
PROEFTUIN ONDERZOEK 'METEN IS WETEN'

Eindrapportage

Onderzoek uitgevoerd door:

Prof. dr. M. Lindeboom
Prof. dr. B. van der Klaauw
S. Vriend MSc

Vrije Universiteit Amsterdam
Faculteit der Economische Wetenschappen en Bedrijfskunde
Afdeling Algemene Economie

In opdracht van:

Centrum Indicatiestelling Zorg

29 november 2013

MANAGEMENTSAMENVATTING

Voor u ligt de eindrapportage over het Proeftuin onderzoek ‘Meten is Weten’. In het kader van dit onderzoek is een veldexperiment uitgevoerd om de effecten van verschillende vormen van toetsingspraktijk op het aanvraaggedrag van zorgaanbieders te bestuderen. Het gaat hierbij om aanvragen voor AWBZ zorg, in het bijzonder om herindicaties via taakmandaat (HiTs) en standaard indicatieprotocollen (SIPs). Er is gevarieerd met twee beleidsinstrumenten in de toetsingspraktijk, te weten het toetsingsmoment en het toetsingspercentage.

HiT aanvragen worden normaliter getoetst voordat aangevraagde zorg door zorgaanbieders kan worden ingezet. In het veldexperiment onderzoeken we de gevolgen van achteraf toetsing, wat inhoudt dat zorgaanbieders niet hoeven te wachten met het inzetten van zorg. Bij achteraf toetsing is er geen mogelijkheid tot correctie wanneer een aanvraag niet conform wordt bevonden. Daarom is een tweede variant onderzocht waarbij het moment van toetsing afhankelijk is van de resultaten van recente toetsingen van een zorgaanbieder. In dit regime worden zorgaanbieders die in de regel veel recente conform toetsingen hebben, vooraf getoetst, terwijl zorgaanbieders met relatief veel niet conform toetsingen achteraf worden getoetst en moeten wachten voordat zorg kan worden ingezet.

De belangrijkste conclusies van dit deel van het veldexperiment zijn:

1. Achteraf toetsing leidt tot meer HiT aanvragen. De omvang van het effect is groot: we vinden een toename van 31% in het aantal ingediende HiT aanvragen (in 2013). De toename in het aantal HiT aanvragen kan echter niet (geheel) worden verklaard uit substitutie met andere type aanvragen. Daarmee heeft dit directe gevolgen voor de grote geldstroom (AWBZ uitgaven). De effecten op het aantal aanvragen zijn het grootst voor zorgaanbieders met een initieel klein aandeel HiT aanvragen in het totale pakket van aanvragen.
2. Ten opzichte van vooraf toetsing is er geen verschil in het aanvraaggedrag van zorgaanbieders waarbij het toetsingsmoment afhangt van uitkomsten van recente toetsingen. Steeds meer zorgaanbieders in de groep met een prestatieafhankelijk toetsingsmoment presteren gedurende het experiment boven de gestelde norm 92% conform en worden daarom achteraf getoetst. Dit geeft zorgaanbieders die het recent goed hebben gedaan de mogelijkheid om zorg snel in te zetten, zonder dat dit een effect heeft op de grote geldstroom (AWBZ uitgaven).

SIP aanvragen worden altijd achteraf getoetst, er is daarom geen mogelijkheid voor correctie van niet conform getoetste aanvragen. Binnen het veldexperiment is het toetsingspercentage voor deze aanvragen aangepast. We maken onderscheid tussen zorgaanbieders die geconfronteerd worden met een laag toetsingspercentage, een hoog toetsingspercentage en een toetsingspercentage dat afhankelijk is van de resultaten van recente toetsingen.

De belangrijkste conclusie van dit gedeelte van het veldexperiment is dat variatie in het toetsingspercentage geen effect heeft op aanvraaggedrag van zorgaanbieders.

Als we beide experimenten gezamenlijk bekijken lijkt de dimensie vooraf/achteraf belangrijk te zijn. Alleen achteraf toetsen is niet beperkend, maar de combinatie met vooraf toetsing bij ondermaatse prestaties is wel zinvol. Zorgaanbieders die vaak conform getoetst worden, worden achteraf getoetst. Daarom kunnen zij snel zorg inzetten, maar wanneer dat nodig blijkt kan worden overgegaan op vooraf toetsing. Deze prikkel is belangrijk en voldoende.

Inhoudsopgave

MANAGEMENTSAMENVATTING	ii
1 Inleiding	1
2 Onderzoeksopzet	3
2.1 Het experiment	4
2.1.1 Beleidsinstrument 1: het moment van toetsing	4
2.1.2 Beleidsinstrument 2: het toetsingspercentage	5
2.2 De selectie van participerende zorgaanbieders	6
3 Verloop van het onderzoek	9
4 Data	11
4.1 Beschrijving van de beschikbare gegevens	11
4.2 Beschrijvende statistieken van de data	12
4.2.1 Kenmerken van zorgaanbieders	12
4.2.2 Inhoudelijke kenmerken van de zorgaanvragen	14
4.2.3 Aantallen aanvragen, toetsingen en fractie conform	16
5 Beschrijvende analyses	19
5.1 Periodieke aanpassingen prestatieafhankelijke toetsingspraktijk	19
5.1.1 Aanpassingen in het toetsingsmoment (HiT prestatie groep)	19
5.1.2 Aanpassingen in het toetsingspercentage (SIP prestatie groep)	21
5.2 Effecten van variatie in toetsingspraktijk: grafische analyses	23
5.2.1 Aantal aanvragen	24
5.2.2 Aantal toetsingen	26
5.2.3 Fractie conform getoetste aanvragen	27
6 Regressieanalyses	30
6.1 Basisanalyses	30
6.1.1 Aantal aanvragen	31
6.1.2 Aantal toetsingen	34
6.1.3 Fractie conform getoetste aanvragen	36
6.1.4 Substitutie-effecten	37
6.2 Heterogene effecten	39
6.2.1 Grootte van de zorgaanbieder in termen van aantal aanvragen	39
6.2.2 Aanvraagmix van de zorgaanbieder	44
6.2.3 Ziekenhuizen ten opzichte van ander soort zorgaanbieders	46
6.3 Robuustheidsanalyses	48

7 Conclusies	53
Appendices	55
A Simulatiestudie looptijd experiment	55
B Indeling van zorgaanbieders over de experimentele groepen	57
C HiT typen en SIP protocolnummers	64
D Gedetailleerde beschrijving van de dataset	66
E Toepassing van selectiecriteria op de dataset	68
F Fractie conform per HiT type	70
G Aanvullende regressieresultaten	73

