

Review NZa-prognosemodel Wlz-kader

18 Augustus 2021

Bram Wouterse
Pieter Bakx

Colofon / Colophon

Review NZa-prognosemodel Wlz-kader

Bram Wouterse, Pieter Bakx, 18 augustus 2021

Uitgever

Marketing & Communicatie ESHPM

Vormgeving

PanArt.nl

Review NZa-prognosemodel Wlz-kader

18 Augustus 2021

Bram Wouterse
Pieter Bakx

Inhoudsopgave

1	Inleiding	5
1.1	Aanleiding en onderzoeksvragen	5
1.2	Aanpak	5
1.3	Leeswijzer	5
2	Situatieschets	6
2.1	De organisatie van de Wlz	6
2.2	Het Wlz-kader	6
2.3	De budgetcyclus	7
2.4	De adviescyclus	7
3	Het Wlz-kader en het advies als beleidsinstrumenten	9
3.1	De rol van het Wlz-kader	9
3.2	De rol van het advies	12
4	De prognoses	14
4.1	Wat is het doel van de basisprognose?	14
4.2	Wat voorspelt het model?	15
4.3	Hoe werkt het basismodel?	15
4.4	Alternatieve scenario's	17
4.5	De interpretatie van de modelresultaten	19
4.6	Hoe goed voorspelt het model?	19
4.7	Zijn er alternatieven voor het huidige model?	24
4.8	Zijn er nog niet gebruikte databronnen?	25
4.9	Modeldocumentatie	27
5	Conclusies	28
	Literatuur	33

1 Inleiding

1.1 Aanleiding en onderzoeksvragen

De Wet Langdurige Zorg (Wlz) regelt de financiering en organisatie van intensieve langdurige zorg. Voor die zorg is ieder jaar een budget (het Wlz-kader) beschikbaar. De NZa (Nederlandse Zorgautoriteit) monitort de besteding van het Wlz-kader en adviseert de Minister van Volksgezondheid, Welzijn en Sport (VWS) over de toereikendheid ervan. Daartoe maakt de NZa meerdere keren gedurende het lopende kalenderjaar een prognose van de uitgaven gedurende de rest van het jaar.

De NZa heeft de Erasmus School of Health Policy & Management (ESHPM) verzocht een review uit te voeren van die prognosemethode voor het Wlz-kader. In de review staan de volgende vragen centraal:

- Hoe functioneert het predictiemodel als beleidsinstrument?
- Hoe goed voorspelt het huidige predictiemodel?
- Hoe gaat het predictiemodel om met de beschikbare data?
- Worden de huidige inputs optimaal gebruikt?
- Kan de prognose worden verbeterd door gebruik te maken van aanvullende databronnen?
- Hoe goed gaat het model om met plotselinge veranderingen (schokbestendigheid)?

Het doel van de review is om een onderzoeksagenda op te stellen waarmee de (voorspel)kwaliteit en bruikbaarheid van de prognose als beleidsinstrument verder verbeterd kunnen worden.

1.2 Aanpak

Deze review is gebaseerd op een analyse van relevante beleidsstukken van de NZa en andere partijen en het analyseren van door de NZa aangeleverde informatie over de werking en prestaties van het prognosemodel. Daarnaast hebben we een aantal gesprekken gevoerd met medewerkers van de NZa, het Ministerie van VWS, het Ministerie van Financiën, Zorgverzekeraars Nederland en Wlz-uitvoerders.

De jaren 2020 en 2021 blijven in deze review grotendeels buiten beschouwing. De Covid-19 pandemie heeft een dusdanige (tijdelijke) verandering in de normale Wlz-zorg veroorzaakt, dat de gemaakte prognoses in deze periode veel sterker gebaseerd zijn op expert opinion dan op de gebruikelijke modelmatige benadering waarover wij gevraagd zijn een review te doen.

1.3 Leeswijzer

Om de bruikbaarheid en kwaliteit van de prognose te kunnen evalueren is het belangrijk om het doel van de prognose in kaart te brengen en de context waarbinnen de prognose wordt uitgevoerd. Dit gebeurt in hoofdstuk 2 (situatieschets) en hoofdstuk 3 (analyse van het Wlz-kader en het advies als beleidsinstrument). Hoofdstuk 4 bespreekt het model, in hoofdstuk 5 geven we een aantal conclusies en formuleren we onderzoeksvragen.

2 Situatieschets

2.1 De organisatie van de Wlz

De Wlz regelt zware, intensieve zorg voor kwetsbare ouderen, mensen met een handicap en mensen met een psychische aandoening (Rijksoverheid 2021). Mensen die deze zorg nodig hebben kunnen een indicatie voor Wlz-zorg aanvragen bij het Centrum Indicatiestelling Zorg (CIZ), dat hun aanvraag beoordeelt. Het CIZ bepaalt of iemand inderdaad Wlz-gefinancierde zorg nodig heeft en welke zorgzwaarte. De personen met een Wlz-indicatie hebben een wettelijk recht op Wlz-gefinancierde zorg. Een persoon met een indicatie maakt met het zorgkantoor afspraken over de manier waarop de zorg wordt geleverd en door welke aanbieder. Mensen kunnen kiezen voor zorg in natura of voor een persoonsgebonden budget (PGB) waarmee ze de zorg zelf in kunnen kopen. Mensen die kiezen voor zorg in natura kunnen de zorg afnemen in een instelling of thuis. Een deel van de cliënten staat op een wachtlijst bij het zorgkantoor, omdat zij nog geen plek (van hun voorkeur) kunnen vinden.¹ Sommige mensen kiezen er ook voor om hun indicatie (nog) niet te verzilveren en melden zich niet bij het zorgkantoor.

De uitvoering is door de overheid gedelegeerd naar 9 Wlz-uitvoerders, waarvan 7 zijn aangewezen als zorgkantoor in één of meer regio's. In totaal gaat het om 31 zorgkantoorregio's. Binnen het regiobudget (zie paragraaf 2.2) kopen de Wlz-uitvoerders de Wlz-zorg in, of verstrekken de zorgkantoren persoonsgebonden budgetten, voor alle cliënten die daar in hun regio recht op hebben. Wlz-uitvoerders voeren de Wlz niet voor eigen rekening en risico uit. De uitvoerders sluiten contracten met de zorgaanbieders over de hoeveelheid zorg die ze kunnen leveren en tegen welke prijs. De zorgaanbieders sturen gedurende het jaar vervolgens declaraties naar de Wlz-uitvoerder voor de vergoeding van de door hen geleverde zorg.

2.2 Het Wlz-kader²

De Minister van VWS stelt voor ieder jaar het Wlz-kader vast voor de totale Wlz. Het (bijgestelde) kader van het voorgaande jaar vormt hierbij het startpunt. Dit kader wordt aangepast op basis van een groei ruimte voor de demografisch ontwikkelingen en toenemende zorgzwaarte, de verwachte loon- en prijsontwikkelingen, en beleidsmatige wijzigingen. Dit kader wordt in het jaar voorafgaand aan het betreffende kalenderjaar vastgelegd in de Voorlopige Kaderbrief (juni) en Definitieve Kaderbrief (oktober).

De NZa verdeelt het Wlz-kader vervolgens via een verdeelmodel over de verschillende zorgkantoorregio's. De Wlz-uitvoerders moeten binnen hun regionale budget blijven en eventuele overschrijdingen binnen het budget oplossen. Tegelijk is Wlz-zorg een recht en moeten er dus altijd genoeg middelen zijn om mensen die zorg nodig hebben ook zorg te bieden. Een deel van de middelen wordt daarom bij aanvang van het jaar achtergehouden door de minister van VWS: dit zijn de zogenaamde herverdelingsmiddelen. Deze herverdelingsmiddelen worden gebruikt om onverwachte verschuivingen in geleverde zorg tussen de regio's op te vangen.³ Verder kan de Minister van VWS gedurende het jaar besluiten om het Wlz-kader op te hogen.

¹ Er wordt onderscheid gemaakt tussen verschillende soorten wachtlijsten maar dat onderscheid staat in dit onderzoek niet centraal.

² Dit rapport focust op de prognose van het Wlz-kader. De verdeling van het Wlz-kader over de regio's en de verdeling over de individuele aanbieders worden gedaan op basis van andere modellen en blijven hier buiten beschouwing.

³ Daarnaast wordt een deel van het Wlz-kader (circa 0,3%) beschouwd als manoeuvreerruimte om het contracteringsproces soepel te laten verlopen. Deze middelen worden verondersteld niet benut te worden.

2.3 De budgetcyclus

Op basis van de definitieve kaderbrief voor het volgende jaar en de beleidsregels (inclusief maximumtarieven per prestatie) van de NZa, maken de Wlz-uitvoerders afspraken met alle individuele aanbieders binnen hun regio over de prijzen en omzet voor de te leveren zorg in natura. De afgesproken prijzen zijn lager dan of gelijk aan de maximumtarieven. De NZa toetst vervolgens in de budgetronde (15 november) of de gemaakte afspraken voor het *volgende* jaar passen binnen de contracteerruimte: het op basis van het Wlz-kader en het verdeelmodel vastgestelde regiobudget. Op basis van deze toetsing stelt de NZa per aanbieder een omzetplafond vast: indien de totale omzet in een regio past binnen het regiobudget dan worden de afspraken tussen de uitvoerder en de aanbieders gehonoreerd. Als het regiobudget wordt overschreden dan past de NZa een generieke korting op regioniveau toe naar rato van de overschrijding. De budgetronde is bedoeld om de aanbieders vooraf een plafond te stellen voor de hoogte van de aanvaardbare kosten.

Op 1 november van het *lopende* jaar is er een herschikkingsronde. Bij de herschikkingsronde kunnen aanbieders een bijstelling van de budgetaanvraag voor het lopende jaar aanvragen. De bijgestelde aanvragen worden wederom door de NZa getoetst aan de beschikbare contracteerruimte. Wanneer de bijgestelde aanvragen de beschikbare ruimte overschrijden, dan wordt op regioniveau een korting toegepast bij de zorgaanbieders die verzoeken om verhoging van hun eerder vastgestelde omzetplafond, naar rato van het aandeel van de aanvragen van deze zorgaanbieders op de totale toename.

Na afloop van het jaar vindt de nacalculatie plaats. De NZa toetst daarbij per zorgaanbieder de gerealiseerde omzet aan het individueel vastgestelde omzetplafond uit de budget- of herschikkingsronde. Overproductie kan – voor zover er (macro) financiële ruimte is – worden verrekend met onderproductie.

2.4 De adviescyclus

Functies van de adviezen

De NZa geeft advies op verzoek van de Minister van VWS. Dit verzoek wordt onder andere gedaan in de definitieve kaderbrief Wlz (zie voor de meest recente Ministerie van VWS (2020)) en specificeert waarover advies wordt gevraagd en wanneer. De adviescyclus, en ook de relaties tussen het advies en het kader, is niet in formele procedures vervat. De adviesverzoeken kunnen dus verschillen in inhoud, timing en frequentie.

Er wordt in de praktijk altijd advies gevraagd over de inzet van herverdelingsmiddelen en of het Wlz-kader toereikend is om de verwachte uitgaven van de Wlz-uitvoerders te dekken.⁴ Daarnaast wordt soms ook advies gevraagd over andere aspecten van de budgettering, zoals bijvoorbeeld de inzet van kwaliteitsbudget en de invloed van de maatregelen als gevolg van de Coronacrisis (Minister van VWS 2021).⁵ In sommige gevallen rekent de NZa op verzoek ook beleidsopties door in het advies: dit was in

⁴ Als we het in het stuk over toereikendheid hebben dan gaat het over deze definitie. Het gaat niet over de toereikendheid van de budgetten van individuele instellingen om hun kosten te dekken.

⁵ In de kaderbrief Wlz 2021 (Ministerie van VWS 2020) wordt de opdracht bijvoorbeeld als volgt geformuleerd:

“Ik verzoek de NZa mij in februari, mei en augustus te adviseren over de inzet van de herverdelingsmiddelen en de ontwikkelingen in het licht van de toereikendheid van het Wlz-kader in 2021. Ik verzoek de NZa dit te bezien in relatie tot de ontwikkeling van de indicaties en de gedeclareerde zorg waarbij ik in 2021 specifiek aandacht vraag voor de budgettaire impact van de instroom van cliënten met een psychische stoornis in de Wlz en het (financiële) effect van corona op het Wlz-kader. Ik verzoek de NZa in februari een raming met mij te delen van de compensatie van de extra kosten die het gevolg zijn van de coronacrisis voor zorg in natura en pgb, in samenhang met de benutting van het kwaliteitskader verpleeghuiszorg. De extra kosten als gevolg van de coronacrisis worden op basis van de betreffende NZa-beleidsregel buiten het Wlz-kader om bekostigd vanuit het Fonds langdurige zorg. Als laatste vraag ik de NZa om mij in de komende adviezen periodiek te informeren over de uitputting en andere ontwikkelingen in relatie tot het kwaliteitsbudget voor verpleeghuizen.”

2020 het geval met het verzoek om te schatten hoeveel extra budget het beheersbaar houden van de wachtlijsten kost (NZa 2020).

De adviezen van de NZa spelen geen expliciete rol in de vaststelling van het Wlz-kader voor het volgende jaar. In die zin zijn de vaststelling van het kader en de daaropvolgende prognose van de toereikendheid ervan gedurende het jaar twee losstaande werelden. Wel zijn de prognoses van de NZa van *indirecte* invloed op het Wlz-kader: bijstellingen van het Wlz-kader gedurende het lopende jaar veranderen ook het startpunt voor het volgende jaar of kunnen structurele doorwerkingen hebben. De NZa adviezen over de toereikendheid van het Wlz-kader om de Wlz-uitvoerders te laten voldoen aan de zorgplicht spelen voor de aanbieders een indirecte rol: de adviezen van de NZa hebben namelijk invloed op de contracteerruimte die bij de herschikkingsronde en het volgende jaar beschikbaar is.

