21). Wij zullen de classifiërs daarom schatten met vier of acht mogelijke uitkomsten. Met het aantal uitkomsten neemt de complexiteit van de schatting toe. Het model moet dan namelijk meer afwegingen maken. Aan elke mogelijke afweging moeten voldoende gegevens ten grondslag liggen. De steekproef moet daarom groter zijn dan wanneer bijvoorbeeld slechts twee uitkomsten met elkaar vergeleken hoeven worden.

*Meer dan twee uitkomstklassen (zorgvraagtypes) betekent dat het model meerdere afwegingen moet maken, niet alleen "1 of 2" maar bijvoorbeeld "1 of 2; 1 of 3; 2 of 3" tegelijk.*

- ad 4. De zorgvraagtypes lijken in sommige gevallen zeer vergelijkbaar. Bijvoorbeeld zorgvraagtypes 1 en 2 (respectievelijk *Psychische aandoening - lichte problematiek* en *Psychische aandoening - lichte problematiek met grotere zorgvraag*) zijn in de pilotgegevens vergelijkbaar en ook het 'Britse' algoritme heeft grote moeite het verschil tussen de twee te identificeren. Dat maakt het moeilijk voor de classifiër om de juiste klasse te kiezen. Het model doet er dan ook langer over om te leren wat het onderscheid tussen de zorgvraagtypes is. Ook daarom is een relatief grote steekproef nodig om het model te trainen.

*Als zorgvraagtypes weinig van elkaar verschillen is meer informatie nodig om het verschil te bepalen.*

- ad 5. Er is grote diversiteit binnen de zorgvraagtypes. Dat is deels de aard van de sector, de zorgvragen zijn immers zeer divers, maar kan ook het gevolg zijn van de huidige indeling die nog niet op de Nederlandse situatie is geënt. Als de verschillen binnen een zorgvraagtype groot zijn is het voor een classifiër moeilijk de grenzen te leren en is een grote steekproef nodig om te trainen.

*Als verschillende zorgvragen binnen een zorgvraagtype veel van elkaar verschillen is het moeilijk de grenzen van het zorgvraagtype te bepalen en is veel informatie nodig.*

Uit bovenstaande is op te maken dat ten behoeve van de doorontwikkeling minimaal alle vragenlijsten nodig zijn in 2022-2023. Zonder de volledige set aan gegevens is het niet mogelijk de patiënten in de juiste zorgvraagtypes in te delen. Het is cruciaal dat hier zekerheid over bestaat. Zolang dit niet het geval is zullen zorgaanbieders grote risico's lopen met de behandeling van verkeerd ingedeelde patiënten die daardoor geen gepaste vergoeding oplevert. Als gevolg daarvan zullen ze zich niet kunnen permitteren op grote schaal zorg te gaan leveren aan complexe patiënten. Afspraken en afrekeningen tussen zorgverzekeraar en zorgaanbieder vinden doorgaans in een jaarlijkse cyclus plaats. We stellen ons daarom tot doel dat alle zorgvragen die in ieder geval éénmaal per jaar voorkomen goed moeten kunnen worden ingedeeld. Als we alle zorgvragen uit een jaar hebben gezien, dan hebben we dus op alle relevante zorgvragen verzadiging (zie subdoel 3). Om hier een minimale mate van zekerheid over te hebben vragen we gedurende twee jaren alle zorgvragen uit. Door twee jaar volledig uit te vragen kunnen we de zorgvragen vergelijken en vaststellen hoeveel nieuwe zorgvragen er nog bij komen. Alleen de uitzonderingsgevallen (< 1 keer per jaar binnen de Nederlandse ggz) vormen dan nog een risico. Deze



risico's zullen moeten worden afgedicht in afspraken tussen zorgverzekeraar en zorgaanbieder, net als de praktijkvariatie die binnen de zorgvraagtypes kan bestaan.