Lijst van figuren

5.1	Gemiddeld aantal HiT aanvragen, per HiT groep.	24
5.2	Gemiddeld aantal SIP aanvragen, per SIP groep.	25
5.3	Gemiddeld aantal HiT toetsingen, per HiT groep.	26
5.4	Gemiddeld aantal SIP toetsingen, per SIP groep.	27
5.5	Gemiddelde fractie conform getoetste aanvragen, per HiT groep.	28
5.6	Gemiddelde fractie conform getoetste aanvragen, per SIP groep.	29
F.1	Fractie conform getoetste HiT type 1 aanvragen, per HiT groep.	71
F.2	Fractie conform getoetste HiT type 2 aanvragen, per HiT groep.	71
F.3	Fractie conform getoetste HiT type 4 aanvragen, per HiT groep.	72
F.4	Fractie conform getoetste overige HiT aanvragen (type 3 en 6), per HiT groep.	72

Lijst van tabellen

2.1	Beschrijvende statistieken over de periode voorafgaand aan het experiment.	8
3.1	Tijdspad periodieke aanpassingen in toetsingspraktijk.	9
4.1	Aantallen aanvragen en toetsingen, 1 januari 2012 tot en met 7 april 2013.	12
4.2	Kenmerken van zorgaanbieders.	13
4.3	Zorginhoudelijke kenmerken van aanvragen en HiT toetsingen, HiT groepen.	15
4.4	Zorginhoudelijke kenmerken van aanvragen, SIP groepen.	16
4.5	Aanvragen, toetsingen en fractie conform getoetst.	17
5.1	Periodieke aanpassingen in het toetsingsmoment.	20
5.2	Aantal wijzigingen in toetsingsmoment.	21
5.3	Aantal keer toetsingsmoment vooraf.	22
5.4	Periodieke aanpassingen in toetsingspercentage.	22
5.5	Variatie in het toetsingspercentage.	23
6.1	Schattingsresultaten voor het effect van variatie in de toetsingspraktijk op het aantal aanvragen.	32
6.2	Schattingsresultaten voor het effect van variatie in de toetsingspraktijk op het aantal aanvragen; uitsplitsing van de experiment periode.	34
6.3	Schattingsresultaten voor het effect van variatie in de toetsingspraktijk op het aantal toetsingen.	35
6.4	Schattingsresultaten voor het effect van variatie in de toetsingspraktijk op het aantal toetsingen; uitsplitsing van de experiment periode.	36
6.5	Schattingsresultaten voor het effect van variatie in de toetsingspraktijk op de fractie conform getoetste aanvragen.	37
6.6	Schattingsresultaten voor het effect van variatie in de toetsingspraktijk op het aantal indicatiemeldingen, spoedaanvragen en reguliere AWBZ aanvragen.	38
6.7	Heterogene effecten: kleine versus grote zorgaanbieders; aantal aanvragen.	41
6.8	Heterogene effecten: kleine versus grote zorgaanbieders; aantal toetsingen.	42
6.9	Heterogene effecten: kleine versus grote zorgaanbieders; conformfractie.	43
6.10	Heterogene effecten: aanvraagmix.	45
6.11	Heterogene effecten: ziekenhuis versus ander soort zorgaanbieders.	47
6.12	Robuustheidsanalyses: verwijderen van inactieve zorgaanbieders.	49
6.13	Robuustheidsanalyses: polynoom in tijd.	51
6.14	Robuustheidsanalyses: kwartaaldummies.	51
6.15	Robuustheidsanalyses: aggregatie per 4 weken.	52
A.1	Resultaten van de simulatiestudie voor het HiT gedeelte van het onderzoek.	55
A.2	Resultaten van de simulatiestudie voor het SIP gedeelte van het onderzoek.	56
B.1	Indeling van zorgaanbieders in HiT groepen.	57
B.2	Indeling van zorgaanbieders in SIP groepen.	60
C.1	Overzicht HiT typen en SIP protocolnummers.	65

E.1	Selectiecriteria en verlies in aantal aanvragen.	69
G.1	Robuustheidsanalyses: polynoom in tijd.	73
G.2	Robuustheidsanalyses: kwartaaldummies.	74

1 Inleiding

Net als veel andere landen wordt ook Nederland geconfronteerd met een vergrijzende bevolking en daarmee gepaard gaande stijgende uitgaven voor vanuit de AWBZ gefinancierde, lange termijn zorg. Deze AWBZ uitgaven bedroegen in 2011 25,3 miljard euro, wat gelijk staat aan 4,2% van het BBP. Dit is een toename van 36% ten opzichte van het niveau in 2002.¹ Deze ontwikkelingen roepen de vraag op hoe de zorguitgaven in de toekomst in de hand gehouden kunnen worden.

Sinds 2005 functioneert het Centrum Indicatiestelling Zorg (CIZ) als poortwachter voor toegang tot AWBZ-gefinancierde lange termijn zorg. Jaarlijks verwerkt het CIZ ongeveer één miljoen aanvragen voor lange termijn zorg. De operationele kosten die gemoeid zijn met het verwerken van deze aanvragen, bedragen rond de 100 miljoen euro. De aanvragen worden meestal ingediend door zorgaanbieders.² Een willekeurige steekproef wordt door het CIZ getoetst op conformiteit met regels en richtlijnen en de benodigde zorg voor de cliënt. Een toetsing kan bijvoorbeeld leiden tot bijstelling van de aangevraagde zorgomvang en/of het type zorg. Aan iedere getoetste aanvraag wordt een label ‘conform’ of ‘niet conform’ gegeven.

Gezien de beperkte middelen beschikbaar voor het uitvoeren van toetsingen bij het CIZ, is het interessant om te bekijken hoe de kleine geldstroom (de kosten die gemoeid zijn met het verwerken van aanvragen en het uitvoeren van toetsingen) de grote geldstroom (de AWBZ uitgaven) kan beïnvloeden. Het potentiële effect van toetsingspraktijk op de AWBZ uitgaven is tweeledig. Enerzijds kan sprake zijn van een direct effect dat voortkomt uit daadwerkelijk uitgevoerde toetsingen: bijstelling van aangevraagde zorg kan de AWBZ uitgaven veranderen. Anders gezegd: meer toetsen betekent dat er meer mogelijkheden zijn voor aanpassingen en daarmee kunnen de AWBZ uitgaven direct worden beïnvloed.