Timing van de adviezen

De NZa brengt op verzoek van de minister van VWS meerdere keren per jaar advies uit: doorgaans gebeurt dat in het Mei-advies en het Augustus-advies; soms wordt daarnaast ook op andere momenten advies gevraagd (bijvoorbeeld het Februari-advies 2021). In alle adviezen rapporteert de NZa over de trend, en de verwachtingen ten aanzien van de toereikendheid van het Wlz-kader. Indien er nog herverdelingsmiddelen gereserveerd zijn adviseert de NZa over de inzet daarvan.

Onderbouwing van de adviezen

De NZa geeft het advies op basis van een set prognosemodellen (vanaf hier het basismodel genoemd), waarin de gedurende het jaar beschikbaar komende informatie over het volume en de prijs per prestatie en per Wlz-uitvoerder wordt gebruikt om de uitgaven voor de resterende maanden van het jaar te schatten. De resultaten worden geaggregeerd naar totale uitgaven (declaraties) voor zorg in natura en (toekenningen voor) persoonsgebonden budgetten (PGB). De resultaten worden meestal ook op Wlz-uitvoerder niveau gepubliceerd. De onderliggende prognoses op prestatieniveau worden alleen incidenteel gebruikt voor de inhoudelijke duiding van specifieke trends.

De adviezen bevatten niet alleen een puntschatting van de benutting van het Wlz-kader en een advies over de inzet van de herverdelingsmiddelen, maar bevatten ook een toelichtende duiding ondersteund met aanvullende analyses.

Consultatie

De NZa consulteert de Wlz-uitvoerders voordat ze advies uitbrengt. Bij die consultatie wordt het Ministerie van VWS in de gelegenheid gesteld om aan te sluiten als toehoorder. Als onderdeel van de consultatie van de Wlz-uitvoerders krijgen de uitvoerders de mogelijkheid om input te geven op een consultatiedocument dat de NZa opstelt. Dit document bevat de voorlopige prognose (op basis van gegevens die beschikbaar zijn op het moment van de consultatie), inclusief informatie over de methode en de input, maar niet het (concept)advies. Na de consultatie stelt de NZa de prognose bij met recentere data, en weegt de NZa af hoe de opgehaalde input uit de consultatie wordt verwerkt in het advies.

3 Het Wlz-kader en het advies als beleidsinstrumenten

In dit hoofdstuk bespreken we de rol van het Wlz-kader (paragraaf 3.1) en het NZa-advies (paragraaf 3.2) als beleidsinstrumenten. Inzicht in deze rollen helpt om de prestatie van het prognosemodel te kunnen evalueren.

Het Wlz-kader en het NZa-advies zijn twee van de beleidsinstrumenten die worden ingezet om de doelen van de Wlz te bereiken. De doelstelling van de Wlz is: het bieden van goede en toegankelijke zorg voor mensen die daar recht op hebben. Daarnaast zijn ook de betaalbaarheid, financiële beheersbaarheid en doelmatigheid van de zorg belangrijke doelen (Tweede Kamer 2013/2014).⁶ Om die doelen te bereiken worden echter ook andere beleidsinstrumenten ingezet. Zo reguleert de indicatiestelling het recht op zorg, zijn er gereguleerde (maximum) tarieven voor zorgprestaties. Hieronder richten wij ons op de vraag hoe het Wlz-kader en de prognose een bijdrage leveren aan de mogelijkheden voor Wlz-uitvoerders om aan de wettelijke zorgplicht te voldoen en aan het stimuleren van betaalbaarheid en doelmatigheid. Daarbij moet echter steeds in het achterhoofd worden gehouden dat de effecten uiteindelijk afhangen van de (soms mogelijk tegenstrijdige) prikkels en beperkingen in de rest van het systeem.⁷

3.1 De rol van het Wlz-kader

De twee functies van het kader

Figuur 1 geeft een schematisch overzicht van de zorguitgaven in de Wlz en de rol van het Wlz-kader. Het kader heeft twee functies. De eerste is dat het kader de zorgvraag volgt: het kader moet ruim genoeg zijn om zorgkantoren in staat te stellen om voldoende zorg in te kopen zodat mensen met een Wlz-indicatie de zorg krijgen waar ze recht op hebben en zo de toegankelijkheid te garanderen.

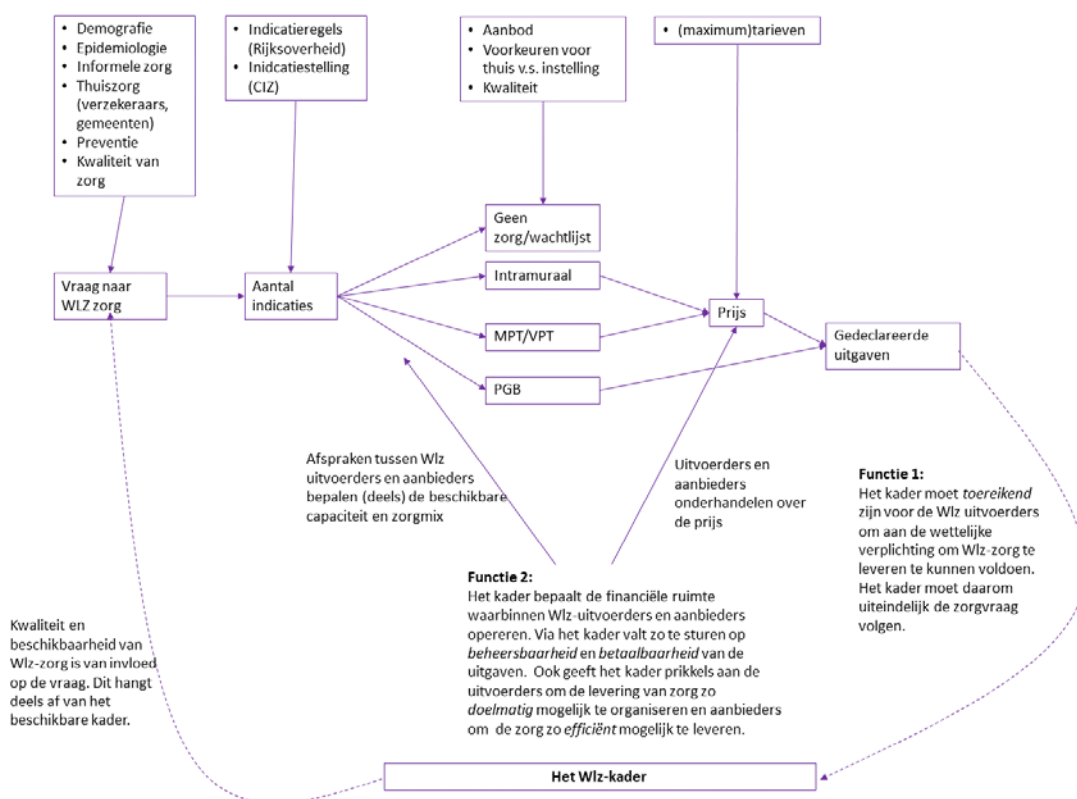
De tweede functie is dat het Wlz-kader de geleverde zorg (en daarmee de uitgaven) beïnvloedt, omdat het de ruimte bepaalt die Wlz-uitvoerders hebben bij het maken van afspraken over volume, prijs en de mix van leveringsvormen van zorg. Zo draagt het kader bij aan de betaalbaarheid en beheersbaarheid van de uitgaven. Ook kan dit prikkels voor doelmatigheid opleveren, zowel op het niveau van de Wlz-uitvoerders (via de verdeling naar het regiobudget) als van de aanbieders (via de budgetcyclus). Voor Wlz-uitvoerders geeft een (relatief) krap Wlz-kader een prikkel om de toegang tot zorg zo te organiseren dat er een optimale balans is tussen kosten, volume en kwaliteit van zorg. Zij kunnen die balans beïnvloeden door te sturen op het aanbod van verschillende leveringsvormen en het selectief inkopen bij verschillende aanbieders. Voor zorgaanbieders geeft het Wlz-kader een (indirecte) prikkel voor een doelmatige levering van zorg en een efficiënte bedrijfsvoering. Indirect kan het Wlz-kader ook nog een effect hebben op doelmatigheid via de zorgvraag wanneer het zorgaanbod invloed heeft op hoeveel potentiële cliënten een (her)indicatie aanvragen.

⁶ Betaalbaarheid gaat over de hoogte van het budget, beheersbaarheid over de omvang en frequentie van budgetoverschrijdingen en doelmatigheid over of het budget aan zorg wordt besteed op een manier die per bestede euro de meeste waarde oplevert voor de cliënt.

⁷ De situatie dat er meerdere instrumenten worden ingezet om de drie doelen van de publieke verzekering te behalen is niet uniek: dat is ook in de Zorgverzekeringswet het geval, waar het macrokader voor de ZVw ook niet het enige instrument is maar de betaalbaarheid verder wordt geregeld via onder meer het pakketbeheer, richtlijnen van beroepsgroepen, actieplannen, de eigen bijdrage, de nominale verzekeringspremie. De situatie is ook niet uniek in internationaal perspectief (zie bijvoorbeeld Bakx et al. 2020).

Bij het bepalen van de hoogte van het Wlz-kader moet de Minister dus rekening houden met deze twee functies van het kader. Die zijn tot op zekere hoogte tegengesteld: een ruim kader wordt eenvoudig voldaan aan het recht op zorg, maar neemt de prikkels voor doelmatigheid deels weg.⁸

Figuur 1: De rol van het Wlz-kader



De dynamiek in het Wlz-kader

Of het Wlz-kader bijdraagt aan het behalen van de doelen hangt onder meer af van de verwachtingen van alle betrokken partijen en hun gedragsreacties. Wanneer het oorspronkelijke Wlz-kader gedurende het jaar niet voldoende blijkt te zijn om aan de vraag naar Wlz-zorg te voorzien, kunnen de herverdelingsmiddelen worden ingezet en kan het kader worden verruimd, wat in de afgelopen jaren is gebeurd (Tabel 1). Dit betekent dat de Minister bij bepalen van het initiële Wlz-kader de doelmatigheids- en betaalbaarheidsoverwegingen een relatief grote rol zou kunnen laten spelen, omdat als het kader te krap blijkt te zijn dit gedurende het jaar kan worden opgevangen op basis van de prognoses van de NZa. Aan deze strategie zitten echter twee tegenovergestelde risico's. Enerzijds kan het vooruitzicht van een ophoging van het kader bij een dreigende overschrijding de prikkels voor doelmatige inkoop en levering bij Wlz-uitvoerders en aanbieders deels wegnemen, omdat ze bij een (collectieve) overschrijding van het Wlz-kader verwachten toch gecompenseerd te worden. Anderzijds kan onzekerheid over eventuele kortingen als gevolg een mogelijke overschrijding van het regiobudget aanbieders er ook van weerhouden om cliënten in zorg te nemen. Ook op langere termijn speelt deze tegengestelde dynamiek, waarbij Wlz-uitvoerders en aanbieders enerzijds belang hebben bij enige overproductie, omdat dit het uitgangspunt vormt voor het Wlz-kader van volgend jaar, en anderzijds onzekerheid over de financiering op lange termijn aanbieders er van kan weerhouden om te investeren in de benodigde extra capaciteit.

⁸ Ook hier geldt dat aanvullende sturingsmechanismen (zoals tariefregulering) kunnen helpen om deze twee verschillende doelen gezamenlijk te behalen (of tegenovergesteld de gewenste prikkels die uitgaan van het kader juist deels teniet kunnen doen).

Tabel 1: aanpassingen aan het Wlz-kader gedurende het jaar, 2018-2021, bedragen in miljoen euro's.

	2018	2019	2020	2021
Wlz-kader bij aanvang, inclusief herverdelingsmiddelen (contracteerruimte en pgb kader)*	21.230	22.144	23.492	25.672
Waarvan herverdelingsmiddelen	260	200	70	100
Wlz-kader bij aanvang, exclusief herverdelingsmiddelen (contracteerruimte en pgb kader)*	20.970	21.944	23.422	25.572
Totale ophoging door VWS gedurende het jaar naar aanleiding van NZa adviezen, mede door middel van inzet herverdelingsmiddelen**	130	720***	550	552**** (t/m juli)

* Inclusief Experiment persoonsvolgende inkoop (2018); exclusief kwaliteitsbudget verpleeghuiszorg (2019-2021)

** Niet altijd wordt exact aangesloten bij de geprognosticeerde tekorten door de NZa (mede vanwege het feit dat de NZa meerdere scenario's presenteert).

*** Waarvan € 60 miljoen incidenteel (bovenop het door de NZa geprognosticeerde tekort), om de capaciteit van de verpleeghuiszorg te vergroten en wachtlijsten weg te werken.

**** Deze ophoging is grotendeels herelayeerd aan de openstelling van de Wlz voor cliënten met een psychische stoornis.

Bronnen: Voorlopige en definitieve kaderbrieven (2017, 2018, 2019, 2020, 2021), NZa (2021a)

Bewijs over de effecten van het Wlz-kader ontbreekt

Bij het geven van het advies over het Wlz-kader door de NZa en het vaststellen van het kader door de Minister zou rekening gehouden moeten worden met de gedragsreacties van de Wlz-uitvoerders en aanbieders. Dit gebeurt maar zeer beperkt, ook omdat er over die gedragsreacties – en daarmee een aantal belangrijke relaties in Figuur 1 – geen sterk empirisch bewijs is. Zo is, vanwege de hierboven genoemde dynamiek, het effect van een krappere Wlz-kader op de beheersbaarheid, betaalbaarheid en doelmatigheid niet duidelijk. Dat geldt voor kortetermijneffecten: hoe gaan Wlz-uitvoerders in hun inkoopbeleid om met de situatie wanneer het regiobudget krappere is en hoe reageren aanbieders? En waar komt eventueel extra budget bij een verruiming van het Wlz-kader terecht: neemt het aantal cliënten in zorg toe, of verbetert de kwaliteit of de hoeveelheid zorg per cliënt? Maar ook voor langetermijneffecten: leidt een krappere Wlz-kader op lange termijn tot lagere uitgavengroei? Of worden alle tekorten gedurende het jaar opgelost door het Wlz-kader te verhogen en is er dus feitelijk sprake van (enigszins geremde of vertraagde) openeindefinanciering?