Belangrijk doel van de uitvraag is dus het goed leren kennen welke combinaties aan scores gelijk staan aan complexe zorgvragen. Het zou dan een overweging kunnen zijn om alleen de complexe zorgvragen uit te vragen. Het is echter niet goed te voorspellen waar de complexe zorgvragen te vinden zullen zijn. Naar schatting is er een jaarlijkse instroom van 41.000 patiënten met ernstige psychiatrische aandoeningen (EPA-patiënten) in de ggz (Feiten en cijfers over mensen met een ernstige psychiatrische aandoening | Vektis.nl ). Van de EPA-patiënten blijkt echter maar een klein deel (12-14%) behandeling met klinische opname te ontvangen (vektis) en ze komen ook voor bij vrijgevestigde aanbieders (eigen onderzoek). Het selecteren van een behandelsetting zal dus niet voldoende zijn. We weten ook dat EPA-patiënten niet allemaal tot de duurste patiënten behoren. Dat kan zijn omdat ze op de verkeerde plaats behandeld worden als gevolg van de lange wachtlijsten, omdat ze bekend staan om zorgmijddend gedrag, of omdat een groot deel van de zorgkosten niet gemaakt worden in het ggz-domein. We zien dus geen aanknopingspunten om een kleinere selectie van patiënten te nemen en toch ook de zorgvragen binnen te krijgen die in ieder geval jaarlijks voorkomen.

## Subsidiariteit

In de praktijk kunnen er aanknopingspunten zijn om de steekproef kleiner te maken in het kader van de subsidiariteit. Bijvoorbeeld als slechts bepaalde vragen voorspellende waarde blijken te hebben of als de zorgvraagtypes homogeen en onderscheidend zijn. Dit is op dit moment nog niet het geval. Wel is de verwachting dat bepaalde HoNOS+-items meer invloed hebben dan andere. Zodra duidelijk is wat de meest relevante HoNOS+-scores zijn in relatie tot de behandelinzet zullen we de uitvraag beperken. Tot die tijd moeten we er echter van uitgaan dat elk van de scores evenveel voorspellende waarde heeft. Dat betekent dat we een volledig beeld moeten hebben van alle waarschijnlijk voorkomende vragenlijsten. Of we dit verzadigingsniveau bereikt hebben kunnen we meten door te kijken naar het aantal vragenlijsten in de dataset waar er slechts één van bestaat. Als dit aantal unieke vragenlijsten lager wordt dan is aannemelijk dat we in de buurt van verzadiging komen. Uit de pilots blijkt dat dit met 20.000 vragenlijsten nog niet het geval was: 79% van de vragenlijsten in de dataset was namelijk nog uniek.

Naast deze afwegingen heeft de NZa ook middels simulatie naar oplossing gezocht. Het simuleren van een groot aantal (8 miljoen) vragenlijsten leverde echter geen verzadiging op. Dit komt doordat we de fijnere relaties tussen de verschillende items niet goed konden simuleren op basis van de 20.000 vragenlijsten uit de pilot. We hebben dus alle vragenlijsten uit de ggz-sector nodig om het model goed de relaties tussen items te leren kennen en vooral te leren welke items het meest belangrijk zijn. Daarover meer onder fase 2.

## Proportionaliteit

Zoals aangegeven willen we bereiken dat patiënten met een complexe zorgvraag eerder in behandeling genomen worden en minder lang op de wachtlijst staan. Dit doel kan nauwelijks worden overschat en wordt onderschreven door o.a. de Algemene Rekenkamer. We willen het doel bereiken door het voor aanbieders financieel aantrekkelijker te maken om patiënten met een zwaardere zorgvraag eerder te behandelen. Dat kan alleen als we de patiënten met een zwaardere zorgvraag kunnen identificeren en de zorgaanbieder en verzekeraar afspraken kunnen maken over de verdeling van de zorgkosten. Dat doen we d.m.v. zorgvraagtypering. Hiervoor is aangegeven waarom het voor een juiste bepaling van de zorgvraagtypering noodzakelijk is dat in ieder geval in 2022 en 2023 een groot aantal HoNOS+ vragenlijsten wordt verzameld en geanalyseerd, wat naar schatting neer zal komen op 1.200.000



vragenlijsten, gebaseerd op de 780.000 patiënten dat jaarlijks in de geestelijke gezondheidszorg wordt behandeld. Dit aantal kan lager uitvallen door ingevulde privacy-verklaringen, maar kan ook hoger zijn doordat de HoNOS+ op verschillende momenten voor een patiënt worden vastgelegd. Omdat we onder het aantal niet uitkomen, worden een aantal maatregelen genomen om de impact voor de betrokkenen te beperken:

- De declaratiegegevens en HoNOS+ vragenlijsten worden afzonderlijk uitgevraagd. De HoNOS+ vragenlijsten bevatten geen direct identificerende kenmerken van de patiënt waardoor deze informatiestroom zonder koppelingen met andere gegevens niet tot een individu te herleiden is.
- Een beperkt aantal personen binnen de NZa heeft toegang tot de HoNOS+ vragenlijsten. Een deel van de informatie gegevens is niet zichtbaar voor personen die rechten hebben om de declaraties in te zien. Toegang tot alle gegevens wordt actief gelogd.
- De HoNOS+ vragenlijsten worden niet voor een ander doel gebruikt dan voor de doeleinden zoals hiervoor beschreven, zie ook sectie 5.2 voor waarborgen in de Regeling.

In de zorginkoop moet onderscheid gemaakt kunnen worden naar complexiteit van de zorgvraag. In dat kader willen we op basis van gegevens uit 2022 en 2023 beter begrijpen welke zorgvragen er bestaan in de ggz en hoe die zorgvragen zich verhouden tot de verschillende zorgvraagtypes, patiëntengroepen en soorten behandelinzet (zie ook sectie 4.1). Met name de groep patiënten met een complexe zorgvraag willen we goed in beeld krijgen. Tegelijk zijn dat de weinig voorkomende zorgvragen en op dit moment weten we niet goed in welke zorgvraagtypes deze patiënten ingedeeld zijn en hebben we ook geen andere methode om de groep te definiëren. In 2023 zullen we op basis van gegevens leren hoe we deze groep kunnen herkennen. Vanaf dat moment kunnen we gericht gegevens uitvragen en daarmee de steekproef beperken zoals hierboven omschreven. Gezien dit doel meent de NZa dat de uitvraag proportioneel is.

## 3.2 Fase 2: 2024-2025

In 2024 verwachten we de zorgvraagtypes ingedeeld te hebben op basis van de Nederlandse situatie. Ook zullen de algoritmes dan gebouwd zijn en gevoed zijn met gegevens van de Nederlandse zorg. Het meest data-intensieve werk is hiermee gedaan en de uitvraag van gegevens kan sterk worden teruggeschroefd. Vanaf 2024 zal de NZa niet meer voor alle patiënten de gegevens uitvragen. Vanaf dat moment zal de NZa hypothesegedreven gegevens uitvragen, naast een relatief kleine monitoringsuitvraag om de stabiliteit van het systeem te waarborgen en eventuele implementatie-effecten te ondervangen. De vorm en omvang van uitvragen na 2023 zijn nog niet nauwkeurig te voorspellen en bij afwijkingen van de plannen zal een nieuwe AVG-afweging gemaakt worden.

Subdoel 4. Geijkt houden algoritmes In fase 2 verwacht de NZa goed zicht te hebben op de zorgvragen die kunnen voorkomen en daarmee op de patiënten met een complexe zorgvraag. Er zullen dan twee nieuwe algoritmes (dynamisch en volledig) geïmplementeerd zijn. Het is te verwachten dat er implementatie-effecten optreden na invoering van de nieuwe algoritmes. Mensen leren immers hoe het nieuwe systeem werkt en wetenschappelijk onderzoek naar gedrag laat zien dat mensen vaak kiezen voor de gemakkelijkste route. Het is niet uit te sluiten dat de gemaksprikkels verschillen tussen de twee algoritmes. Voor de zorginkoop is het belangrijk dat zorgtrajecten met een bepaald zorgvraagtype onderling vergelijkbaar zijn, ongeacht de variant van zorgvraagtypering die is gebruikt. Daarom is het belangrijk te monitoren dat de twee algoritmes dezelfde resultaten opleveren in de vorm van de voorspelde en gekozen zorgvraagtypes en behandelinzet. Daarom is in fase 2 een steekproef nodig van vragenlijsten en bijbehorende behandelinzet. We kunnen naar verwachting volstaan met een steekproef van 1% van de

dynamische zorgvraagtyperingen met bijbehorende volledige vragenlijst en daarbij ook een even grote groep volledige zorgvraagtyperingen. Uit de gegevens uit 2022-2023 moet duidelijk worden hoe vaak de dynamische zorgvraagtypering gebruikt wordt. Als blijkt dat de dynamische zorgvraagtypering niet of nauwelijks wordt gebruikt vervalt dit subdoel. We zullen dan de mogelijkheid voor dynamische zorgvraagtypering afschaffen.