Anderzijds kan sprake zijn van een indirect effect, via de inrichting van de toetsingspraktijk: de mogelijkheden voor inhoudelijke correctie van zorgaanvragen (toetsingsmoment) en de kans waarmee een bepaalde aanvraag wordt getoetst (toetsingspercentage), kunnen een effect hebben op het aanvraaggedrag van zorgaanbieders en daarmee op de AWBZ uitgaven. De vraag is dan welke prikkels in de toetsingspraktijk positieve effecten hebben op bijvoorbeeld het percentage aanvragen dat conform wordt getoetst, dus op de zorgvuldigheid van zorgaanbieders in het indienen van aanvragen. Wanneer de kans op toetsing bijvoorbeeld erg groot is, kan het zijn dat zorgaanbieders extra nauwkeurig worden bij het indienen van aanvragen om zo correcties te voorkomen.

Om de effecten van verschillende vormen van toetsingspraktijk te analyseren, is over de

¹Gebaseerd op cijfers van CBS StatLine (juli 2013). De cijfers over 2011 zijn voorlopig.

²Het is ook mogelijk dat de cliënt zelf een aanvraag voor AWBZ zorg indient. Echter, in ongeveer 85% van de gevallen is het de zorgaanbieder die de aanvraag verzorgt.

periode van 17 september 2012 tot en met 7 april 2013 een veldexperiment uitgevoerd. In dit experiment is gevarieerd met twee beleidsinstrumenten: het moment van toetsing, wat de mogelijkheden voor inhoudelijke correctie van zorgaanvragen bepaalt, en het toetsingspercentage. Zowel exogene als endogene (gebaseerd op eerdere prestaties) variatie in deze beleidsinstrumenten is daarbij bekeken. Het onderzoek en de resultaten worden in dit rapport beschreven. In de volgende sectie wordt de opzet van het onderzoek besproken. Het verloop van het onderzoek is beschreven in sectie 3. De dataset wordt beschreven in sectie 4. Sectie 5 geeft een beschrijvende analyse van (de gevolgen van) de variatie in de toetsingspraktijk. In sectie 6 worden vervolgens regressieanalyses gepresenteerd en worden de resultaten van het onderzoek weergegeven. Tenslotte kunnen de conclusies in sectie 7 worden gevonden.

2 Onderzoeksopzet

Het onderzoek heeft als doel het effect van (variatie in) het moment van toetsing en de hoogte van het toetsingspercentage op het aanvraaggedrag van zorgaanbieders te bepalen. Aanvraaggedrag heeft hierin twee dimensies: het aantal ingediende aanvragen en de fractie conform beoordeelde aanvragen. Er kunnen diverse soorten aanvragen door zorgaanbieders worden ingediend:

- standaard indicatieprotocollen (SIPs);
- herindicaties via taakmandaat (HiTs);
- reguliere aanvragen;
- spoedaanvragen;
- indicatiemeldingen.

In dit onderzoek richten wij ons op HiT en SIP aanvragen. Van de 1.000.000 aanvragen die per jaar worden ingediend, zijn ongeveer 50.000 HiT aanvragen en 340.000 SIP aanvragen. HiT en SIP aanvragen worden direct aan de ‘voorkant’ afgehandeld. Deze aanvragen worden steekproefsgewijs getoetst door de afdeling Externe Toetsing.³ SIP aanvragen worden achteraf getoetst; er vindt geen correctie plaats als een SIP niet conform wordt geacht, alhoewel wel terugkoppeling plaatsvindt door Relatiebeheer. HiT aanvragen worden vooraf getoetst. In geval van een niet conform toetsing kan de geïndiceerde zorg afwijken van de aangevraagde zorg. Zo’n aanpassing kan bijvoorbeeld in termen van hoeveelheid zorg (klasse of zorgzwaartepakket) of soort zorg (functies) zijn. Het is ook mogelijk dat de aanvraag volledig wordt afgekeurd en een nieuwe aanvraag moet worden ingediend.

In de volgende sectie bespreken we de opzet van het experiment. Sectie 2.2 beschrijft de manier waarop participerende zorgaanbieders zijn geselecteerd. Om vast te stellen hoe lang het experiment dient te lopen om statistisch gegronde conclusies te kunnen trekken, hebben we voor aanvang van het experiment een simulatiestudie uitgevoerd, een zogenaamde *power*-analyse. In zo’n simulatiestudie wordt gekeken hoe groot de kans is dat bij een bepaalde steekproefomvang een gegeven effect van de wijziging in toetsingspraktijk gevonden wordt. De resultaten laten zien dat, met de gekozen looptijd van zes maanden, effecten van 5%-punt met voldoende statistische betrouwbaarheid kunnen worden gevonden. De resultaten, zoals ook gerapporteerd in de tussenrapportage die op 18 januari 2013 is opgeleverd, zijn beschreven in Appendix A.

³Indicatiemeldingen, die betrekking hebben op intramurale zorg voor cliënten van 80 jaar of ouder, worden direct omgezet in een indicatiebesluit, zonder dat daar een inhoudelijke beoordeling of toetsing aan vooraf gaat. De betreffende zorg kan daarmee direct worden ingezet. De reguliere en spoedaanvragen worden door indicatiestellers behandeld en vertaald in een indicatiebesluit. Dit kan worden gezien als een 100% toetsing vooraf.

2.1 Het experiment

Het experiment bestaat uit twee onderdelen. In het eerste deel variëren we het toetsingsmoment voor HiT aanvragen. In het tweede gedeelte van het experiment wordt geëxperimenteerd met het toetsingspercentage voor SIP aanvragen.

2.1.1 Beleidsinstrument 1: het moment van toetsing

Met het eerste onderdeel van het experiment onderzoeken we het effect van de mate waarin correctie van de zorgaanvraag mogelijk is op het aanvraaggedrag van zorgaanbieders. Er zijn twee mogelijkheden voor het moment van toetsing: (1) *vooraf*, dat wil zeggen voordat de aangevraagde zorg kan worden ingezet door de zorgaanbieder, of (2) *achteraf*, dat wil zeggen nadat de aangevraagde zorg is geëffectueerd. Wanneer sprake is van vooraf toetsing, kan de uiteindelijk toegekende zorg afwijken van de aangevraagde zorg, bijvoorbeeld in termen van type zorg en zorgomvang. Als een aanvraag vooraf wordt geselecteerd voor toetsing, betekent dit dat de zorgaanbieder maximaal twee weken, en in de regel één week, moet wachten voordat een indicatiebesluit is afgegeven en de zorg kan worden ingezet. Bij achteraf toetsing kan de aangevraagde zorg direct worden ingezet. Een toetsing kan dan evengoed nog leiden tot een niet conform beoordeling van de aanvraag, maar dit heeft geen (financiële) gevolgen voor de zorgaanbieder of de inzetbare zorg. Een negatieve toetsing wordt wel gecommuniceerd naar en besproken met de betreffende zorgaanbieder. Voor de zorgaanbieder heeft het als voordeel dat de zorg sneller kan worden ingezet.