Daarnaast is niet bekend in hoeverre het Wlz-kader kan bijdragen aan een doelmatige invulling van de zorgplicht door de Wlz-uitvoerders.⁹ Het is niet evident hoe zo'n doelmatigheids-effect tot stand komt, want Wlz-uitvoerders hebben slechts beperkte beleidsvrijheid (Bakx et al. 2020). Zo heeft iedereen met een indicatie recht op zorg en kan een zorgkantoor dus geen invloed uitoefenen op het aantal personen dat zorg ontvangt. Ook zijn mensen in principe vrij om zelf een leveringsvorm te kiezen. Zorgkantoren kunnen sturen op de prijs en volume middels de contracten die ze met aanbieders afspreken maar ook die

⁹ De NZa (2021b) documenteert regionale verschillen in indicaties, verzilvering van indicaties, gebruik en toegankelijkheid van langdurige zorg. Er is echter geen zicht op wat deze verschillen verklaart; regionale verschillen in uitgaven zijn bijvoorbeeld niet zomaar toe te wijzen aan doelmatigheidsverschillen (CPB 2020; Ministerie van Financiën 2020).

mogelijkheden zijn beperkt.¹⁰ Verder kan het Wlz-kader een prikkel geven om een andere mix van leveringsvormen in te kopen. VPT- en MPT-tarieven zijn lager dan die voor reguliere instellingszorg.¹¹

De rol van het Wlz-kader ligt niet vast

Een groot deel van het benodigde Wlz-kader is objectief vast te stellen: het aantal cliënten met recht op Wlz-zorg en de (gereguleerde) tarieven van de verschillende (gereguleerde) prestaties bepalen in hoge mate hoeveel budget er nodig is. Juist bij de veranderingen in het kader bovenop deze objectief vast te stellen hoeveelheid, waar de besluitvorming in de praktijk om gaat, is de hoogte van het kader echter deels een politieke afweging. Dat komt in de eerste plaats omdat de afweging tussen het recht op zorg enerzijds en kostenbeheersing en doelmatigheid anderzijds uiteindelijk een politieke keuze is. Hetzelfde geldt voor de combinatie van beleidsinstrumenten – het Wlz-kader, de aanspraak, of andere restricties aan de vraag- of aanbodzijde – die van invloed zijn op de rol en effectiviteit van het Wlz-kader. In de tweede plaats is niet volledig bekend welke mate de budgettering bijdraagt aan de doelstellingen betaalbaarheid en doelmatigheid.

Dat de hoogte van het Wlz-kader uiteindelijk deels een politieke keuze is, is van belang voor het advies en de prognose, omdat welk model en advies het beste is afhangt van wat de politiek met het Wlz-kader beoogt. Dit bespreken we verder in paragraaf 3.3. De ambiguïteit over het doel van het Wlz-kader is daarnaast een reden waarom de prognose en het advies door de verschillende gebruikers (of stakeholders) verschillend gewaardeerd en geïnterpreteerd worden.

3.2 De rol van het advies

De Minister van VWS verzoekt de NZa om te adviseren over (onder meer) de inzet van de herverdelingsmiddelen en de toereikendheid van het Wlz-kader. In de recente adviesvragen aan de NZa wordt alleen toereikendheid van de middelen genoemd; andere mogelijke doelen zoals betaalbaarheid en beheersbaarheid worden niet genoemd. Dit betekent ook dat beoogde effecten op deze doelen, voor zover die in de vaststelling van het Wlz-kader wel een rol spelen¹², in de prognose over de toereikendheid van het kader niet expliciet worden meegenomen.

Aannames over doelmatigheid spelen wel een rol in de prognose. Om een inschatting te maken van de toereikendheid van het Wlz-kader moeten er door de NZa namelijk - impliciet of expliciet - veronderstellingen worden gemaakt over hoe Wlz-uitvoerders en zorgaanbieders, gegeven een bepaald budget, reageren op de vraag naar zorg. Stel dat het aantal afgegeven indicaties gedurende het jaar harder stijgt dan verwacht, leidt dit dan tot olopende wachtlijsten of kan de extra vraag binnen het huidige kader worden opgevangen door een grote aandeel extramurale zorgvormen? Dit is in feite een doelmatigheidsvraag waar de NZa in haar adviezen aannames over maakt.

¹⁰ Met behulp van volumeafspraken kan de uitvoerder weliswaar de door een bepaalde zorgaanbieder geleverde hoeveelheid zorg beperken, maar niet de totale hoeveelheid zorg binnen een regio. De mogelijkheden voor prijssturing worden beperkt door de tariefregulering vanuit de NZa en de rechterlijke uitspraak over generieke kortingen in 2020 (Skipr 2020, Skipr 2021). Tot slot is vanwege de krapte op de markt voor langdurige zorg weinig (dreiging van) concurrentie tussen aanbieders om contracten: het niet contracteren van een te dure of niet-doelmatige aanbieder is vanwege de krapte vrijwel nooit een optie. Daarnaast speelt mogelijk op de achtergrond een rol dat er relatief weinig informatie beschikbaar is over de kwaliteit van de geleverde zorg op basis waarvan uitspraken gedaan kunnen worden over verschillen tussen aanbieders in doelmatigheid (Bakx et al. 2020).

¹¹ Voor een (groot) deel zijn deze leveringsvormen goedkoper omdat er kosten verschoven worden, naar zorgontvangers (met name: kosten voor verblijf) en mogelijk in mindere mate naar andere begrotingsposten (Zvw, huurtoeslag). Zie CPB (2020). Daarnaast is niet bekend wat het effect op de totale maatschappelijke kosten is. Zie Hussem et al. (2020), Portrait en Koolman (2021).

¹² De NZa en het prognosemodel spelen geen expliciete rol in de vaststelling van het Wlz-kader. De prognoses en adviezen hebben indirect wel invloed op het Wlz-kader omdat het definitieve kader voor het huidige jaar het uitgangspunt vormt voor het komende jaar en veranderingen in het huidige kader langdurige consequenties kunnen hebben

Deze aannames worden tot op zekere hoogte geëxpliciteerd in de adviezen. De adviezen bevatten niet alleen een puntschatting van de benutting van het Wlz-kader en een advies over de inzet van de herverdelingsmiddelen, maar bevatten ook een toelichtende duiding ondersteund met aanvullende analyses. In veel gevallen hebben die aanvullende analyses direct betrekking op de aannames over de gedragsrespons. Zo bevatten de meest recente ramingen meerdere scenario's (zie hoofdstuk 4) die gebruikt kunnen worden om de benutting van het Wlz-kader onder verschillende aannames te laten zien.

Uiteindelijk hangt wat de beste prognose en advies is af van de doelen van de politiek (de Minister van VWS). Bij de prognose moet altijd een afweging gemaakt worden tussen verschillende risico's: wanneer de prognose relatief hoog blijkt te zijn, dan leidt dit tot hogere zorgkosten en mogelijk tot ondoelmatigheid. Blijkt de prognose aan de krappe kant te zijn, dan kan dit leiden tot ongewenste kortingen bij aanbieders of toenemende wachtlijsten. De optimale afweging tussen deze risico's hangt af van de weging van de onderliggende doelen door de politiek.¹³

¹³ Voor een beslisser is de voorspelling die het dichtst bij de verwachte werkelijkheid ligt niet altijd de optimale uitkomst om vanuit te gaan bij beleidskeuzes. Vaak vinden beslissers overschrijdingen van het budget bijvoorbeeld vervelender dan meevallers. In dat geval is het optimaal om het budget net wat ruimer te maken dan de beste voorspelling. Het CPB hanteerde in het verleden om een vergelijkbare reden bijvoorbeeld een voorzichtige raming voor het begrotingsbeleid, die lager was dan de meest waarschijnlijke raming (Don, 2007). In het geval van de Wlz gaat het vooral om een afweging voor de Minister tussen de twee doelen: kostenbeheersing en doelmatigheid aan de ene kant en voldoen aan de wettelijke plicht om zorg te leveren (voorkomen van wachtlijsten) aan de andere kant. De optimale hoogte van het budget hangt af van het relatieve gewicht dat de Minister geeft aan beide doelen en hoe onzeker de voorspelling en de gedragsreacties van cliënten en aanbieders zijn.

4 De prognoses

4.1 Wat is het doel van de basisprognose?

Het primaire doel van de basisprognose is de NZa ondersteunen in haar advies over de inzet van de herverdelingsmiddelen en de toereikendheid van het Wlz-kader. Maar de prognose heeft bredere relevantie en moet daarom ook bruikbaar zijn voor de andere partijen die in de besluitvorming en consultatie over het kader een rol hebben: het ministerie van VWS, Financiën en de Wlz-uitvoerders. De eisen die aan het model gesteld worden hangen af van de functies die het model kan hebben: voorspellen, verklaren en interpreteren.

Voorspellen

Het model moet volgens alle betrokken partijen een zo goed mogelijke en objectieve voorspelling maken van de benutting van het Wlz-kader. Vertaald naar statistische termen betekent dit dat het model een zo precies mogelijke puntschatting moet geven met een zo klein mogelijke voorspelfout. Dit is het standaardcriterium voor elk voorspelmodel. Er zijn ook talrijke statistische methoden ontwikkeld om een model volgens deze criteria te kiezen en te evalueren. In het geval van het prognosemodel voor de Wlz wordt dit doel echter gecompliceerd door een tweetal redenen: de performativiteit van het model en het ontbreken van een duidelijke doelstellingsfunctie.

Performativiteit wil zeggen dat de voorspelling van het model zelf een effect heeft op de uitkomsten. De prognoses van de NZa bepalen in hoge mate of er herverdelingsmiddelen of extra middelen worden ingezet, en dit beïnvloedt weer hoeveel middelen er uiteindelijk worden uitgegeven. Er is dus geen van het model losstaande werkelijkheid waartegen de voorspelkwaliteit van het model getoetst kan worden. Als het model iets anders zou voorspellen, dan zijn de uiteindelijke realisaties waarschijnlijk ook anders.

Wat de 'beste' voorspelling is hangt af van de doelstelling van de uiteindelijke gebruiker. De ontbrekende consensus over de weging van de verschillende doelen en het ontbreken van informatie over de (optimale) gedragsreacties van de verschillende partijen die het geven van het advies bemoeilijken spelen in dezelfde mate bij het model. Het is moeilijker om een zo goed mogelijke voorspelling te geven van de benutting van het Wlz-kader in de rest van het jaar als niet duidelijk is hoe uitvoerders en aanbieders, maar ook de Minister, reageren op veranderingen zoals een onverwachte stijging van de zorgvraag.

Voorspellen of verklaren?

Statistische modellen kunnen gebruikt worden voor verschillende doeleinden: voorspellen, het identificeren van specifieke (causale) relaties en het doorrekenen van verschillende beleidsalternatieven (wat-als scenario's). Een statistisch model dat geoptimaliseerd is voor een nauwkeurige voorspelling is meestal niet goed bruikbaar voor het identificeren van causale relaties of het doorrekenen van beleidsalternatieven. De doelstelling van het basismodel is om de uitgaven te voorspellen. De functionele vorm van het model en de manier waarop het wordt geschat sluiten aan bij deze doelstelling. Het model is daardoor minder geschikt voor de andere doelen en wordt hier ook niet voor gebruikt. Wel worden in het advies vaak aanvullende analyses gepresenteerd die trachten in te gaan op de onderliggende gedragsreacties of veranderingen (zie hoofdstuk 3).

Interpreteren

Ook, of misschien wel juist, voor een voorspelmodel is het van belang dat de gebruikers van het model de uitkomsten kunnen interpreteren. Het gaat er dan niet per se om dat alle oorzakelijke of statistische verbanden die (impliciet) ten grondslag liggen aan de voorspelling voor de gebruikers duidelijk zijn, maar

wel dat de gebruikers de aannames en trends die onder de voorspelling liggen in voldoende mate begrijpen om de beslissing die ze nemen op basis van de prognose te kunnen onderbouwen. Het gaat hierbij dan om vragen als:

- In hoeverre wijkt de voorspelling af van de vorige voorspelling en wat zijn de belangrijkste verklaringen voor dit verschil?
- Hoe verhoudt de voorspelling zich tot voor gebruikers van de adviezen (Minister en Wlz-uitvoerders) herkenbare ijkpunten zoals de hoeveelheid budget die nodig is om de huidige cliënten de rest van het jaar in zorg te houden?
- Welke invloed hebben de belangrijkste aannames op de voorspelling?

4.2 Wat voorspelt het model?

Het prognosemodel wordt meerdere malen per jaar gebruikt om de benutting van het Wlz-kader in het huidige jaar in te schatten. Dat gebeurt op basis van de meest recente trends in het gebruik en de prijs van zorg. Voor de gebruikers van de prognoses zijn vooral de totalen op macroniveau van belang. Aangezien er aparte kaders zijn voor zorg in natura en zorg gefinancierd door een pgb, worden deze apart geraamd. Prognoses op het niveau van individuele prestaties lijken voor de gebruikers van weinig intrinsiek belang te zijn, al kan dit soort detailinformatie soms worden gebruikt om trends op macroniveau nader te duiden.

4.3 Hoe werkt het basismodel?

De NZa maakt aparte prognoses van de uitgaven per prestatie en Wlz-uitvoerder. Deze prognoses worden achteraf opgeteld om tot totalen op nationaal (of Wlz-uitvoerder-) niveau te komen. Voor iedere prestatie per Wlz-uitvoerder worden de volumes voorspeld en vervolgens vermenigvuldigd met (een inschatting van) de prijs.

Prijzen

De prijs voor een prestatie wordt bepaald op basis van de gemiddelde gedeclareerde prijs per Wlz-uitvoerder gedurende het lopende jaar. Indien er nog geen gedeclareerde prijzen over het lopende jaar beschikbaar zijn worden andere bronnen gebruikt (bij voorkeur de declaratiegegevens van het voorgaande jaar, en de daaruit afgeleide afgesproken kortingspercentages op het NZa-maximumtarief).