Subdoel 5. Onderzoek/verbeteren bijzonderheden Naar verwachting ligt er na het ijkken van het algoritme en het vastleggen van de zorgvraagtypes op de Nederlandse situatie in fase 1 een goed systeem van zorgvraagtypering. Het valt echter niet uit te sluiten dat het systeem op bepaalde punten verder onderzoek of verbeteringen nodig heeft. Als dat het geval is kan het zijn dat we aanvullende gegevens op moeten vragen. Het opvragen van gegevens kan dan hypothesegestuurd gebeuren. We zien twee waarschijnlijke routes waaruit aanvullende vragen voort kunnen komen. Allereerst is daar de set met declaratiegegevens waar de NZa over beschikt als onderdeel van haar wettelijke taken. Uit analyse van deze gegevens kunnen bijzonderheden naar voren komen, zoals een zeer grote variatie van behandelinzet binnen een zorgvraagtype of het veel voorkomen van zorgvraagtypes in een behandelsetting waar ze onwaarschijnlijk zouden zijn. Deze bijzonderheden zullen eerst getoetst worden bij de adviescommissie zorgvraagtypering op hun validiteit en mogelijke verklaringen uit de praktijk. Als uit deze consultatie een goede verklaring naar voren komt kan het mogelijk zijn dat we van de groep zorgvraagtypes die het betreft aanvullende gegevens zullen opvragen. De tweede route waarlangs we bijzonderheden verwachten te constateren is via kwalitatief onderzoek en gesprekken met en signalen uit de zorgpraktijk. Uit deze hoek zijn signalen te verwachten zoals het niet goed passen van patiëntengroepen in de zorgvraagtypes. Als de signalen voldoende sterk en breed gedragen blijken na toetsing in de adviescommissie kunnen we een specifieke aanvullende uitvraag opzetten. De grootte van de uitvraag zal afhangen van de geconstateerde bijzonderheid en zal altijd subsidiariteit als uitgangspunt hebben.

### 3.3 Fase 3: na 2025

In fase 3 hoeven we slechts een zeer beperkte uitvraag te doen. We kunnen er immers van uit gaan dat de grote implementatie-effecten zijn neergedaald en zorgvraagtypering relatief stabiel is. Ook is het niet te verwachten dat er plotseling grote veranderingen optreden van voorkomende zorgvragen of dat de best passende behandelinzet bij een zorgvraagtype verandert. We gaan er immers van uit dat er geen grote veranderingen van inzicht zullen zijn over wat goede zorg is. Daarom hoeven we geen gegevens uit te vragen om sterke fluctuaties in kaart te brengen. We kunnen onze informatiebehoefte grotendeels baseren op hypothesen. De hypothesegedreven informatiebehoefte is dermate onvoorspelbaar dat we hem in dit stuk niet verder zullen uitwerken, behalve de toezegging dat de overwegingen en keuzes goed zullen worden vastgelegd en proportionaliteit en subsidiariteit ook hierbij het uitgangspunt zullen zijn. De noodzaak voor een uitvraag wordt door de NZa in overleg met de veldpartijen jaarlijks bepaald, waarbij deze uitvraag als een nieuwe uitvraag zal worden gezien en separaat getoetst op de AVG.

Uitzondering op de hypothesegestuurde informatiebehoefte is het monitoren van de stabiliteit. Zoals ook in fase 2 besproken is de stabiliteit van en vergelijkbaarheid tussen zorgvraagtypes belangrijk voor de zorginkoop. Daarom moeten we ook op de langere termijn monitoren of de twee algoritmes op elkaar geijkt blijven en of de grenzen tussen zorgvraagtypes nog op de juiste plaats lopen. Hiervoor hebben we een doorlopende steekproef nodig, gebaseerd op de zorgvragen en de behandelinzet. Het is mogelijk een steekproefgrootte te berekenen op basis van gegevens verkregen in fases 1 en 2 of de declaraties. De keuze van de verkregen gegevens waarop we de steekproef baseren is afhankelijk van het type vraag. We gaan ervan uit dat de variatie van de zorginzet het gevolg is van vergelijkbare variatie in de zorgvraag. In dat geval kunnen we een relatief eenvoudige waarde (kosten van de zorginzet) gebruiken om de



steekproef voor beiden te bepalen. We kunnen het gemiddelde en de standaarddeviatie van de kosten van de zorginzet nemen en invoeren in onderstaande formule.