Het eerste onderdeel van het experiment is als volgt opgezet. Zorgaanbieders geselecteerd voor participatie in dit deel van het experiment worden willekeurig verdeeld over drie groepen. Vóór de start van het experiment werd een willekeurige 25%-steekproef van alle HiT aanvragen vooraf getoetst. In het experiment maken we daarnaast ook een groep waarin achteraf wordt getoetst en een groep waarin het toetsingsmoment afhankelijk is van prestatie. In deze laatste groep wordt het moment van toetsing als volgt bepaald: indien het percentage aanvragen dat in de afgelopen 8 weken conform is getoetst groter is dan 92% wordt achteraf getoetst en anders wordt er vooraf getoetst. Deze doelstelling van 92% conform is in overleg met het CIZ vastgesteld. Aanpassingen van het toetsingsmoment worden iedere 2 weken berekend en doorgevoerd. Alle zorgaanbieders beginnen aan de start van het experiment op vooraf toetsing. Het prestatieafhankelijke systeem corrigeert daar waar nodig: de goed presterende zorgaanbieders ondervinden geen administratieve vertraging in het inzetten van zorg en de inzet van zorg wordt alleen vertraagd indien dat nodig is. In al deze groepen wordt hetzelfde toetsingspercentage gehanteerd als voor de start van het experiment.⁴ De drie groepen zijn dus:

- HiT controle: vooraf toetsing (controlegroep);
- HiT achteraf: achteraf toetsing;

⁴Dit toetsingspercentage is afhankelijk van het type HiT aanvraag dat het betreft. Bovendien is voor de start van het experiment voor een aantal zorgaanbieders het toetsingspercentage verlaagd vanwege goede prestaties. Alle toetsingspercentages liggen echter tussen 20% en 25%.

- HiT prestatie: moment van toetsing afhankelijk van prestatie.

Er kan een aantal hypothesen over de effecten van de beschreven vormen van variatie in toetsingspraktijk worden geformuleerd. Enerzijds kan men verwachten dat de overgang op achteraf toetsing negatieve effecten heeft op het conformpercentage of het aantal ingediende aanvragen. Immers, het ontbreken van de mogelijkheid tot correctie kan zorgaanbieders prikkelen om meer aanvragen in te dienen of minder nauwkeurig te zijn bij het indienen. Anderzijds kan het overgaan op achteraf toetsing worden gezien als een teken van vertrouwen. Indien sprake is van een succesvolle toetsingspraktijk gebaseerd op vertrouwen, zou men geen negatieve effecten verwachten van het verwijderen van de mogelijkheid tot bijstelling op aanvraagdgedrag, of zelfs positieve effecten. In het systeem van een prestatieafhankelijk moment van toetsing zijn andere prikkels actief. Enerzijds zijn er directe prikkels om het goed te doen, aangezien dat leidt tot achteraf toetsing. Daarnaast kan het behoud van vertrouwen wanneer zorgaanbieders boven de gestelde 92%-norm blijven en het opnieuw invoeren van bijstellingsmogelijkheden bij prestaties beneden de norm, eveneens een stimulans geven voor goede prestaties.

2.1.2 Beleidsinstrument 2: het toetsingspercentage

In het tweede onderdeel van het experiment bestuderen we het effect van de kans op toetsing op het aanvraagdgedrag van zorgaanbieders. We kijken daarbij naar het effect van het verlagen van het toetsingspercentage, het verhogen van het toetsingspercentage, en het endogeen (prestatieafhankelijk) aanpassen van het toetsingspercentage.

In dit onderdeel van het onderzoek kijken we naar SIP aanvragen. Er zijn diverse soorten SIPs die kunnen worden aangevraagd. Elke soort SIP wordt aangemerkt als laag risico, midden risico of hoog risico.⁵ De risicocategorie bepaalt het gehanteerde toetsingspercentage. Voor de start van het experiment werd voor laag (hoog) risico SIPs een percentage van 6% (16%) gehanteerd. We laten dit onderscheid intact gedurende het experiment en hanteren daarom in iedere experimentele groep twee verschillende toetsingspercentages. We onderscheiden de volgende vier experimentele groepen in dit gedeelte van het experiment:

- SIP controle: toetspercentages 6% en 16% voor respectievelijk laag en hoog risico SIPs (controlegroep);
- SIP laag: minder strenge toetsing, toetspercentages 2% en 10% voor laag en hoog risico SIPs;
- SIP hoog: strengere toetsing, toetspercentages 10% en 26% voor laag en hoog risico SIPs;
- SIP prestatie: toetspercentage afhankelijk van prestatie.

In de groep met prestatieafhankelijk toetsingspercentage wordt het toetspercentage twee-

⁵In totaal worden in 2013 27 verschillende indicatieprotocollen onderscheiden. 4 daarvan vallen in de hoog risico categorie, de overige indicatieprotocollen zijn aangemerkt als laag risico. Voor een overzicht van SIP protocollen zie Appendix C.

wekelijks als volgt aangepast. Voor elke conform getoetste aanvraag in de afgelopen twee weken wordt het toetspercentage met 0,2 procentpunt verlaagd. Anderzijds leidt iedere niet conform getoetste aanvraag tot een verhoging van het toetspercentage met 3 procentpunt. Deze getallen impliceren dat het toetsingspercentage constant blijft wanneer 93,75% van de getoetste aanvragen conform wordt bevonden. De toetspercentages in de SIP prestatie groep zijn onderhevig aan een minimum van 2% en een maximum van 26%.⁶ Alle zorgaanbieders in de SIP prestatie groep beginnen met een toetsingspercentage van 6% (16%) voor laag (hoog) risico SIPs.