Volumes

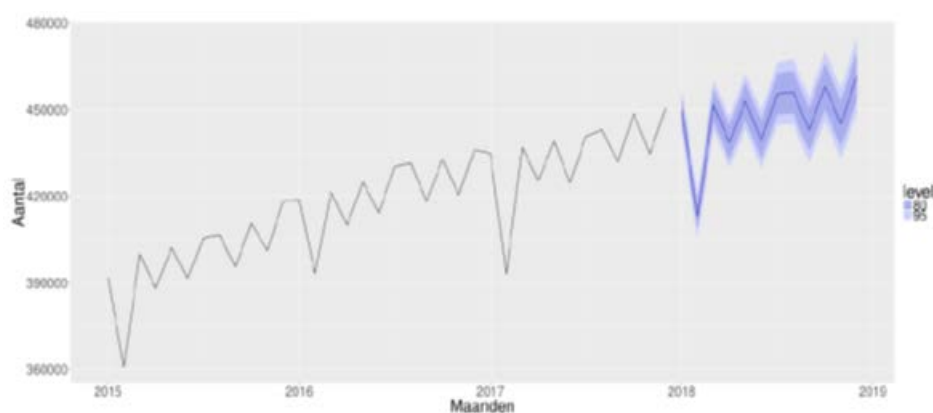
Om de volumes te voorspellen wordt per Wlz-uitvoerder en prestatie een apart model geschat. Dat model voorspelt vervolgens voor iedere resterende maand van het lopende jaar het volume. Het model wordt geschat op basis van maandelijkse tijdreeksdata over het aantal gedeclareerde prestaties. Deze data zijn voor de meeste prestaties beschikbaar vanaf 2015. Nieuwe data komen beschikbaar met een vertraging van 5 maanden.

Het model dat wordt gebruikt is een Error, Trend, Seasonality (ETS) model (Otexts, 2021). Het model maakt voor iedere resterende maand van het jaar een voorspelling op basis van de gerealiseerde uitkomsten uit het verleden. Een kenmerkende eigenschap van een ETS model is dat alle realisaties uit het verleden meetellen om een voorspelling te maken, maar dat realisaties van langer geleden een kleiner (exponentieel afnemend) gewicht krijgen dan de meest recente realisaties. Het model houdt rekening met tijdstrends in de voorspelling. Als er in de uitkomsten in het verleden een stijgende of juist dalende trend over de tijd te zien is, dan wordt die trend in de voorspellingen voor de toekomst doorgetrokken. Op eenzelfde manier wordt ook rekening gehouden met seizoenseffecten: als de uitkomsten in het verleden in het begin van het jaar bijvoorbeeld altijd relatief laag zijn en aan het eind van het jaar relatief hoog, dan wordt dat seizoenspatroon doorgetrokken in de voorspelling voor de komende maanden.

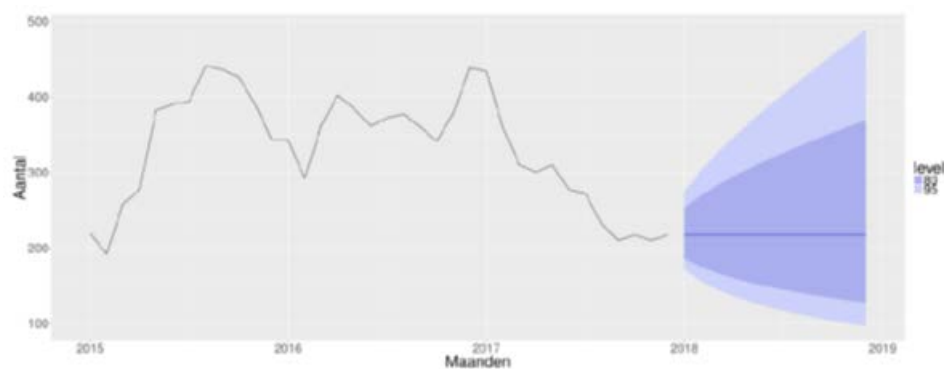
Er moeten een aantal keuzes worden gemaakt over de functionele vorm van het model. De modelleur moet kiezen of hij een tijdstrend en seizoenseffecten aan het model toevoegt en op welke manier (gedempt of ongedempt, additief of multiplicatief). In totaal zijn er 9 verschillende versies van het model die gebruikt kunnen worden. Welk van deze versies de beste voorspelling oplevert kan per uitkomst verschillen. Figuur 2 illustreert dit punt op basis van de voorspellingen van de ontwikkeling van twee willekeurige zorgprestaties voor het jaar 2018: voor de eerste is een duidelijke trend en seizoenseffect te zien in de periode 2014-2017, die vervolgens ook in de voorspelling van het model wordt meegenomen. Voor de tweede prestatie is geen duidelijke trend of seizoenseffect te zien en de voorspelling is daarom een rechte lijn. De figuur laat ook zien dat de voorspellingen omgeven zijn met onzekerheid. De blauwe vlakken geven het bereik aan waarbinnen de echte waarden met een bepaalde mate van zekerheid (95 respectievelijk 80 procent) zullen vallen.

Figuur 2: Voorbeelden van twee voorspellingen door het model voor twee zorgprestaties van een bepaalde Wlz-uitvoerder.

2a Prestatie Z053



2b Prestatie Vervoer dagbesteding GHZ



Het model maakt de keuze tussen de 9 mogelijke specificaties in iedere prognoseronde opnieuw voor ieder van de uitkomsten. Daarbij kiest het model de specificatie die het beste past bij de historische trend. Dit gebeurt op basis van een vergelijking van de waarde van het Akaike Information Criterion (AIC) of een variant hiervan die rekening houdt met kleine steekproefgroottes (AICc) tussen de verschillende modelversies. Zowel het AIC als het AICc zoeken het model uit dat met de minste parameters het beste op de eerdere waarnemingen past. De manier waarop de NZa de keuze voor het model maakt houdt dus in dat de voorspellingen voor verschillende uitkomsten gebaseerd (kunnen) zijn op verschillende specificaties en dat de modelspecificatie voor een uitkomst kan verschillen tussen prognoserondes.

Aanvullende data en aannames

Additionele bronnen (b.v. herschikkings- en nacalculatiegegevens) worden gebruikt om een vollediger beeld te krijgen van een drietal posten:

- Crisis-bedden. Zorgaanbieders krijgen een beschikbaarheidsvergoeding voor niet-bezette crisisbedden die niet via de standaard declaratiestroom verloopt. Hiervoor wordt op basis van data over gedeclareerde crisis-bedden en aannames over de bezettingsgraad een inschatting gemaakt.
- Na-ijleffecten. Een deel van de declaraties door zorgaanbieders komt pas binnen na de (initiële) data aanlevering. De data voor de meest recente data zijn dus een onderschatting van het daadwerkelijke aantal declaraties. Daarom wordt een correctie gedaan op basis van de gemiddelde na-ijleffecten over de afgelopen 4 jaar per Wlz-uitvoerder per maand. Alleen de gerealiseerde maanden worden opgehoogd.
- Na-ijleffecten meerzorg. Voor meerzorg geldt hetzelfde principe, alleen is de na-ijl daar groter dan bij reguliere zorg. Vandaar dat een aparte ophoging wordt gedaan.

Voor de prognoses voor het pgb kader wordt gebruik gemaakt van de afgegeven verlengingsbeschikkingen per maand.

Bijstellingen

De resultaten worden getoetst op hun plausibiliteit en eventueel bijgesteld. Die toets gebeurt niet alleen intern maar ook samen met Wlz-uitvoerders, door middel van een consultatie. Op die manier worden ook andere zaken zoals recente beleidsontwikkelingen of informatie uit de zorginkoop en uit andere bronnen gebruikt. Aanpassingen kunnen relevant zijn wanneer bijvoorbeeld blijkt dat er ontbrekende gegevens zijn in de door het model gebruikte brondata, of dat bijvoorbeeld recente trends nog niet in de door het model gebruikte data zijn opgepikt.

4.4 Alternatieve scenario's

Nul-scenario

In de meeste recente adviezen (vanaf 2020) presenteert de NZa naast het basismodel twee alternatieve scenario's. Het eerste scenario is een "naïeve raming" of "nul-scenario". In dit scenario worden de uitgaven over de rest van het jaar bepaald op basis van de cliënten die bij de start van het jaar al zorg gebruiken, zonder rekening te houden met een groei in het aantal cliënten gedurende het jaar. De uitgaven die het komende jaar nodig zijn om de cliënten die er in het begin van het jaar al zijn wordt het overloopeffect genoemd. Aangezien het aantal cliënten normaliter (alleen al op basis van demografie) gedurende het jaar zal stijgen, kan het nul-scenario worden gezien als een ondergrens voor de benodigde middelen.¹⁴

Vraag-gebaseerd scenario

Het tweede alternatieve scenario is gebaseerd op de trend in het aantal indicaties. Het idee is dat indicatiedata eerder beschikbaar komen (er zit tijd tussen indicatie en verzilvering) en zo eerder belangrijke ontwikkelingen in de zorgvraag op kunnen pikken. De alternatieve prognose wordt uitgevoerd op landelijk niveau en is gebaseerd op gemiddelde groei in het aantal indicaties per zorgzwaarteprofiel in de voorgaande 12 maanden. In tegenstelling tot het ETS-model worden voor dit scenario geen gegevens over een langere periode gebruikt en vervolgens op basis van een model gesmoothed en ontdaan van trend en seizoenseffecten, maar wordt de trend simpelweg lineair doorgetrokken. De relatie tussen indicatie en gedeclareerde uitgaven hangt af van of en wanneer de indicatie daadwerkelijk wordt verzilverd en welke

¹⁴ De benodigde middelen zouden ook lager kunnen zijn dan die in het nul-scenario, wanneer er gedurende het jaar meer cliënten uitstromen (bijvoorbeeld door overlijden) dan dat er instromen. Gezien de vergrijzing ligt dit (in ieder geval voor de VV-sector) in normale omstandigheden niet voor de hand.

leveringsvorm wordt gekozen. Als benadering voor de kosten per zorgprofiel worden de (landelijk gemiddelde) kosten per persoon met een indicatie in het voorgaande jaar genomen.

Waar het basismodel een voorspelling maakt op basis van geleverde zorg, wordt in dit alternatieve scenario geraamd op basis van de zorgvraag. Een belangrijk verschil is dat een prognose op basis van zorgvraag geen rekening houdt met eventuele veranderingen in leveringsvorm, die wel (impliciet) in het basismodel zitten (zie ook paragraaf 3.2). De Wlz-uitvoerders of aanbieders zouden tot op bepaalde hoogte de zorguitgaven kunnen remmen door extra zorgvraag op te vangen met (relatief goedkope) extramurale leveringsvormen of (vanwege een tekort aan passend aanbod) door de wachtlijsten op te laten lopen. Dit effect zit niet in het scenario op basis van zorgvraag, waardoor dit scenario ook wel het "ongeremde groei"-scenario wordt genoemd. Daarmee is dit in zekere zin een bovengrens van de te verwachten groei. Waar het echter vrij evident is dat het nul-scenario een absolute ondergrens vormt, is het tweede scenario niet een absolute bovengrens: de groei zou ook nog hoger kunnen zijn. Bovendien is de interpretatie van de prognose op basis van indicaties als "ongeremde groei" weliswaar intuïtief, maar het is niet zo dat de verschillen tussen beide prognoses per se alleen door dit verschil komen. Zo speelt ook het feit dat indicatiedata voorlopen op declaraties bijvoorbeeld een rol. Er is bovendien geen zicht op in hoeverre WLZ-uitvoerders (of aanbieders) daadwerkelijk op doelmatigheid kunnen sturen (paragraaf 3.1) en dus of er op korte termijn een remmende werking van uitgaat.

Tabel 2: overzicht van de verschillen en overeenkomsten tussen het basismodel en de alternatieve scenario's

	Basismodel	Nulscenario	Vraaggebaseerd-model
Input	Declaratiedata per prestatie en Wlz-uitvoerder per maand vanaf 2015.	De (ingeschatte) stand van het aantal zorggebruikers aan het begin van het jaar.	Indicaties, per zorgzwaarteprofiel, per maand over de afgelopen 12 maanden.
Output	Uitgaven per maand gedurende de rest van het lopende jaar.	Totale (deel van de) jaaruitgaven dat besteed wordt aan bestaande cliënten.	Uitgaven per maand gedurende de rest van het lopende jaar.
Schattingsniveau	Per prestatie, per uitvoerder	Nationaal niveau	Per zorgzwaarteprofiel
Type model	ETS model, waarin trends en seizoenseffecten uit het verleden op een gewogen manier worden meegenomen.	Naïef model: het aantal cliënten wordt gedurende het hele jaar gelijk verondersteld aan het aantal aan het begin van het jaar.	Extrapolatie van een lineaire trend.

De rol van de alternatieve scenario's

Het doel van de alternatieve scenario's die de NZa presenteert is niet volledig geëxpliciteerd. Zij lijken deels ook voort te komen uit de expliciete verzoeken van de Minister van VWS aan de NZa om in het advies rekening te houden met de ontwikkelingen in de indicaties.¹⁵

De scenario's kunnen helpen om drie verschillende doelen te bereiken. Ten eerste kunnen de scenario's de gebruikers van de prognoses helpen om de prognose te begrijpen en te duiden. Het nul-scenario vormt bijvoorbeeld een intuïtief startpunt, dat ook voor Ministers en Wlz-uitvoerders herkenbaar en bruikbaar is. Met dit scenario valt meteen te zien hoeveel van de in het Wlz-kader ingeruimde groeimiddelen aan het begin van het jaar al 'op' zijn. Het maakt ook inzichtelijk hoeveel extra groei er in de basisprognose wordt verondersteld *boven op* de zorg die nodig is voor de huidige cliënten.

¹⁵ Het verzoek in de Voorlopige kaderbrief Wlz 2021 (Ministerie van VWS 2020) luidt bijvoorbeeld: "Ik verzoek u om mij in uw komende adviezen te blijven adviseren over de inzet van de herverdelingsmiddelen en de ontwikkelingen in het licht van de toereikendheid van het Wlz-kader. Ik verzoek u dit te bezien in relatie tot de ontwikkeling van de indicaties en gedeclareerde zorg, waarbij ik u in 2021 specifiek aandacht vraag voor de budgettaire impact van de instroom van cliënten met een psychische stoornis in de Wlz."