$$n = \left( \frac{Z * \sigma}{E} \right)^2$$

In deze formule is Z afkomstig uit de (standaard) Z-tabel en is 1.96 voor het gewenste betrouwbaarheidsinterval van 95%. De  $\sigma$  is de standaardafwijking van de zorgkosten binnen elk zorgvraagtype. E is de foutmarge die we acceptabel vinden voor de steekproef. We willen met het oog op de rol in de zorginkoop dat de schatting van de gemiddelde zorgkosten binnen 10% van de gemiddelde zorgkosten van elk zorgvraagtype valt. De precieze grootte van de steekproef kunnen we pas bepalen als we gegevens uit fase 1 hebben. Als voorbeeld en zeer ruwe inschatting kunnen we alvast de (onbetrouwbare) pilotgegevens gebruiken in deze berekening. In dat geval komen de schattingen uit tussen 250 en 2700 vragenlijsten met behandelinzet per zorgvraagtype. De steekproef wordt snel kleiner als de standaardafwijking kleiner wordt ten opzichte van de gemiddelde kosten.

In samenvatting kunnen we de fases en benodigde gegevens beschrijven als hieronder.

| Fase               | Doel                           | Subdoel                                    | Benodigde gegevens                | Aantal vragenlijsten                                                 |
|--------------------|--------------------------------|--------------------------------------------|-----------------------------------|----------------------------------------------------------------------|
| 1. (2022-2023)     | Algoritme                      | 1. Dynamisch algoritme                     | Zorgvraagtypering                 | Sectorbreed                                                          |
| 1. (2022-2023)     | Algoritme                      | 2. Vervangen Brits algoritme               | Zorgvraagtypering                 | Sectorbreed                                                          |
| 1. (2022-2023)     | Algoritme                      | 3. Inschatten 'ruimte' complexe zorgvragen | Zorgvraagtypering + behandelinzet | Sectorbreed                                                          |
| 1. (2022-2023)     | IJken zorgvraagtypes           | IJken zorgvraagtypes                       | Zorgvraagtypering + behandelinzet | Sectorbreed                                                          |
| 2. (2024-2025)     | Algoritme                      | 4. Geijkt houden algoritme                 | Zorgvraagtypering + behandelinzet | Steekproef (1%)                                                      |
| 2. (2024-2025)     | Algoritme/IJken zorgvraagtypes | 5. Onderzoek bijzonderheden                | Zorgvraagtypering + behandelinzet | Opvallende zorgvraagtypes                                            |
| 3. (2026 en later) | Algoritme                      | 4. Geijkt houden zorgvraagtypes            | Zorgvraagtypering + behandelinzet | Monitorings-steekproef (grootte bepaald volgens formule in de tekst) |
| 3. (2026 en later) | Algoritme/IJken zorgvraagtypes | 5. Onderzoek bijzonderheden                | Zorgvraagtypering + behandelinzet | Opvallende zorgvraagtypes                                            |

## 4 Alternatief onderzoek

Uit het voorgaande blijkt dat een volledige uitvraag noodzakelijk, maar ook proportioneel en subsidiair is. In dit hoofdstuk komt de subsidiariteitsvraag nogmaals aan bod. In aanloop naar de invoering van het Zorgprestatiemodel zijn diverse alternatieven op een volledige uitvraag overwogen. De AP geeft aan dat de verwerping van deze overwogen alternatieven onvoldoende is gemotiveerd. Hieronder zal uitdrukkelijker worden ingegaan op de motivering. De rode draad is dat niet duidelijk is welke zorgvragen in de ggz-praktijk voorkomen en wat de relatie is tussen de zorgvraag, het zorgvraagtype en de behandelinzet. Zorgvraagtypering is een nieuw systeem en er bestaat dus nog weinig kennis over de praktische knelpunten en toepassingen. Daarnaast is het aantal mogelijke zorgvragen zeer groot. Het is daarbij onmogelijk vooraf in te schatten welke van deze zorgvragen daadwerkelijk in de praktijk voor zullen komen en welke het meest zorginzet vragen. Hieronder lopen we de verschillende alternatieven op een volledige uitvraag na en geven we waar mogelijk aan op welk moment het alternatief een reële optie vormt.