Eenzijds is te verwachten dat er een negatief (positief) effect is van een verlaging (verhoging) van het toetsingspercentage op de zorgvuldigheid waarmee zorgaanbieders aanvragen voor AWBZ zorg indienen. Anderzijds kan het onconditioneel verlagen van het toetsingspercentage ook door zorgaanbieders worden geïnterpreteerd als een teken van vertrouwen. Als dit geven van vertrouwen succesvol is, dan zou men verwachten dat de eerder genoemde negatieve effecten op accuratesse uitblijven en zou anderzijds het verhogen van het toetsingspercentage de intrinsieke motivatie om accuraat te handelen kunnen onderdrukken. In het systeem met prestatieafhankelijke toetsingspercentages is een extra prikkel aanwezig voor zorgaanbieders om nauwkeurig te zijn in hun aanvraaggedrag, aangezien ze in geval van slechte prestaties geconfronteerd zullen worden met extra toetsingen. Bij het voorgaande dient wel te worden opgemerkt dat de gevolgen van een toetsing in het geval van SIP aanvragen beperkt zijn: in alle gevallen vindt toetsing achteraf plaats, waardoor zorgaanbieders weinig hinder ondervinden van de toetsing en er geen inhoudelijke aanpassing van de aanvraag kan plaatsvinden. Dit kan invloed hebben op de mate waarin zorgaanbieders reageren op variatie in het toetsingspercentage. In een regime van prestatieafhankelijk toetsingspercentage wordt het effect op het conformpercentage bepaald door een selectie-effect en een gedragseffect. Het selectie-effect houdt in dat er relatief meer aanvragen van slecht presterende zorgaanbieders worden getoetst. Het gedragseffect anderzijds houdt in dat zorgaanbieders een gedragsreactie kunnen hebben naar aanleiding van een aanpassing in het toetsingspercentage. Een verhoging van het toetsingspercentage kan een zorgaanbieder bijvoorbeeld aanzetten tot het nauwkeuriger indienen van aanvragen en daarmee een positieve invloed hebben op het conformpercentage.

2.2 De selectie van participerende zorgaanbieders

In totaal zijn er ongeveer 1.500 zorgaanbieders die HiT of SIP aanvragen kunnen indienen. Voor het eerste (tweede) onderdeel van het experiment hebben we zorgaanbieders geselecteerd die minstens 4 HiT (SIP⁷) aanvragen hebben ingediend in april 2012. Deze

⁶Bovendien wordt gewerkt met afgeronde toetspercentages. Uitgangspunt voor de tweewekelijkse aanpassing is telkens het ruwe toetspercentage, dat wil zeggen het toetspercentage vóór toepassing van het minimum en maximum en vóór afronding.

⁷Omdat we in dit onderzoek in eerste instantie naar een selectie van alle soorten SIPs zouden kijken, hebben we in de selectie van zorgaanbieders voor het tweede onderdeel van het experiment gekeken naar

selectieprocedure leidt er toe dat er 315 participerende zorgaanbieders in het eerste gedeelte (toetsingsmoment) van het experiment zijn en 252 zorgaanbieders in het tweede onderdeel (toetsingspercentage). Een zorgaanbieder kan tegelijk onderdeel zijn van beide delen van het experiment. Er zijn in totaal 473 verschillende zorgaanbieders die meedoen aan het experiment. De zorgaanbieders in het eerste gedeelte van het experiment zijn goed voor 89% van alle HiT aanvragen in de maand april 2012; de SIP zorgaanbieders nemen 93% van alle SIP aanvragen in april 2012 voor hun rekening. De geselecteerde zorgaanbieders worden willekeurig verdeeld over de verschillende experimentele groepen zoals beschreven in de vorige sectie. De indeling is opgenomen in Appendix B.

Op 1 januari 2013 is ERAI (Experiment Regelarme Instellingen) gestart. Een aantal van de zorgaanbieders geselecteerd voor deelname aan het experiment ‘Meten is Weten’, maakt ook deel uit van ERAI. Omdat deelname aan ERAI er voor zorgt dat er geen HiT en SIP aanvragen meer ingediend worden, laten wij deze zorgaanbieders in de analyses van ‘Meten is Weten’ buiten beschouwing. In totaal gaat het om 16 zorgaanbieders in het eerste gedeelte van het experiment en 24 zorgaanbieders in het tweede gedeelte van het experiment.⁸ Welke zorgaanbieders het betreft is aangegeven in Appendix B. Wanneer we de ERAI zorgaanbieders buiten beschouwing laten, hebben we in totaal 299 zorgaanbieders in het experiment waarin het toetsingsmoment wordt gevarieerd en 226 zorgaanbieders⁹ in het experiment waarin het toetsingspercentage wordt gevarieerd. In totaal gaat het om 443 verschillende zorgaanbieders die deelnemen aan het onderzoek.

Om uiteindelijk de effecten van de verschillende wijzigingen in toetsingspraktijk te kunnen analyseren, is het van belang dat de zorgaanbieders in de verschillende groepen voor aanvang van het experiment vergelijkbaar zijn. Tabel 2.1 geeft enkele beschrijvende statistieken. Deze cijfers zijn gebaseerd op gegevens over de periode van 9 januari 2012 tot de start van het experiment op 17 september 2012. De tabel laat zien dat, voor zowel de HiT als de SIP groepen, de gerapporteerde gemiddelden dicht bij elkaar liggen. Een formele test om te toetsen of de steekproeven uit dezelfde populatie komen, is de Kruskal-Wallis test. Op basis van de gerapporteerde p-waarden voor deze test kan de

aanvragen voor vijf soorten SIPs, te weten protocollen 03, 18, 28, 29, 30 en 34. Over de periode van 1 januari 2011 tot en met 30 april 2012, vormen deze vijf soorten SIPs 44% van het totale aantal SIP aanvragen voor alle zorgaanbieders die SIP aanvragen mogen indienen (CIZ Rapportage Relatiebeheer & Externe Toetsing, mei 2012).

⁸ERAI zorgaanbieders in de HiT (SIP) groepen zijn verantwoordelijk voor 10,3% (15,0%) van alle HiT (SIP, exclusief SIP 34 en SIP 35) aanvragen over de periode 1 januari 2012 tot 17 september 2012. Om precies te zijn gaat het om vier zorgaanbieders in de HiT controle groep, zes zorgaanbieders uit de HiT achteraf groep en zes zorgaanbieders in de HiT prestatie groep. In het tweede gedeelte van het experiment nemen zeven zorgaanbieders uit de SIP controle groep, zeven uit de SIP laag toetsingspercentage groep, vier uit de SIP hoog toetsingspercentage groep en zes zorgaanbieders uit de SIP prestatie groep deel aan ERAI.

⁹Er zijn twee zorgaanbieders extra verwijderd uit het experiment waarin het toetsingspercentage wordt gevarieerd. Deze zorgaanbieders hebben alleen SIP 34 of SIP 35 aanvragen ingediend, en hebben daarom effectief telkens 0 aanvragen omdat deze typen SIP aanvragen buiten beschouwing worden gelaten. Zie sectie 4.1 voor meer details.

hypothese dat de steekproeven in de HIT of SIP groepen uit dezelfde populatie komen, niet worden verworpen. Dit geldt voor elk van de gerapporteerde uitkomstvariabelen. Daarmee duiden de cijfers in onderstaande tabel op een valide randomisatie.