Ten tweede kunnen de scenario's dienen als 'zijlichten' om de resultaten uit het prognosemodel te kunnen duiden. Een grote afwijking tussen het basismodel op basis van geleverde zorg en het scenario op basis van indicaties kan bijvoorbeeld duiden op een plotselinge verandering in de zorgvraag die nog niet in de declaratiedata wordt opgepikt. Dat kan dan aanleiding zijn om deze verandering nader te onderzoeken en de prognose of het advies bij te stellen.

Ten derde kunnen de scenario's ook een manier zijn om de onzekerheid in de ramingen richting Minister te presenteren. Het basismodel voorspelt de verwachte omvang van het Wlz-kader, maar die voorspelling is omgeven met onzekerheid. De Minister moet een keuze maken tussen een kader dat niet te hoog is, om zo een zekere doelmatigheidsprikkel te behouden, en een kader dat ruim genoeg is om aan de wettelijke verplichting tot het leveren van zorg te voldoen. In dat geval zou er iets voor te zeggen zijn om de Minister een meer gematigd en een hoog scenario te presenteren, en hem zelf de afweging tussen de twee verschillende doelen te laten maken (zie paragraaf 3.2). Concreet zou het hoge-uitgavenscenario kunnen bestaan uit een projectie waarbij er voor iedereen met een indicatie voldoende budget is voor de rest van het jaar. Het vraag-gebaseerd scenario dat de NZa als alternatief presenteert is dit in feite al, hoewel de verhouding tot de basisraming daar gecompliceerd wordt door de andere modelaanpak.

4.5 De interpretatie van de modelresultaten

Zoals beschreven in paragraaf 4.1 is het voor de kwaliteit van de besluitvorming over het Wlz-kader van belang dat de verschillende gebruikers van het model in staat zijn de prognose te interpreteren en te relateren aan hun eigen voorspellingen en/of herkenbare ijkpunten. De adviezen van de NZa bieden met het nul-scenario en het vraag-gestuurd scenario twee duidelijke ijkpunten die ook voor gebruikers van de adviezen herkenbaar zijn. Uit de gesprekken met de verschillende stakeholders blijkt echter dat de presentatie van de prognoses op twee punten nog verbeterd kan worden.

Ten eerste lijken de verschillende gebruikers het moeilijk te vinden om uit de adviezen de informatie te destilleren die voor hen van belang is. Ze zoeken vooral inzichten op macroniveau die direct vergelijkbaar zijn met ijkpunten zoals het aantal cliënten dat op dit moment (of aan het begin van het jaar) zorg gebruikt, en de relatie met de geobserveerde macrotrends in declaraties en indicaties. Deze informatie is veelal wel beschikbaar in de adviesbrieven, maar wordt kennelijk niet op zo'n manier gepresenteerd dat die voor de gebruikers meteen hanteerbaar is. Een heldere (grafische) weergave van de verhouding van de basisprognose ten opzichte van de ijkpunten en het (ontbreken van) een consistente presentatie van de resultaten over de adviezen heen lijken daarbij van belang te zijn.

Ten tweede worden in de consultatieronde vaak hele specifieke elementen uitgevraagd (bijvoorbeeld over de prestatie vervoer) waarvan de relevantie voor het totale kader voor de geconsulteerde partijen niet altijd duidelijk is. Hier lijkt sprake van een tegenstelling in verwachtingen over het doel van de consultatie tussen enerzijds de NZa en anderzijds de geconsulteerde partijen, met name de Wlz-uitvoerders. Voor de NZa is het doel een zo goed mogelijk onderbouwing van gehanteerde *methode*, in termen van de plausibiliteit van specifieke data en aannames. Voor de geconsulteerde partijen is vooral de mogelijkheid om input te geven over de *uitkomst* (de prognose van de benutting van het Wlz-kader) van belang.

4.6 Hoe goed voorspelt het model?

Voorspelkwaliteit en performativiteit

De voorspelkwaliteit is niet het enige, maar wel een belangrijk criterium om de prestaties van het prognosemodel te beoordelen: hoe onzeker zijn de prognoses die uit het model komen, hoe groot is de foutmarge, en zijn de prognoses structureel hoger of lager dan de uiteindelijke realisaties? En zijn de foutmarges van prognose hanteerbaar voor de beslisser (de Minister)? Normaliter zouden we de

voorspelkwaliteit van het model kunnen beoordelen door de voorspellingen uit het verleden te vergelijken met de realisaties¹⁶. Een belangrijke complicatie is echter de performativiteit van het model: de prognoses van de NZa hebben een indirecte invloed op de hoeveelheid zorg die uiteindelijk wordt geleverd. Dit betekent dat een vergelijking met de gerealiseerde uitgaven aan het eind van het jaar beperkt informatief is. Het is daarom bijvoorbeeld lastig om te toetsen of een model waarin aanvullende informatie wordt toegevoegd of dat op een ander aggregatieniveau wordt geschat beter zou presteren: het huidige model bepaalt namelijk deels de werkelijkheid en scoort alleen om die reden al goed.

Desondanks blijft een analyse van hoe het model structureel presteert zinvol. De gerealiseerde uitgaven zijn immers niet alleen een functie van de modelvoorspelling, maar ook van andere factoren, zoals de autonome ontwikkeling van de zorgvraag. Een analyse van de voorspelkwaliteit van het model over meerdere jaren, en hoe die zich verhoudt ten opzichte van andere modellen ontbreekt op dit moment. Wel heeft de NZa op ons verzoek met terugwerkende kracht de prognoses uit 2019 vergeleken met de uiteindelijke realisaties. Er is gekozen voor het jaar 2019, omdat het prognosemodel in dat jaar relatief slecht presteerde. Deze vergelijking geeft inzicht in hoe het model werkt en waardoor afwijkingen tussen de prognose en de realisatie kunnen ontstaan, maar zegt dus weinig over hoe het model structureel presteert. Daarvoor is een soortgelijke analyse over meerdere jaren nodig.

Hoe presteerde het model in 2019?

Figuur 3 vergelijkt de voorspelde uitgaven voor zorg in natura in 2019 op verschillende adviesmomenten¹⁷ (Augustus advies, Addendum (september data), Addendum (oktober data)) met de daadwerkelijk gerealiseerde uitgaven. De voorspelde uitgaven zijn uitgedrukt als een fractie van de gerealiseerde uitgaven: een waarde van 1 betekent dat de prognose overeenkomt met de realisatie, een waarde kleiner dan 1 betekent dat de prognose de gerealiseerde uitgaven onderschat en een waarde groter dan 1 betekent een overschatting.

Een aantal dingen vallen op. Ten eerste neemt de voorspelkwaliteit gedurende het jaar toe: de voorspellingen later in het jaar hebben een waarde dichterbij de 1 dan de voorspellingen uit augustus. Dit patroon is ook logisch: de meer recente prognoses hebben beschikking over meer en meer recente data (deels ook al de realisaties gedurende het jaar). Ten tweede was er in 2019, als we kijken naar de totale uitgaven, sprake van een initiële onderschatting van de benodigde middelen, die pas in de prognose van september volledig is verdwenen. Ten derde illustreert de exercitie dat het ingewikkeld is de voorspelkwaliteit van het model goed te waarderen. De augustusprognose voor het totale kader zat ongeveer 1 procent beneden de uiteindelijk gerealiseerde waarde. Voor een voorspelmodel is zo'n afwijking op het eerste gezicht niet per se een slechte score¹⁸. In euro's gaat het echter wel om een bedrag van 200 miljoen. Tot slot is er onder de raming voor het totale kader sprake van aanzienlijke verschillen tussen sectoren in hoe goed het model voorspelt.

Figuur 4 laat de maand-op-maand voorspellingen en uiteindelijke realisaties zien voor verschillende sectoren. Voor de sector Verpleging en Verzorging (VV) weet het model al vanaf de augustusraming consistent de uitgaven per maand, en bijvoorbeeld het seizoenspatroon daarin, goed te voorspellen. Voor de GGZ voorspelt het model, op basis van trends die zich in voorgaande jaren kennelijk hebben voorgedaan, in de eerste twee rondes een dalende trend in de uitgaven per maand. Dat die dalende trend zich in 2019 niet voordoet, wordt in het model pas in de oktoberraming opgepikt. Voor de gehandicaptenzorg (GHZ) is in de augustusraming sprake van een substantiële onderschatting, maar die is

¹⁶ Of door het model te herschatten op slecht een deel van de beschikbare observaties en vervolgens te onderzoeken hoe het model het niet meegenomen deel van de observaties kan voorspellen.

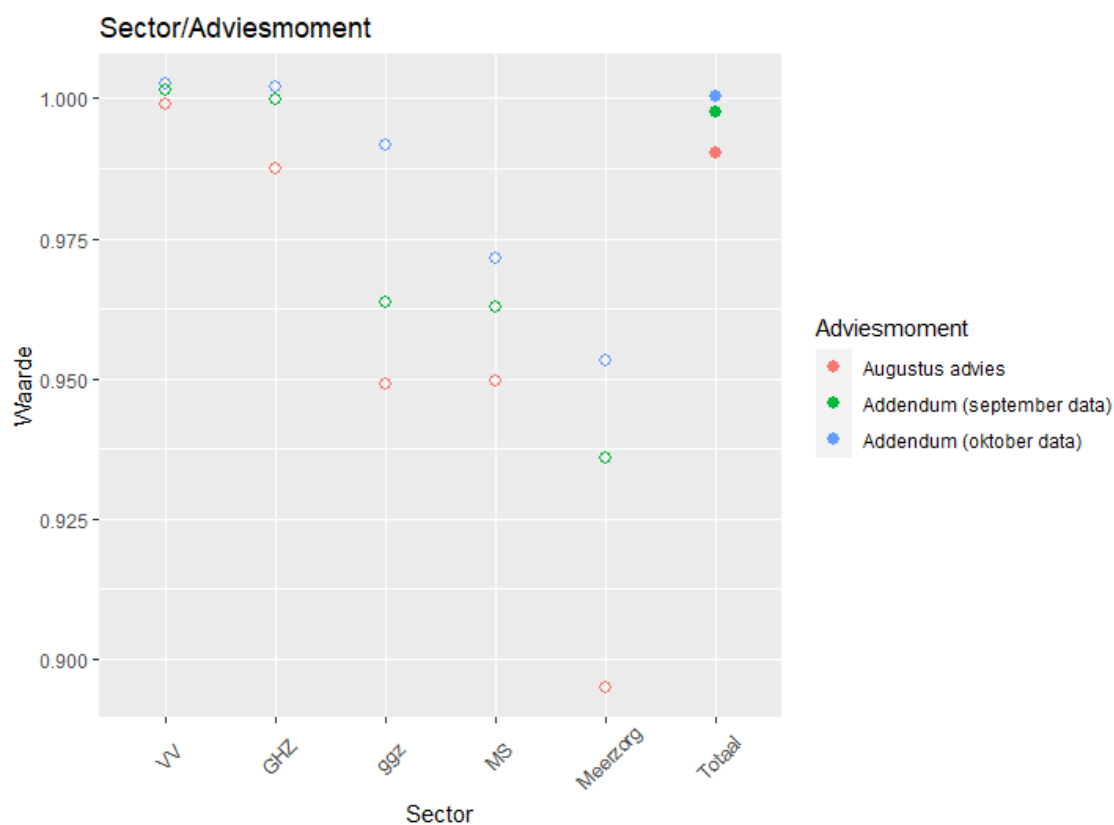
¹⁷ Replicatie van de prognose uit het mei-advies was om data technische redenen op korte termijn niet mogelijk. De prestaties van deze eerdere prognose zouden in de door ons geadviseerde uitbreiding van de evaluatie van de voorspelkwaliteit over meerdere jaren moeten worden meegenomen.

¹⁸ De enige manier om de kwaliteit vanuit modeloptiek echt goed in te kunnen schatten is de vergelijking met andere modellen: zijn er alternatieven die het structureel beter doen?

in de daaropvolgende ramingen zo goed als verdwenen.¹⁹ Voor Meerzorg is sprake van een onderschatting van de uitgaven die zelfs in de laatste ramingsronde niet volledig wordt ingelopen.²⁰

In 2018 en 2019 steeg de zorgvraag in de Wlz relatief snel, en bleek het prognosemodel deze vraag, vooral in het augustusadvies, te onderschatten. Hieruit zou wellicht de indruk kunnen ontstaan dat een ander model, of andere data, tot betere voorspellingen zou leiden. Het is de vraag in hoeverre de specifieke omstandigheden in deze periode iets zeggen over de *structurele prestaties* van het model. Figuren 3 en 4 laten dan ook vooral zien hoe zo'n prestatie-analyse eruit zou kunnen zien.

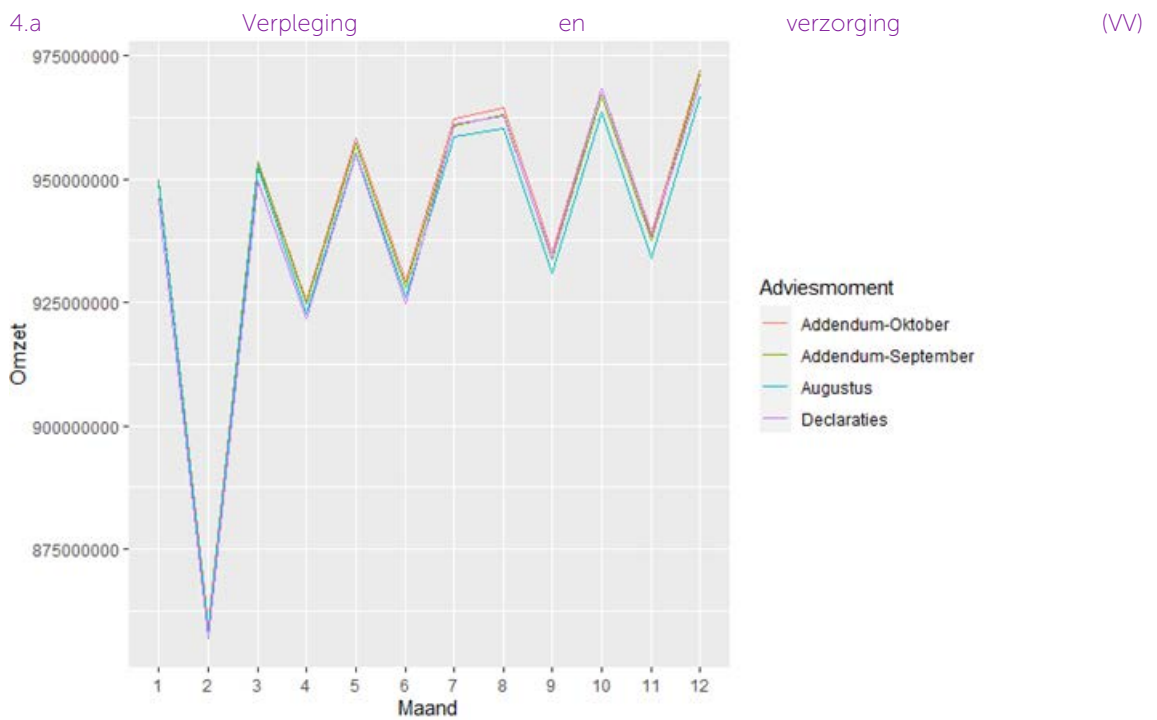
Figuur 3: Hoe goed voorspelde het prognosemodel in 2019? De voorspelling op verschillende momenten in het jaar als percentage van de uiteindelijk gerealiseerde omzet.