## 4.1 Alleen registraties uit 2022 opvragen

De NZa heeft allereerst overwogen de uitvraag te beperken tot alleen zorgvraagtyperingen uit 2022. Om een goed werkend algoritme te ontwikkelen is het echter belangrijk dat de onderliggende gegevens stabiel zijn. Op basis van deze gegevens wordt de ontwikkeling van zorgvraagtypering over tijd zichtbaar. Daarmee kunnen we per zorgvraagtype vaststellen wanneer zorgvraagtypering stabiel is. Dit kan enkel op basis van de volledige gegevensset, juist als deze nog (deels) instabiel is. Het is belangrijk dat zorgverzekeraars, ggz-aanbieders en –behandelaren de objectieve werkelijkheid van de algoritmes en indeling in zorgvraagtypes accepteren. Anders zal zorgvraagtypering immers niet gebruikt worden om de benodigde verschuiving van zorg richting complexe patiënten voor elkaar te krijgen. In 2022 is zorgvraagtypering een nieuw ingevoerd systeem. Dit betekent dat de behandelaars nog moeten wennen aan het invullen van de HoNOS+-vragenlijst en nog hun weg moeten leren kennen tussen de zorgvraagtypes. Het is daarom aan te nemen dat zorgvraagtypering nog niet erg stabiel is in het eerste jaar. Daarnaast is breed gebruik gemaakt van de mogelijkheid om patiënten pas vanaf eind 2022 of begin 2023 een zorgvraagtype toe te kennen en de verplichting om de zorgvraagtyperingsgegevens te registreren is pas in juli 2022 ingegaan. Ook zorgverzekeraars en zorgaanbieders hebben aangegeven dat het voor hen noodzakelijk is dat de algoritmes worden gebouwd op gegevens uit een stabiele situatie. De inschatting is dat die stabiele situatie eind 2023 een feit is. Om de stabiliteit goed in te kunnen schatten is het nodig dat we de gegevens over tijd met elkaar kunnen vergelijken en dus moeten we voor een (iets) langere periode gegevens ontvangen. We zullen de gegevens die eind 2022, begin 2023 en eind 2023 aangeleverd worden met elkaar vergelijken om zo in te schatten op welk moment de gegevens stabiel genoeg zijn om de algoritmes te schatten en zorgvraagtypes te iken op de Nederlandse situatie. Na 2023 zullen we toe kunnen met een veel kleinere steekproef, zoals in sectie 3.2 en 3.3 hierboven beredeneerd.

## 4.2 Kwalitatief onderzoek

Het tweede alternatief dat de NZa heeft overwogen is kwalitatief onderzoek. Hoewel kwalitatief onderzoek belangrijk is, kan het een kwantitatief onderzoek niet vervangen. Kwalitatief onderzoek in de vorm van het verzamelen van signalen, interviews en gesprekken met behandelaren en vragen tijdens *webinars* en informatiesessies zal moeten bijdragen aan het vinden van knelpunten met de zorgvraagtypering en ontbrekende of overlappende zorgvraagtypes. Daarom heeft de NZa bij de adviescommissie die adviseert over zorgvraagtypering ook gevraagd om een werkgroep die adviseert over de praktijk. Deze werkgroep heeft als deelopdracht:

*Doe (veld)onderzoek naar de toepassing van zorgvraagtypering in de praktijk. Inventariseer knelpunten bij de toepassing van zorgvraagtypering en benoem hoe deze knelpunten verholpen kunnen worden. Vraag in dit veldonderzoek ook naar de bruikbaarheid van hertypering en doe eventueel voorstellen ter verbetering.*

De uitkomsten van het genoemde onderzoek zullen dienen als bron om de uitvraag voor subdoel 5 (onderzoek bijzonderheden) in fase 2 richting te geven.