Tabel 2.1: Beschrijvende statistieken over de periode voorafgaand aan het experiment.

	# ZABs	gem. # aanvragen	gem. # toetsingen	gem. fractie conform
HiT controle	101	6,6 (0,2)	1,5 (0,1)	0,82 (0,01)
HiT achteraf	99	7,1 (0,2)	1,7 (0,1)	0,83 (0,01)
HiT prestatie	99	6,5 (0,2)	1,4 (0,1)	0,83 (0,01)
<i>p-waarden</i>		0,30	0,28	0,94
SIP controle	56	37,1 (1,2)	2,8 (0,1)	0,87 (0,01)
SIP laag	55	35,9 (1,2)	2,8 (0,1)	0,85 (0,01)
SIP hoog	58	37,3 (1,3)	2,9 (0,1)	0,87 (0,01)
SIP prestatie	57	34,0 (1,0)	2,6 (0,1)	0,88 (0,01)
<i>p-waarden</i>		0,97	0,99	0,10

Beschrijvende statistieken op basis van gegevens over de periode 9 januari 2012 tot 17 september 2012. Gemiddelden over zorgaanbieder-periode observaties. Een periode is gedefinieerd als twee weken. Standardfouten staan tussen haakjes weergegeven. P-waarden geven de resultaten van Kruskal-Wallis testen.

3 Verloop van het onderzoek

Het experiment is gestart op maandag 17 september 2012. Voorafgaand hieraan, begin september, zijn de directies van de zorgaanbieders die onderdeel uitmaken van het experiment geïnformeerd middels een brief. In deze brief zijn het doel en de opzet van het onderzoek en de verwachte looptijd gecommuniceerd. De relatiebeheerders van het CIZ hebben daarnaast hun contactpersoon bij de betreffende zorgaanbieders middels een e-mail ingelicht over het onderzoek en er heeft een bericht over het onderzoek in CIZ Info van 31 oktober 2012 gestaan. Zorgaanbieders in de HiT prestatie groep en de SIP prestatie groep ontvingen bovendien rond de ingangsdatum van de tweewekelijkse aanpassingen een e-mail waarin zij geïnformeerd werden over de wijzigingen van dat moment.

Voor zorgaanbieders in de HiT controlegroep, de HiT achteraf groep en de SIP controle, SIP laag en SIP hoog groepen is op maandag 17 september 2012 de hiervoor beschreven toetsingspraktijk ingesteld. Voor deze zorgaanbieders veranderde de toetsingspraktijk niet meer gedurende het onderzoek. De eerste prestatiegerelateerde aanpassingen van de toetsingspraktijk voor zorgaanbieders in de HiT prestatie en SIP prestatie groepen zijn ingegaan op maandag 15 oktober 2012. De tweede ronde van aanpassingen is per maandag 5 november 2012 doorgevoerd en vanaf dat moment heeft iedere twee weken een update plaatsgevonden. In totaal is de toetsingspraktijk 12 keer aangepast. Tabel 3.1 beschrijft het tijdsplan van het experiment en laat zien welke gegevens gebruikt zijn voor het bepalen van de periodieke aanpassingen.

Tabel 3.1: Tijdsplan periodieke aanpassingen in toetsingspraktijk.

Aanpassingsronde	Gegevensperiode SIP aanpassing	Ingangsdatum
Ronde 1	17-09-2012 t/m 30-09-2012	15 oktober 2012
Ronde 2	01-10-2012 t/m 28-10-2012	5 november 2012
Ronde 3	29-10-2012 t/m 11-11-2012	19 november 2012
Ronde 4	12-11-2012 t/m 25-11-2012	3 december 2012
Ronde 5	26-11-2012 t/m 09-12-2012	17 december 2012
Ronde 6	10-12-2012 t/m 23-12-2012	31 december 2012
Ronde 7	24-12-2012 t/m 06-01-2013	14 januari 2013
Ronde 8	07-01-2013 t/m 20-01-2013	28 januari 2013
Ronde 9	21-01-2013 t/m 03-02-2013	11 februari 2013
Ronde 10	04-02-2013 t/m 17-02-2013	25 februari 2013
Ronde 11	18-02-2013 t/m 03-03-2013	11 maart 2013
Ronde 12	04-03-2013 t/m 17-03-2013	25 maart 2013

Voor de aanpassingen in de HiT prestatie groep werd telkens data over de meest recente periode van 8 weken genomen om de aanpassingen te berekenen. De einddatum was daarbij telkens de einddatum genoemd in de tweede kolom.

Bij de berekening en doorvoering van de eerste aanpassingen zijn enkele opstartproblemen opgetreden. Hier is, waar nodig, voor gecorrigeerd in latere aanpassingen.¹⁰ De commu-

¹⁰Opstartproblemen hebben zich voorgedaan tot en met de vierde ronde van aanpassingen. De belangrijkste zijn geweest:

1. Het maximum en minimum op de toetsingspercentages in de SIP prestatie groep is in enkele

nicatie van de aanpassingen is in de eerste rondes van het onderzoek soms lastig gebleken. Een bijeenkomst met relatiebeheerders op 4 december 2012 en een document met uitleg over de aanpassingen heeft voor opheldering gezorgd. Daarnaast is vanaf 1 januari 2013 SIP 35 uitgesloten van het experiment. Voor dit protocolnummer was, in verband met een inhoudelijke wijziging per die datum, een hoger toetsingspercentage gewenst. Het toetsingspercentage voor dit type SIP is vanaf 1 januari in alle groepen op 50% gezet en daarna niet meer periodiek aangepast.

De laatste aanpassingen in de prestatieafhankelijke toetsingspraktijk voor zorgaanbieders in de HiT prestatie groep en de SIP prestatie groep zijn doorgevoerd per 25 maart 2013. Het experiment is geëindigd op 8 april 2013. Op dat moment is de toetsingspraktijk teruggezet naar de situatie voor de start van het experiment, maar voor de HiT aanvragen is vanaf dat moment een gewijzigde vorm van toetsingspraktijk ingesteld waarbij het toetsingspercentage wordt aangepast op basis van prestaties in het verleden.

gevallen niet doorgevoerd.