¹⁹ Deze onderschatting zou te maken kunnen hebben met de wijziging van de tarieven in de GHZ per 1 januari 2019 (NZa, 2019a).

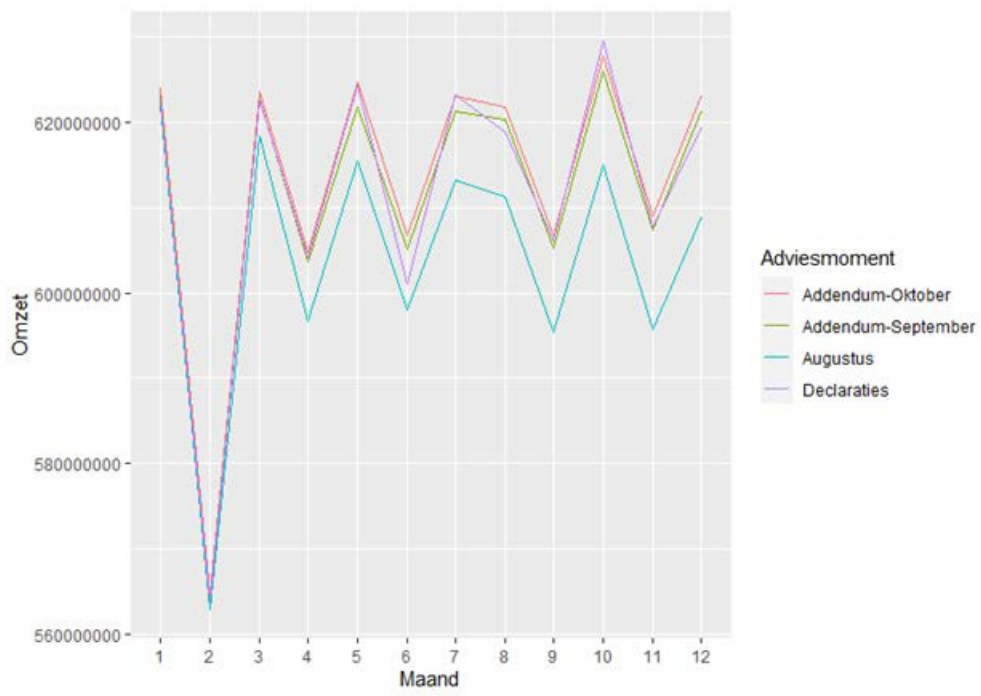
²⁰ Per 1 januari 2019 is de meerzorgregeling verruimd, wat invloed zou kunnen hebben op de kwaliteit van de prognose (NZa, 2019b).

Figuur 4: De maand-voor-maand voorspellingen va het prognosemodel op verschillende momenten, per sector.²¹

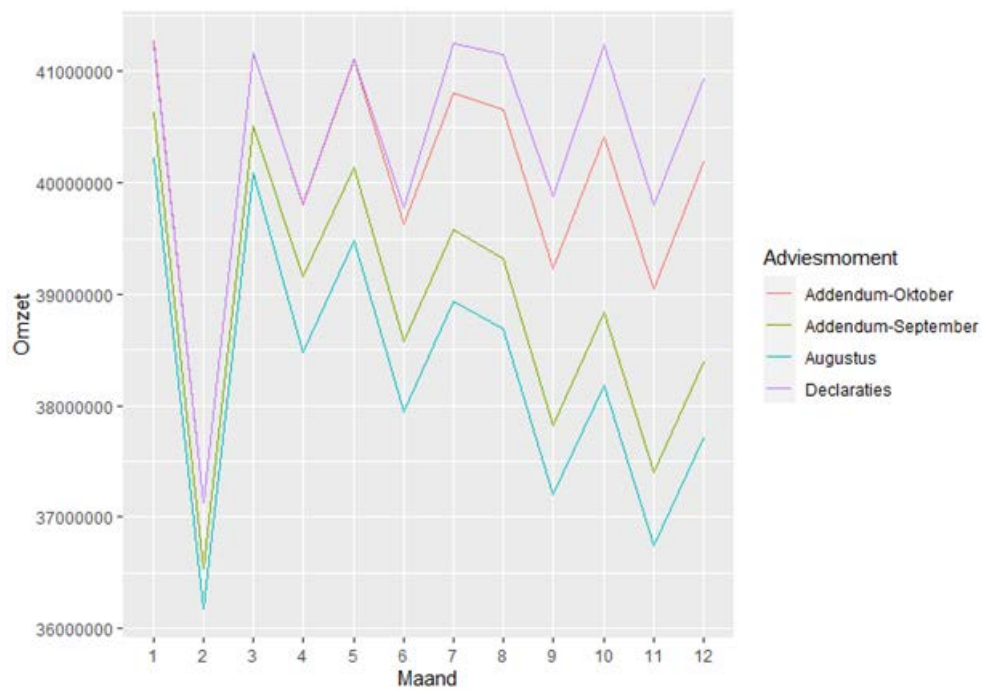


4.b Gehandicaptenzorg (GHZ)

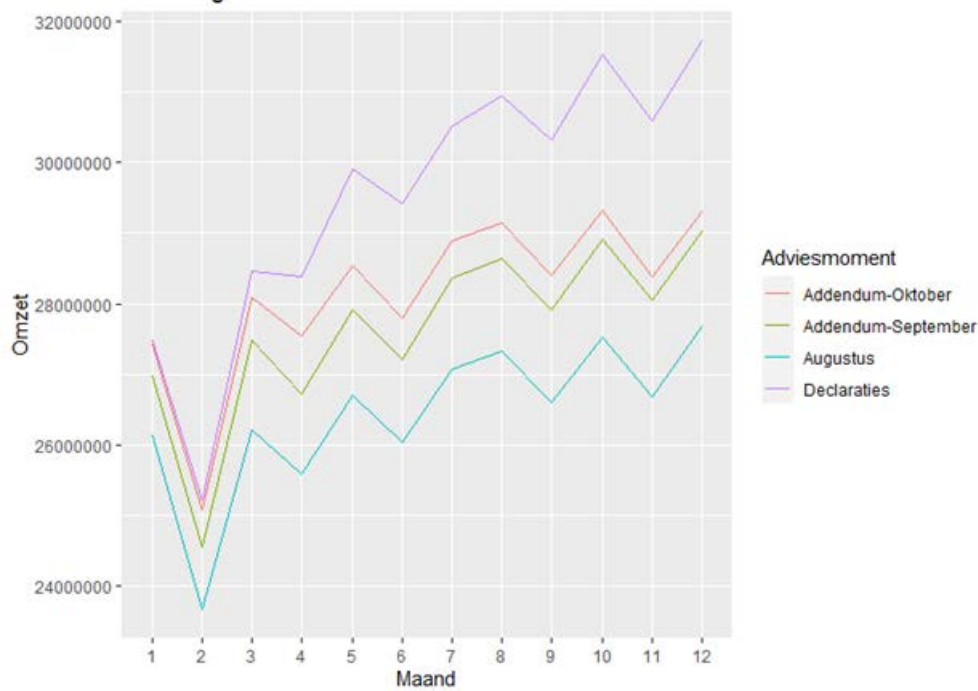
²¹ Ook voor de modelprognoses later in het jaar zit er een kleine afwijking tussen de prognose en de realisatie (terwijl data over het feitelijk aantal indicaties in die maanden op het moment van de prognose al beschikbaar is). Dit kleine verschil komt door de manier waarop in de prognose met na-ijl wordt omgegaan.



4.c Geestelijke gezondheidszorg (GGZ)



4.d Meerzorg



4.7 Zijn er alternatieven voor het huidige model?

Het door de NZa gekozen ETS-model past goed bij de structuur van de data en de doelstellingen van de prognose. Er moet een maand-op-maand voorspelling gemaakt worden over de relatief korte periode van maximaal 1 jaar. Zorgprestaties hebben in veel gevallen een tijdstrend en er kan sprake zijn van seizoenseffecten. Wanneer ervoor wordt gekozen om iedere individuele prestatie per WLZ-uitvoerder te ramen, zoals de NZa doet, dan ligt het ook voor de hand om de modelspecificaties te laten variëren tussen de prestaties, omdat niet iedere prestatie eenzelfde tijds patroon heeft. Gezien het grote aantal prestaties ligt de data-gedreven en geautomatiseerde aanpak om die specificaties te kiezen aan de hand van de in-sample fit voor de hand.

Veelgebruikte alternatieven om voorspellingen te maken voor data zoals deze zijn tijdreeksmodellen (zoals de verschillende vormen van het ARIMA-model). Deze modellen hebben een andere achtergrond en ook een enigszins andere functionele vorm dan het ETS-model, maar doen in wezen hetzelfde: het maken van voorspellingen op basis van (gewogen) realisaties uit het verleden, rekening houdend met eventuele tijdstrends en seizoenseffecten. Vanwege de grote gelijkheid in aanpak is het onwaarschijnlijk dat tijdreeksmodellen veel beter zullen voorspellen dan het huidige ETS-model. Strikt genomen is dit echter een (onbeantwoorde) empirische vraag. Ook in termen van interpreteerbaarheid of hanteerbaarheid bieden tijdreeksmodellen geen evidente voordelen ten opzichte van de huidige aanpak. Wel is het zo dat het meenemen van extra verklarende variabelen (zoals bijvoorbeeld indicatiedata, zie 4.8) eenvoudiger is in een ARIMA-model dan in een ETS-model. En extra verklarende variabelen kunnen mogelijk leiden tot een betere voorspelling.

Een ander alternatief ligt in de hoek van de datascience. Onder de noemer van datascience (ook wel machine learning of big data) zijn de afgelopen jaren een aantal verschillende methoden ontwikkeld om op basis van een grote variëteit aan vaak ongestructureerde data zo goed mogelijk voorspellingen te

maken. Deze modellen bezitten een grote mate van flexibiliteit waardoor ze gebruik kunnen maken van niet van tevoren gespecificeerde verbanden tussen een grote hoeveelheid aan variabelen. Ook voor deze modellen geldt dat ze niet per se beter zullen presteren dan het huidige ETS model. Data science modellen zijn namelijk vooral goed in het omgaan met zeer grote hoeveelheden data (zowel in termen van aantallen observaties als aantallen mogelijk relevante variabelen), terwijl de omvang van de data die de NZa gebruikt relatief beperkt is. Bovendien zijn dit soort modellen in veel gevallen een 'black-box', waaruit achteraf moeilijk te destilleren valt waarom het model tot een bepaalde voorspelling is gekomen. Data science methoden zouden wel gebruikt kunnen worden voor exploratieve analyses om mogelijke relevante verbanden of nieuwe informatiebronnen te identificeren.

Naast de keuze voor een ander type model zijn er ook andere keuzes die binnen de huidige aanpak gemaakt kunnen worden: over het aggregatieniveau en over de samenhang tussen de verschillende voorspellingen. De NZa kiest ervoor om iedere prestatie per Wlz-uitvoerder apart te voorspellen. De voorspellingen geven dus informatie over de ontwikkeling van het zorgvolume op detailniveau. In de praktijk wordt deze informatie echter nauwelijks gebruikt, door de NZa zelf noch door de andere belanghebbende partijen. Vooral de prognose voor het gehele kader (de som van alle individuele prognoses) is van belang, of eventueel de totale ontwikkeling per Wlz-uitvoerder. Alleen in zeer specifieke gevallen, zoals bij de introductie van een nieuwe prestatie waarvan de ontwikkeling met veel onzekerheid is omgeven, wordt er ingezoomd op individuele prestaties. Het valt dus te overwegen om ook de prognose zelf op een hoger aggregatieniveau uit te voeren. Of een geaggregeerd model betere voorspellingen oplevert dan een gedisaggregeerd model lijkt, ondanks bepaalde theoretische inzichten hierover, in de praktijk vooral een empirische vraag (Barker en Pesaran 2011). Naast de voorspelkwaliteit zijn echter ook andere overwegingen, zoals het doel van de voorspelling en de begrijpelijkheid voor de gebruiker belangrijke overwegingen. Een hoger aggregatiemodel zou een makkelijker te begrijpen model kunnen opleveren en de communicatie over de modeluitkomsten kunnen bevorderen. Het helpt bovendien om hoofdzaken en details van elkaar te onderscheiden.²²

Daarnaast valt het te verwachten dat er een bepaalde samenhang is tussen de verschillende prestaties. Als er bijvoorbeeld substitutie plaatsvindt van lage naar hogere zorgzwaartepakketten, dan geeft de ontwikkeling van de lage zorgzwaartepakketten informatie over de te verwachten ontwikkeling van de hoge zorgzwaartepakketten. Dit soort informatie wordt nu niet benut. Als deze informatie wel benut wordt, bijvoorbeeld middels hiërarchische modellen, levert dat mogelijk betere voorspellingen op prestatieniveau op. Tegelijkertijd valt op dit moment bij het sommeren van de individuele prognoses op prestatieniveau een deel van de voorspelfouten op prestatieniveau tegen elkaar weg.