Kwalitatief onderzoek is niet voldoende om een betrouwbare *classifier* te schatten (hoofddoel 1) of om betrouwbaar de grenzen tussen de zorgvraagtypes te bepalen (hoofddoel 2). De algoritmes en indeling in zorgvraagtypes zijn geheel kwantitatief van aard. Met de kwalitatieve informatie kunnen slechts knelpunten in de toepassing van het algoritme worden opgespoord. Dit terwijl een fundamentele ijking noodzakelijk is om het systeem in de praktijk te laten functioneren en daarmee te laten bijdragen aan het oplossen van de grote problemen in de Nederlandse ggz. Deze ijking is enkel mogelijk op basis van een volledige uitvraag. Wanneer de algoritmes en indeling compleet zijn (fase 2 en 3) is kwalitatief onderzoek een belangrijke bron om verdere ontwikkeling van zorgvraagtypering vorm te geven.



## 4.3 Beginnen bij declaratiegegevens

De NZa heeft ten slotte nog overwogen om het systeem van zorgvraagtypering verder te ontwikkelen op basis van declaratiegegevens. We zouden in die gegevensbron bijzonderheden kunnen constateren die aanleiding geven tot een nader onderzoek. Dat zou een reële optie zijn op het moment dat het systeem al volledig ontwikkeld is. We weten echter dat de algoritmes nog niet goed genoeg zijn (sectie 2.2) en dat de indeling van patiënten in zorgvraagtypes nog geen recht doet aan de werkelijkheid (sectie 2.3). Problemen met de algoritmes zijn niet te herkennen in de declaratiegegevens, omdat het gekozen zorgvraagtype het enige zorgvraagtyperingselement is dat in die gegevens aanwezig is. Om goed te kunnen dienen als referentiepunt in de zorginkoop en adviesinstrument voor de behandelaar moeten de algoritmes gelden voor alle patiënten. Het is voor het schatten van het algoritme (hoofddoel 1) dus niet voldoende om gegevens van enkele opvallende zorgvraagtypes uit te vragen omdat de algoritmes dan niet breed toepasbaar zijn.

Wanneer de zorgvraagtyperingen goed zijn gedefinieerd in fase 1 kunnen de declaratiegegevens wel gebruikt worden voor het aanwijzen van zorgvraagtypes waar verdere verfijning nodig is. We zullen dat dan ook doen onder subdoel 5 (onderzoeken bijzonderheden) in fase 2.

# 5 Waarborgen in de Regeling

## 5.1 Horizonbepaling

De AP vraagt verder aandacht voor duidelijkheid op het punt van de looptijd van de verplichte aanlevering (een horizonbepaling). In art. 4.2 lid 4 van de Regeling staat dat de aanbieders ieder half jaar verplicht de HoNOS+ gegevens bij de NZa aanleveren. Niet is aangegeven wanneer deze verplichting vervalt. Gelet op het voorgaande hebben we na afloop van 2023 een beter inzicht en zal daarna worden volstaan met het uitvragen van een steekproef. De tekst in de Regeling zal daarom als volgt worden aangepast.

### 2023

In de toelichting bij de Regeling geestelijke gezondheidszorg en forensische zorg voor 2023 zal de NZa in de sectie 'Artikelsgewijs' na de toelichting op de informatie-elementen 'd1 en d2. DSM-hoofdgroep en DSM-5 classificatie' de volgende toelichting toevoegen. Deze tekst staat in de toelichting, omdat de Regeling bindend is voor zorgaanbieders en niet voor de NZa. De Regeling wordt inclusief toelichting gepubliceerd via de Staatscourant, zodat publiekelijk helder is hoe de NZa zal omgaan met de informatie-uitvraag.

#### d7. Antwoorden op HoNOS+ vragen

De NZa gebruikt de antwoorden op HoNOS+ vragen alleen voor het ijken van het algoritme zorgvraagtypering en de verdeling van zorgvraagtyperingen. De NZa verwijdt deze gegevens twee jaar nadat de gegevens compleet bij de NZa zijn aangeleverd of aangeleverd hadden moeten zijn. Het nieuwe systeem van zorgvraagtypering moet de functies van de DSM-diagnosehoofdgroep gaan overnemen na een transitiefase van twee jaar. Na deze transitiefase, dus vanaf 2024, vervalt de verplichting om deze gegevens halfjaarlijks uit eigen beweging aan te leveren. Zorgaanbieders hoeven deze gegevens vanaf 2024 alleen op uitvraag van de NZa aan te leveren. Vanaf dan zal de doorontwikkeling van zorgvraagtypering vanuit wetenschap en praktijk centraal staan. De NZa zal de uitvraag daarom beperken tot de informatiebehoefte van de Adviescommissies zorgvraagtypering ggz en fz.