2. In de eerste aanpassingsronde waren enkele toetsingen niet opgenomen in de data. Dit is gecorrigeerd in de tweede ronde van aanpassingen.
3. Van 13 zorgaanbieders zijn tot en met de vierde aanpassingsronde geen toetsingen geregistreerd. Bij de berekening van de vijfde aanpassingsronde is voor deze zorgaanbieders verder gewerkt met de (ongewijzigde) toetsingspraktijk zoals deze tot en met de vierde ronde was doorgevoerd.

4 Data

Ten behoeve van het onderzoek zijn periodiek rapportages gemaakt met daarin informatie over de uitgevoerde toetsingen en hun resultaat. Op basis daarvan zijn telkens de periodieke aanpassingen in de groepen met prestatieafhankelijke toetsingspraktijk berekend. Na afloop van het experiment is een uitgebreidere dataset opgesteld. Deze laatste dataset wordt gebruikt voor analysedoeleinden in deze eindrapportage. In sectie 4.1 worden de toegepaste selecties op de data en de inhoud van de uiteindelijke dataset kort beschreven. Vervolgens worden in sectie 4.2 enkele beschrijvende statistieken gepresenteerd. Deze geven een eerste beeld van de omvang van de dataset in termen van verschillende soorten aanvragen, uitgesplitst voor de diverse groepen en voor de periode voor en tijdens het experiment.

4.1 Beschrijving van de beschikbare gegevens

De definitieve dataset bevat informatie op aanvraagniveau. De dataset bevat alle aanvragen met een aanvraagdatum tussen 1 januari 2012 en 7 april 2013 ingediend door een in het experiment participerende zorgaanbieder. In totaal hebben we informatie over 1.200.825 aanvragen.¹¹ Dit betreft reguliere AWBZ aanvragen, spoedaanvragen, indicatiemeldingen (IM), HiT aanvragen en SIP aanvragen. Voor HiT en SIP aanvragen weten we bovendien welk type HiT of welk SIP protocolnummer de aanvraag betreft.

We observeren voor iedere aanvraag de soort, de datum van aanvraag, datum van toetsing (indien van toepassing), hoeveelheid (klassen en/of zorgzwaartepakketten (ZZP)) en soort (functies) aangevraagde en geïndiceerde zorg. Daarnaast weten we, voor HiT en SIP aanvragen, welk toetsingspercentage en toetsingsmoment voor de betreffende aanvraag is gebruikt, de uitvoerende toetsers, het toetsingsresultaat ('conform' of 'niet conform') en de (gedetailleerde) reden voor dat toetsingsresultaat. We hebben eveneens basisinformatie (leeftijd, geslacht, postcodecijfers) over de cliënt voor wie zorg is aangevraagd. Een volledig overzicht van de beschikbare informatie is te vinden in Appendix D.

We passen een aantal selectiecriteria toe op deze dataset om te komen tot de dataset waarop de analyses worden gedaan. Details omtrent het toepassen van de selectiecriteria zijn beschreven in Appendix E.

Om een beeld te krijgen van de grootte van de dataset en het aandeel van diverse soorten aanvragen, geeft Tabel 4.1 voor diverse soorten aanvragen weer hoeveel er daarvan zijn ingediend over de periode van 1 januari 2012 tot en met 7 april 2013. Voor zorgaanbieders in de HiT groepen geldt dat 17% van de ingediende aanvragen een HiT aanvraag is geweest. Daarnaast is 33% van de ingediende aanvragen een SIP aanvraag en is 22%

¹¹Aanvragen van zorgaanbieders die in zowel een HiT groep als een SIP groep voorkomen, worden hierin dubbel geteld. Het aantal unieke aanvragen is 799.291.

van de aanvragen een spoedaanvraag. Van alle ingediende HiT aanvragen zijn er bijna 17.000 getoetst, dit komt neer op een toetsingspercentage van 23,3%. Voor zorgaanbieders in de SIP groepen geldt dat het overgrote gedeelte van de ingediende aanvragen een SIP aanvraag is geweest (53%). Daarnaast is 19% van de ingediende aanvragen een spoedaanvraag. De HiT aanvragen vormen een aanzienlijk kleiner gedeelte van de totale aanvragenstroom dan in de HiT groepen het geval is. Van alle ingediende SIP aanvragen, zijn er ruim 22.000 getoetst. Dat impliceert een toetsingspercentage van 8,3%.

Tabel 4.1: Aantallen aanvragen en toetsingen, 1 januari 2012 tot en met 7 april 2013.

	HiT groepen		SIP groepen	
	<i>aantal</i>	<i>fractie</i>	<i>aantal</i>	<i>fractie</i>
Aantal aanvragen (totaal)	438.496		511.888	
HiT	72.446	0,17	38.630	0,08
SIP	145.552	0,33	269.142	0,53
Indicatiemeldingen	33.224	0,08	26.762	0,05
AWBZ Spoedaanvragen	94.310	0,22	95.316	0,19
Reguliere AWBZ aanvragen	92.947	0,21	82.023	0,16
Aantal HiT / SIP toetsingen	16.876		22.279	

4.2 Beschrijvende statistieken van de data

In deze sectie worden beschrijvende statistieken van de data gepresenteerd. Sectie 4.2.1 geeft informatie over diverse kenmerken van zorgaanbieders. In sectie 4.2.2 wordt een beschrijving gegeven van zorginhoudelijk kenmerken. Een beeld van het gemiddeld aantal ingediende aanvragen, het aantal toetsingen en de fractie conform getoetste aanvragen wordt tenslotte gepresenteerd in sectie 4.2.3.

4.2.1 Kenmerken van zorgaanbieders

Tabel 4.2 geeft informatie over het type zorgaanbieder (bv. ziekenhuis)¹² en, voor intramurale zorg, de deskundigheid van de zorgaanbieder (verpleging & verzorging (V&V), geestelijke gezondheidszorg (GGZ) of gehandicaptenzorg (GZ)). Hierin bestaan verschillen tussen HiT en SIP groepen: in HiT groepen zijn er geen ziekenhuizen en is er meer spreiding in deskundigheid, terwijl in SIP groepen tussen 32% en 43% van de zorgaanbieders als ziekenhuis is geclassificeerd en er nauwelijks variatie is in de deskundigheid.

Op basis van informatie over aantallen ingediende HiT (SIP) aanvragen door zorgaanbieders in de diverse HiT (SIP) groepen, genereren we een indicator die aangeeft of een zorgaanbieder geclassificeerd kan worden als kleine, middelgrote of grote zorgaanbieder.

¹²Er zijn verschillende mogelijkheden voor het type zorgaanbieder. Het belangrijkste onderscheid is dat tussen ziekenhuizen en andere zorgaanbieders. De fractie zorgaanbieders geclassificeerd als ziekenhuis is gerapporteerd in de tabel.