4.8 Zijn er nog niet gebruikte databronnen?

Het prognosemodel van de NZa gebruikt maandelijkse declaratiedata op prestatie- en Wlz-uitvoerdersniveau als belangrijkste input. De vraag is of het gebruik van andere of aanvullende data tot betere voorspellingen leidt. Daarbij zijn twee specifieke punten van belang: (1) zijn er data die aanvullende voorspellende informatie bevatten over het zorggebruik? En (2) zijn er data die sneller beschikbaar zijn en zo eerder inzicht geven in recente veranderingen in de zorgvraag dan de declaratiedata? In een verkenning van de beschikbare databronnen en in de gesprekken met de verschillende stakeholders zijn vier mogelijke databronnen naar voren gekomen: demografische/gezondheidsdata, indicaties, cliënten in zorg en wachtlijsten.

Demografische data

²² Dit zou een wijziging van de aanpak inhouden: de NZa raamt nu de volumes van iedere prestatie en hangt daar dan achteraf prijzen aan. Om verschillende prestaties bij elkaar op te tellen zouden prestaties eerst beprijsd moeten worden, en worden er feitelijk uitgaven in plaats van volumes geraamd.

Op de middellange- tot lange-termijn wordt de vraag naar Wlz-zorg voor een deel veroorzaakt door de leeftijdsontwikkeling van de bevolking (de vergrijzing) en de gezondheidsontwikkeling. Het CPB (Mot et al., 2016) baseert zijn ramingen van de zorguitgaven bijvoorbeeld deels op de bevolkingsontwikkeling per leeftijd, waarbij een correctie voor een met de levensverwachting meegroeiende verbetering van de gezondheid per leeftijd wordt toegepast. Voor de korte tijdsperiode waarop de prognose voor het Wlz-kader betrekking heeft (een raming van de uitgaven over het lopende jaar) heeft het toevoegen van demografische of epidemiologische gegevens echter weinig toegevoegde waarde: de veranderingen in deze variabelen zijn simpelweg te klein gedurende het jaar om van grote invloed te zijn en veranderingen worden waarschijnlijk grotendeels al opgepikt in de tijdstrends van de declaratiecijfers.

Indicatiegegevens

Er zit een logisch verband tussen het aantal afgegeven indicaties en de hoeveelheid gedeclareerde zorg: een indicatie geeft recht op zorg, dus valt te verwachten dat als het aantal afgegeven indicaties stijgt, ook de hoeveelheid geleverde zorg (uiteindelijk) stijgt. De samenhang is echter niet perfect: niet iedereen met een indicatie voor Wlz-zorg verzilverd die ook en de gedeclareerde kosten hangen af van de gekozen leveringsvorm. Meestal zit er ook tijd tussen het krijgen van een indicatie en het daadwerkelijk ontvangen van zorg: de indicatiegegevens lopen daarom voor op de declaratiegegevens. De relatie tussen de indicatiegegevens en de uiteindelijk gedeclareerde zorg en type leveringsvorm kan bovendien veranderen over de tijd (zie de discussie over doelmatigheid in hoofdstuk 3).

Wanneer de zorgvraag zich geleidelijk ontwikkelt maakt het verschil in timing tussen indicaties en declaraties weinig uit voor de prognoses, maar wanneer zich plotselinge veranderingen in de zorgvraag voordoen dan worden die eerder opgepikt in de indicatiegegevens dan in de declaraties. In de tweede helft van 2018 tot het einde van 2019 steeg het aantal indicaties bijvoorbeeld relatief snel ten opzichte van de trend in het voorgaande jaar. Deze toename was in eerste instantie nog niet te zien in de declaratiegegevens: zowel omdat deze data pas met vertraging beschikbaar komen, maar ook omdat het enige tijd duurt voordat alle nieuwe indicaties daadwerkelijk verzilverd zijn. In dit specifieke geval had het gebruik van (langjarige) indicatiegegevens tot een betere voorspelling geleid, maar dit is niet per se het geval in andere omstandigheden.

De NZa hanteert nu naast het prognosemodel op basis van declaratiegegevens ook een alternatief scenario op basis van de ontwikkeling in het aantal indicaties. Daarmee functioneren de indicatiegegevens als een kanarie in een kolenmijn: het indicatiescenario pikt schokken in de zorgvraag snel op en een grote discrepantie tussen de standaardprognose en het indicatiescenario kan reden zijn om de prognose of het advies aan te passen.

Het hanteren van een zijlicht is echter iets anders dan het direct gebruiken van indicatiegegevens als (extra) input om het aantal declaraties te voorspellen. Binnen het ETS model zal dit niet eenvoudig te doen zijn, maar het verder verkennen van de toegevoegde waarde van de indicatiegegevens in de standaardprognose zou wel met andere modellen kunnen worden onderzocht.

Aantal cliënten in zorg

Wanneer een cliënt bij een zorgaanbieder start met zorg gebruiken wordt dit door de zorgaanbieder geregistreerd. Terwijl zorgdeclaraties met enige vertraging worden ingediend en verwerkt, geeft het aantal cliënten in zorg een zeer recent beeld van de hoeveelheid zorg die op dit moment wordt afgenomen. Deze data wordt nu niet door de NZa benut. Deze informatie zou vooral een beter zicht kunnen geven op de na-ijl: hoeveel extra declaraties vallen er voor een bepaalde maand nog te verwachten? Data over cliënten in zorg zijn in ieder geval bekend bij de Wlz-uitvoerders. Wij kunnen niet inschatten of deze data (zonder aanvullende administratieve lasten) beschikbaar gemaakt kunnen worden voor de NZa.

Wachlijsten

Het aantal cliënten dat op een wachtlijst voor zorg staat geeft informatie over het verschil tussen de (geïndiceerde) vraag en het aanbod. Wanneer het aantal mensen op de wachtlijst oploopt dan is dit een signaal dat het aanbod ontoereikend zou kunnen zijn. Het is daarbij wel van belang om verschillende

soorten wachtenden te onderscheiden: maar voor een beperkt deel van de mensen die op de wachtlijst staan geldt dat ze urgent zorg nodig hebben maar er geen plek beschikbaar is. Voor een groter deel van de wachtenden geldt dat er wel een plek beschikbaar is, maar niet bij de aanbieder van voorkeur. Mensen wachten in zo'n geval op een plek bij de aanbieder die zij op het oog hebben, en gebruiken in de tussentijd vaak Wlz-zorg thuis.

De toegevoegde waarde van het gebruiken van wachtlijstdata in het prognosemodel lijkt beperkt. Het prognosemodel heeft als doel een zo 'n goed mogelijke inschatting te maken van de te verwachten gedeclareerde zorg. Voor zover onderschatting van de zorgvraag in de declaratiedata daarbij een probleem zou kunnen zijn, ligt het aanvullend gebruik van indicatiedata meer voor de hand dan het gebruik van wachtlijstdata. De wachtlijstdata geven vooral inzicht in meer structurele verschillen tussen het aanbod en de vraag, waarvoor aanvullend beleid nodig is. Zo heeft de NZa op verzoek van VWS in 2020 een inschatting gemaakt van hoeveel het beheersbaar houden van de wachtlijsten zou kosten, *bovenop* de basisprognose (NZa 2020).

4.9 Modeldocumentatie

Er is beperkte openbare documentatie over het prognosemodel beschikbaar. Deze documentatie is terug te vinden in (de appendices) van diverse adviezen (zie bijvoorbeeld NZa (2019c)). In opeenvolgende adviezen is min of meer hetzelfde document beschikbaar, dat steeds op een paar specifieke punten is geüpdatet of aangepast. De documentatie geeft een globale beschrijving van het model en de gebruikte data (vergelijkbaar met de informatie in dit stuk). Detailinformatie over het model ontbreekt, bijvoorbeeld over de gekozen specificaties voor de verschillende zorgprestaties, de modelprestatie (bijvoorbeeld de in-sample fit) en de modeluitkomsten (de voorspellingen op prestatie- en/of uitvoerniveau).

5 Conclusies

De NZa adviseert de Minister van VWS en aantal keer per jaar over de inzet van de herverdelingsmiddelen en of het Wlz-kader toereikend is om de verwachte uitgaven van de Wlz-uitvoerders te dekken. De basis voor dat advies is de basisprognose, die eventueel wordt bijgesteld op basis van expert opinion. Daarnaast presenteert de NZa in de adviezen alternatieve modellen. Om tot de basisprognose te komen wordt een ETS-model gebruikt.

Onze hoofdconclusie is dat de door de NZa gebruikte prognosemethode geschikt is voor het beoogde doel. Op basis van de beschikbare data, de kenmerken van het basismodel en eventuele alternatieven zijn er geen aanwijzingen dat een ander soort model tot structureel betere voorspellingen zou leiden. Wel zien wij een risico dat bij het leveren van één puntschatting of advies over de benutting van het Wlz-kader de afweging tussen toegankelijkheid en betaalbaarheid en doelmatigheid impliciet verschuift van de politiek naar de NZa. Wij raden de NZa daarom aan om meerdere scenario's voor de benutting van het Wlz-kader te blijven hanteren en daarbij expliciet aandacht te besteden aan de duiding van de verschillen in termen van de onderliggende doelstellingen. Indien het geven van één scenario toch gewenst is dan zou dit hand in hand moeten gaan met een nadere explicitering van de onderliggende afweging tussen toegankelijkheid en betaalbaarheid en doelmatigheid.

Hieronder werken we onze conclusies uit en doen we een aantal suggesties om de bruikbaarheid en onderbouwing van het prognosemodel verder te versterken. Dat doen we aan de hand van de zes vooraf door de NZa gestelde vragen. Tot slot formuleren we op basis van die conclusies een onderzoeksagenda.

Hoe functioneert het predictiemodel als beleidsinstrument?

Sluit de aanpak aan bij het doel van de prognose?

Het doel van de prognose en het advies is om een inschatting te geven van de benutting van het Wlz-kader en een advies uit te brengen over de inzet van de herverdelingsmiddelen. Hiervoor is een objectieve en nauwkeurige raming nodig. Een groot deel van het benodigde Wlz-kader valt objectief af te leiden uit het aantal cliënten dat reeds zorg ontvangt en de hoeveelheid openstaande indicaties. Voor de aanpassingen aan het kader waarover in de praktijk besloten wordt – hoeveel groeirimte is er extra nodig? – geldt dit echter niet.

Door de ambigue doelstelling van het Wlz-kader – enerzijds voorzien in de zorgvraag en anderzijds sturen op betaalbaarheid en doelmatigheid – zijn aanpassingen van het kader uiteindelijk deels een politieke afweging. Hoeveel budget er nodig is om extra zorgvraag op te vangen hangt af van de keuze van de politiek om via het kader op doelmatigheid te sturen (en bijvoorbeeld tijdelijk oplopende wachtlijsten te accepteren) en de mate waarin Wlz-uitvoerders en aanbieders het zorggebruik willen en kunnen sturen.

Die ambiguïteit in de doelstelling is niet weg te nemen via modelaanpassingen. Dat maakt het niet alleen lastig om de trend te voorspellen en om de voorspelling te beoordelen, maar betekent ook dat verschillende opvattingen over het belang van de doelen van de budgettering leiden tot verschillende opvattingen over de modeluitkomsten: iemand die vindt dat het Wlz-kader alleen ruim genoeg moet zijn om voor toegankelijke zorg te zorgen zal bijvoorbeeld de raming sneller te krap vinden.

Binnen de huidige inrichting van het systeem zijn de vaststelling van het Wlz-kader en de daaropvolgende beoordeling van de toereikendheid van dat kader twee grotendeels gescheiden werelden.²³ Bovendien worden betaalbaarheids- en doelmatigheidsoverwegingen die mee (kunnen) spelen in de vaststelling van het Wlz-kader door de politiek zelden expliciet gemaakt. Binnen deze context is de opdracht aan de NZa eigenlijk om de consequenties van (de verandering van) dat waardeoordeel zo goed mogelijk te benaderen. Dat doet de NZa door realisaties, die beïnvloed zijn door de beleidsafwegingen en het gedrag van uitvoerders en aanbieders in het verleden, door te trekken naar het lopende jaar.

Iedere raming is omgeven met onzekerheid. Er zitten zowel risico's aan het overschatten als het onderschatten van de benodigde middelen. Uiteindelijk is het niet aan de NZa maar aan de Minister om gegeven die onzekerheid een keuze te maken over de hoogte van het kader. De manier waarop de NZa in de meest recente adviezen omgaat met deze onzekerheid, door zowel de basisprognose als een ruimer scenario op basis van de ontwikkelingen van de zorgvraag te presenteren sluit aan bij dit uitgangspunt. De bruikbaarheid van dit ruimere scenario zou verder kunnen worden vergoed door de gehanteerde methode (type model en de observatieperiode) dichter op die van de basisprognose te laten aansluiten. Wanneer er toch gekozen wordt voor het geven van één puntschatting, dan zou dat gepaard moeten gaan met een in het advies en in de adviesaanvraag explicieter geformuleerde doelstelling.

De adviescyclus, en ook de relaties tussen het advies en het kader, is niet in formele procedures vastgelegd. De adviezen (en adviesverzoeken) kunnen dus verschillen in inhoud, timing en frequentie. Dit geeft flexibiliteit, maar zorgt ook voor risico's. De (impliciete) doelstelling onder de adviezen kan door de tijd verschillen, en veranderingen in de timing en frequentie van het advies hebben gevolgen voor de kwaliteit en stabiliteit van de prognoses en beïnvloeden de dynamiek in de besluitvorming.

Een deel van de onzekerheid in de prognose wordt veroorzaakt door een gebrek aan kennis over de meest doelmatige manier om aan de zorgvraag op een bepaald moment te voldoen en over het effect dat het Wlz-kader heeft op het gedrag van de Wlz-uitvoerders en zorgaanbieders. Mede vanwege die gedragsreacties is het inzicht in het effect van budgetveranderingen op de hoeveelheid geleverde zorg, het aantal afgegeven indicaties, de wachtlijsten en de doelmatigheid beperkt. De adviezen en de gebruikte modellen zijn gericht op voorspelling en niet op het verklaren van dit soort mechanismen. Aanvullend onderzoek, buiten het advies en de prognosemodellen om, naar doelmatigheidsverschillen tussen aanbieders of Wlz-uitvoerders en de effecten van budgettering op korte- en langetermijninvesteringen door aanbieders kan bijdragen aan een betere duiding van de ramingen en zo ook aan beter uiteindelijke beslissingen over het Wlz-kader.