## 2024

De NZa zal in de Regeling geestelijke gezondheidszorg en forensische zorg voor 2024 in artikel 4.2, vierde lid, de volgende tekst opnemen:

De zorgaanbieder levert halfjaarlijks, conform de Gegevensaanleverstandaard en onderstaande tabel, een overzicht van de volgende informatie elementen: (...)

Vervangen door:

De zorgaanbieder levert op verzoek van de NZa, conform de Gegevensaanleverstandaard en onderstaande tabel, een overzicht van de volgende informatie elementen: (...)

## 5.2 Doelbinding

Zoals hierboven aangegeven onder 5.1 zal de NZa in de toelichting bij de regeling opnemen dat zij de HoNOS+ vragenlijsten alleen zal gebruiken voor de doeleinden zoals beschreven in dit document. Voor die doeleinden zal het nooit nodig zijn te koppelen met andere gegevensbronnen. Zoals in de eerdere documenten uit de adviesvraag aan de AP is aangegeven is het verder technisch bijna onmogelijk gemaakt de gegevens te herleiden tot het individu en heeft slechts een zeer beperkte set mensen toegang tot de gegevens op een need-to-know basis.

Daarnaast voert de NZa gegevensbeschermingseffectbeoordelingen (GEB's) uit op het niveau van haar wettelijke taken. Voor onderliggende verwerkingen wordt een rechtmatigheidsbeoordeling uitgevoerd. De GEB wordt voorgelegd aan de FG en CISO van de NZa voor advies. De directeur van de verantwoordelijke directie ondertekent de GEB voor akkoord. Ook de rechtmatigheidsbeoordeling wordt voorgelegd aan de FG voor advies. In de rechtmatigheidsbeoordeling wordt aangegeven tot welke gegevens analisten van de NZa toegang hebben, zodat daarop kan worden gestuurd en de beoordeling van bijgehouden gegevens van dataverwerking kan worden uitgevoerd.

## 5.3 Bewaartermijn

De AP signaleert dat in de Regeling geen bewaartermijnen worden genoemd, maar dat de NZa in de eerder gegeven toelichting heeft aangegeven dat het bewaren van de gegevens na twee jaar niet langer noodzakelijk is.

De NZa zal daarom de in de toelichting bij de Regeling geestelijke gezondheidszorg en forensische zorg specificeren dat zij alle ontvangen gegevens in de uitvraag zorgvraagtypering zal verwijderen twee jaar nadat alle behandelinzet bij een zorgvraagtypering volledig is ontvangen of de volledige twaalf maanden ontvangen hadden moeten zijn. Het verwijderen zal gebeuren elke keer dat nieuwe gegevens worden ingelezen in de database. Zie paragraaf 5.1 voor de toelichting die de NZa zal toevoegen.

# 6 Informatie aan betrokkenen

Het is belangrijk dat patiënten van hun privacy-rechten gebruik kunnen maken. Voordat zij dat kunnen doen moet de patiënt de juiste informatie hebben over het delen van gegevens over zorgvraagtypering. Daarnaast moet de patiënt zich realiseren dat de mogelijkheid bestaat tot het ondertekenen van een privacyverklaring zoals genoemd in artikel 4.3 van de Regeling. Voor dit doeleinde werkt de NZa samen



met MIND, de organisatie die patiënten uit de ggz en hun naasten vertegenwoordigt. Uit deze samenwerking zal in ieder geval een folder met voorlichting over de informatieverplichting zorgvraagtypering volgen. In die folder zullen we ook wijzen op de mogelijkheid tot het ondertekenen van een privacyverklaring

## Toelichting grondslagen

In dit document kunt u secties vinden die onleesbaar zijn gemaakt. Deze informatie is achterwege gelaten op basis van de Wet open overheid (Woo). De letter die hierbij is vermeld correspondeert met de bijbehorende grondslag in onderstaand overzicht.

### **J** Art. 5.1 lid 2 sub e

Het belang van de openbaarmaking van deze informatie weegt niet op tegen het belang van de eerbiediging van de persoonlijke levenssfeer van betrokkenen