Zijn de uitkomsten van de prognose bruikbaar en herkenbaar voor de gebruikers?

Uit onze inventarisatie komt naar voren dat de belangrijkste gebruikers van de prognoses (de Wlz-uitvoerders, de ministeries van VWS en Financiën) vooral geïnteresseerd zijn in de raming op macro- en op uitvoerder-niveau. Voor hen is vooral van belang hoe de prognoses zich verhouden tot ijkpunten zoals de uitgaven voor de cliënten in zorg, de groeiruimte in het Wlz-kader, trends in het aantal indicaties. In het geval van de Wlz-uitvoerders is ook de relatie tot de door henzelf gemaakte prognoses van belang.

Hoewel de hiervoor benodigde informatie veelal wel in de adviezen terug te vinden is, blijken de gebruikers vaak toch moeite te hebben met het duiden van de resultaten. Een meer consistente presentatie van de resultaten over de verschillende adviezen heen, gericht op de hoofdzaken, en met een sterkere nadruk op relatie tussen de prognose en herkenbare ijkpunten kan tot een beter hanteerbaar advies leiden. In deze presentatie zou ook de onzekerheid achter de raming, en door welke (beleidsmatige) ontwikkelingen deze

²³ Een vergelijking van het Wlz-kader met de trend die voorspeld wordt op basis van het basismodel geven ook/vooral inzicht in de afwijking van het kader zoals vastgesteld door de Minister van VWS ten opzichte van de trend in voorgaande jaren. Prognoses over de toereikendheid van het Wlz-kader gedurende het jaar die de NZa doet op basis van die vergelijking doen dus mogelijk automatisch een groot deel van de eventuele veranderingen die de Minister van VWS via het kader wilde doorvoeren teniet.

wordt gedreven, een grotere aandacht kunnen krijgen. De infographics bij de economische ramingen van het Centraal Planbureau zijn een goed voorbeeld van hoe dit zou kunnen (zie bijvoorbeeld CPB 2021).

Modeldocumentatie

Open-access, zowel van de gebruikte data als modellen, is een steeds belangrijker thema geworden. Idealiter zijn door de overheid gehanteerde modellen en data zo veel mogelijk toegankelijk, zodat eenieder die dat wil de resultaten kan controleren. Op zijn minst is een grondige en recente documentatie nodig die de uitgangspunten en belangrijkste keuzes in het model beschrijft. Op dit moment is documentatie over het prognosemodel in beperkte mate beschikbaar in de appendix van verschillende adviezen (bijvoorbeeld in de bijlage bij het augustusadvies uit 2019 (NZa 2019a)). Het verdient de aanbeveling om een geüpdatete en uitgebreidere versie van de modeldocumentatie te maken en deze te plaatsen in een los vindbaar document of een website, zodat anderen altijd de meeste recente versie kunnen vinden.

Hoe goed voorspelt het huidige predictiemodel?

Een gestructureerde evaluatie van de voorspelkwaliteit van het prognosemodel ontbreekt. Ondanks dat de uitkomsten van het model zelf invloed hebben op de werkelijkheid die het probeert te voorspellen, zou zo'n gestructureerde analyse wenselijk zijn. In 2018 en 2019 steeg de zorgvraag in de Wlz relatief snel, en bleek het prognosemodel deze vraag aanvankelijk te onderschatten. Hieruit zou wellicht de indruk kunnen ontstaan dat een ander model, of andere data, tot betere voorspellingen zou leiden. Het is de vraag in hoeverre de specifieke omstandigheden in deze periode iets zeggen over de *structurele prestaties* van het model.

Op basis van de kenmerken van het basismodel en de beschikbare alternatieven is het niet evident dat een andere keuze tot betere prognoses zou leiden. Uiteindelijk is dit echter een empirische vraag, waarbij alle adviesmomenten over verschillende jaren heen moeten worden meegenomen.

Bij het beoordelen van de voorspelkwaliteit is het van belang dat kleine relatieve afwijkingen in de prognose relatief grote budgettaire consequenties kunnen hebben, omdat het om zeer grote totaalbedragen gaat. Het totale Wlz-kader is voor 2022 meer dan 28 miljard euro (Ministerie van VWS, 2021). Dat maakt dat in het beleid en de politiek, waar men vooral geïnteresseerd zal zijn in de afwijking van de prognose in absolute euro's, al snel de indruk zou kunnen ontstaan dat het model niet goed presteert. Het sterker onderbouwen van de voorspelkwaliteiten van het huidige model kan dan bijdragen aan het communiceren van het feit dat het doen van prognoses altijd omgeven is met een bepaalde mate van onzekerheid.

Hoe gaat het predictiemodel om met de beschikbare data? En worden de huidige inputs optimaal gebruikt?

Voor het basismodel wordt de hoeveelheid geleverde zorg apart geraamd per prestatie en uitvoerder en daarvoor wordt alleen de hoeveelheid gedeclareerde zorg uit het verleden gebruikt. Deze volumeraming wordt vervolgens vermenigvuldigd met de ingeschatte prijs. Voor de volumeraming wordt een ETS-model gebruikt. Hoewel er alternatieven zijn, zoals ARIMA-modellen en datasciencemodellen, is er geen reden om aan te nemen dat die met dezelfde input beter zullen voorspellen. Wel is het raadzaam om deze premisse expliciet te testen door in de analyse van de voorspelkwaliteit van het huidige model ter vergelijking ook alternatieve modellen op te nemen.

De prognoses worden gemaakt op het laagst mogelijke niveau (per prestatie en uitvoerder). Een hoger aggregatieniveau leidt niet per se tot een hogere voorspelkwaliteit, maar kan wel een makkelijker te interpreteren en hanteerbaarder model opleveren. Zeker gezien het feit dat de meeste gebruikers van het model weinig of geen gebruik maken van de prognoses op prestatieniveau, valt het maken van prognoses op een hoger aggregatieniveau te overwegen.

Kan de prognose worden verbeterd door gebruik te maken van aanvullende databronnen? En hoe goed gaat het model om met plotse veranderingen?

De basisprognose maakt alleen gebruik van declaratiedata. Plotse veranderingen in de zorgvraag werken met enige vertraging door in de prognose: eerst moet een verandering in de zorgvraag (het aantal afgegeven indicaties) zich ook vertalen in een toename van het daadwerkelijke (gedeclareerde) gebruik, en vervolgens komt de informatie over de declaraties zelf met vijf maanden vertraging terecht in het model. Er is daarom mogelijk een rol voor data die sneller beschikbaar is, zoals bijvoorbeeld data over het aantal cliënten in zorg, of voor leading indicators: andere variabelen die voorlopen op veranderingen in het gebruik.

Uit deze inventarisatie komen twee mogelijke aanvullende databronnen naar voren: het aantal afgegeven indicaties en het aantal cliënten in zorg. Het aantal indicaties loopt voor op de gebruikte gedeclareerde zorg en kan op dus eerder veranderingen in de zorgvraag signaleren. De informatie over het aantal afgegeven indicaties wordt op dit moment al gebruikt door de NZa in een van de alternatieve scenario's. Op die manier kunnen eventuele plotselinge veranderingen in de zorgvraag, die nog niet door de declaratiedata worden opgepikt, nu al worden meegewogen in het advies. De indicatiedata hebben daarnaast mogelijk toegevoegde waarde en kunnen in aanvulling op de volumegegevens gebruikt worden binnen het basis-prognosemodel.

Het aantal cliënten in zorg geeft de meeste recente informatie over het zorggebruik. Op dit moment wordt deze informatie nog niet benut. Wij kunnen niet beoordelen of deze informatie eenvoudig door de NZa kan worden ontsloten. De data zouden vooral bruikbaar zijn om een beter inzicht te krijgen in de na-ijl (hoeveel nagekomen declaraties zijn er nog te verwachten?).

Naast aanvullende databronnen kan ook de rol van expert opinion versterkt worden. Bij een grote (verwachte) discrepantie tussen recente observaties en de voorspelde trends zal de modelleur op basis van expert opinion zelf moeten inschatten of er sprake is van een structurele verandering. In het geval van expert opinion is een gestructureerde aanpak, inclusief documentatie en onderbouwing, van belang voor de transparantie en reproduceerbaarheid van het advies. De rol van expert opinion is naar verwachting niet alleen waardevol bij het inschatten van incidentele schokken als gevolg van beleidswijzigingen zoals de overheveling van de langdurige GGZ naar de Wlz of externe factoren zoals de COVID-19 pandemie, maar ook bij het inschatten van de effecten van terugkerende vragen zoals het inschatten van het volume later ontvangen en verwerkte declaraties van aanbieders.

Onderzoeksagenda

Op basis van bovenstaande bevindingen komen we tot de volgende onderzoeksagenda voor de NZa:

1. Heroverweeg het aggregatieniveau van het prognosemodel.
2. Evalueer de voorspelkwaliteit van de basisprognose over meerdere jaren en vergelijk die met die van concurrerende modellen.
3. Maak de drie scenario's – de basisprognose, het nul-scenario en het vraag-gebaseerde scenario – onderling beter vergelijkbaar, bijvoorbeeld op het type schatting, de lengte van de observatieperiode en het aggregatieniveau.
4. Doe vervolgonderzoek naar de toegevoegde waarde van het gebruik van data over het aantal afgegeven indicaties en over het aantal cliënten in zorg in het prognosemodel.
5. Doe, los van het advies, onderzoek naar de (on)mogelijkheden van doelmatigheidsverbeteringen middels prikkels. Onderzoek daarbij bestaande doelmatigheidsverschillen op aanbieder- en uitvoerdersniveau. En onderzoek daarnaast de gedragsreacties van aanbieders en uitvoerders op aanpassingen aan het Wlz-kader.

6. Onderzoek, samen met de belangrijkste gebruikers, hoe de prognoses op zo'n manier gepresenteerd kunnen worden dat ze de meest bruikbare informatie opleveren. Neem in dit onderzoek ook de manier waarop de belangrijkste risico's en onzekerheden onder de raming kunnen worden gecommuniceerd mee.
7. Onderzoek hoe de bijstelling op basis van expert opinion gestructureerd kan worden en geef dit onderdeel van het proces een duidelijk herkenbare plek in het advies.

Literatuur

Bakx P, Schut E, Wouterse B. (2020). Price setting in long-term care in the Netherlands. *Erasmus School of Health Policy & Management Research Report 2020.011*. ESHPM, Rotterdam.

Barker T, Pesaran MH. 2011. *Disaggregation in econometric modelling*. London: Routledge.

CPB (2020). *Zorgkeuzes in Kaart*. CPB, Den Haag.

CPB (2021). www.cpb.nl/sites/default/files/omnidownload/CPB-Infographic-Juniraming-2021.pdf

Don, H. (2007). Ambitie en voorzichtigheid in het economisch beleid. *TPE Digitaal*. 1(1): 67-94.

Hussem A, Bakx P, Tenand M. 2020. Publieke middelen voor particuliere ouderenzorg. *Netspar Design Paper 143*. Netspar, Tilburg.

Ministerie van Financiën. (2020). Naar een toekomstbestendig zorgstelsel. *Brede Maatschappelijke Heroverwegingen 2*. Ministerie van Financiën, Den Haag.

Ministerie van VWS. (2020). Definitieve kaderbrief Wlz 2021. 8 oktober 2020. Kenmerk: 1749665-210825-LZ.

Ministerie van VWS. (2021). Voorlopige kaderbrief Wlz 2022. 2 juli 2021. Kenmerk: 3215792-1011616-LZ

Mot, E., Stuu, K., Westra, P., & Aalbers, R. (2016). Een raming van de zorguitgaven 2018-2021. *CPB Achtergronddocument*.

NZa. (2019a). Informatiekaart beleid en regels voor de langdurige zorg 2019. puc.overheid.nl/nza/doc/PUC_245618_22/1/

NZa. (2019b). Beleidsregel prestatiebeschrijving en tarief zzp-meerzorg Wlz - BR/REG-19123. puc.overheid.nl/doc/PUC_239127_22/1/#b51eb868-35cc-480c-9b0b-9dabe1a30afb

NZa. (2019c). Augustus-advies benutting budgettair kader Wlz 2019. puc.overheid.nl/nza/doc/PUC_289712_22/1/

NZa. (2020). Maartbrief benutting budgettair kader Wlz. puc.overheid.nl/nza/doc/PUC_305276_22/1/

NZa. (2021a). Meibrief benutting budgettair kader Wlz 2021. puc.overheid.nl/nza/doc/PUC_642299_22/1/

NZa. (2021b). Monitor Ouderenzorg - regionale verschillen (deel 1). puc.overheid.nl/nza/doc/PUC_636235_22/1/

Otexts. (2021). ETS forecasting. <https://otexts.com/fpp3/ets-forecasting.html>.

Portrait, F. R. M., & Koolman, X. (2021). Is ageing-in-place an alternative for nursing home admission? A comparison of the survival, days in care and costs of older individuals.

Rijksoverheid. (2021). www.rijksoverheid.nl/onderwerpen/verpleeghuizen-en-zorginstellingen/wet-langdurige-zorg-wlz

Skipr. (2020). www.skipr.nl/nieuws/rechter-veegt-inkoopbeleid-langdurige-zorg-van-tafel/

Skipr. (2021). www.skipr.nl/nieuws/wlz-inkoopkader-weer-voor-de-rechter/

Tweede Kamer. (2013/2014). Regels inzake de verzekering van zorg aan mensen die zijn aangewezen op langdurige zorg (Wet langdurige zorg). Kamerstuk 33891, nr. 3

Erasmus University Rotterdam

Erasmus School of Health Policy & Management

Bayle Building

Burgemeester Oudlaan 50

3062 PA Rotterdam, The Netherlands

T +31 10 408 8555

E communicatie@eshpm.eur.nl

W www.eur.nl/eshpm