We classificeren een zorgaanbieder in een HiT (SIP) groep als kleine zorgaanbieder indien gemiddeld per periode van twee weken (gemeten over de periode van 2 januari 2012 tot 17 september 2012) minder dan 2 (7) HiT (SIP) aanvragen zijn ingediend door de betreffende zorgaanbieder. Een zorgaanbieder in een HiT (SIP) groep wordt geclassificeerd als middelgroot wanneer deze gemiddeld meer dan 2 (7) maar minder dan 7 (45) HiT (SIP) aanvragen heeft ingediend. Tenslotte wordt een zorgaanbieder aangemerkt als grote zorgaanbieder indien deze gemiddeld meer dan 7 (45) HiT (SIP) aanvragen heeft ingediend. De fracties kleine, middelgrote en grote zorgaanbieders in iedere groep zijn gerapporteerd in Tabel 4.2.

Tabel 4.2: Kenmerken van zorgaanbieders.

HiT groepen				
	<i>controle</i>	<i>achteraf</i>	<i>prestatie</i>	
Ziekenhuis	0,00	0,00	0,00	
V&V deskundigheid	0,50	0,52	0,61	
GGZ deskundigheid	0,22	0,25	0,23	
GZ deskundigheid	0,26	0,21	0,15	
Kleine zorgaanbieder	0,14	0,19	0,30	
Middelgrote zorgaanbieder	0,58	0,48	0,41	
Grote zorgaanbieder	0,28	0,32	0,28	
<i>Regionale specialisatie</i>				
Gespecialiseerd	0,94	0,92	0,92	
Noord Nederland	0,07	0,15	0,11	
Oost Nederland	0,22	0,23	0,15	
Zuid Nederland	0,29	0,22	0,24	
West Nederland	0,37	0,31	0,41	
SIP groepen				
	<i>controle</i>	<i>laag</i>	<i>hoog</i>	<i>prestatie</i>
Ziekenhuis	0,32	0,44	0,38	0,39
V&V deskundigheid	0,95	0,93	0,93	0,91
GGZ deskundigheid	0,02	0,02	0,00	0,02
GZ deskundigheid	0,00	0,00	0,00	0,02
Kleine zorgaanbieder	0,20	0,15	0,24	0,16
Middelgrote zorgaanbieder	0,52	0,58	0,43	0,51
Grote zorgaanbieder	0,29	0,27	0,33	0,33
<i>Regionale specialisatie</i>				
Gespecialiseerd	0,91	0,88	0,98	0,89
Noord Nederland	0,16	0,11	0,12	0,09
Oost Nederland	0,11	0,20	0,15	0,18
Zuid Nederland	0,18	0,25	0,20	0,23
West Nederland	0,46	0,32	0,51	0,40

Gerapporteerd zijn fracties zorgaanbieders in de betreffende groep.

Voor iedere ingediende aanvraag hebben we basisinformatie over de cliënt voor wie zorg wordt aangevraagd. Onderdeel daarvan is de vier cijfers van de postcode. Die informatie kunnen we gebruiken om te bepalen of een zorgaanbieder voornamelijk actief is in een bepaalde regio van Nederland. Een zorgaanbieder is gedefinieerd als regionaal, wanneer het percentage van het totaal aantal aanvragen in die regio minstens 80% is. Het kan dus zijn dat een zorgaanbieder verspreid over meerdere regio's actief is en daarom niet wordt geclassificeerd als regionaal. In Tabel 4.2 is de fractie regionale zorgaanbieders gerapporteerd. Bovendien geeft de tabel per regio weer welke fractie van het totaal aantal zorgaanbieders in de groep voornamelijk actief is in die specifieke regio. De tabel laat zien dat het grootste deel van de zorgaanbieders regionaal is: in de meeste groepen is meer dan 90% van de zorgaanbieders gespecialiseerd. De meeste zorgaanbieders zijn voornamelijk actief in de regio West Nederland, wat te maken kan hebben met een grotere zorgvraag en groter aantal inwoners in deze regio.

4.2.2 Inhoudelijke kenmerken van de zorgaanvragen

De uiteindelijke dataset bevat eveneens zorginhoudelijke informatie, zoals de geïndiceerde functie en zorgomvang. In Tabel 4.3 wordt, per experiment groep, aangegeven welke fractie van alle ingediende aanvragen tot indicatie van een bepaalde functie heeft geleid. Een aanvraag kan meerdere geïndiceerde functies omvatten. In de bepaling van deze fractie worden alle soorten aanvragen (HiT, SIP, IM, spoedaanvragen en reguliere AWBZ aanvraag) meegenomen. De functie persoonlijke verzorging (PV) wordt verreweg het meeste geïndiceerd, in meer dan de helft van de aanvragen door zorgaanbieders in HiT groepen. Voor SIP groepen ligt deze fractie zelfs nog hoger (70% van de aanvragen). Ook verpleging (VP) is een veel geïndiceerde functie (30% van de aanvragen). In ongeveer 24% (16%) van de aanvragen door zorgaanbieders in een HiT (SIP) groep wordt een ZZP pakket geïndiceerd en betreft het dus intramurale zorg. De gemiddelde geïndiceerde klasse¹³ van diverse functies varieert tussen SIP en HiT groepen; de gemiddelde klasse begeleiding individueel (BGI) is bijvoorbeeld aanzienlijk hoger in SIP groepen.

De meeste indicaties worden afgegeven voor vrouwen: ruim 60% van de aanvragen betreft zorg voor een vrouwelijke cliënt. De gemiddelde leeftijd van cliënten waarvoor zorg wordt aangevraagd door zorgaanbieders in de HiT groepen ligt rond 71 jaar; de gemiddelde leeftijd bij aanvragen door zorgaanbieders in de SIP groepen ligt iets hoger. Dit heeft te maken met een verschil in samenstelling van het soort aanvragen zoals kan worden opgemerkt uit Tabel 4.1.

Tenslotte is voor niet conform getoetste HiT aanvragen gekeken naar de gerapporteerde redenen voor niet conform toetsingsresultaten. Voor iedere niet conform getoetste HiT aanvraag kunnen meerdere redenen voor niet conform toetsing worden opgegeven. De cijfers in Tabel 4.3 geven aan dat voor veel niet conform getoetste aanvragen de aangevraagde klasse te hoog is (ongeveer 33% van alle niet conform toetsingen); een te lage

¹³Bij het bepalen van de gemiddelde geïndiceerde klasse wordt alleen gekeken naar de eerste volgtijdige indicatie.

Tabel 4.3: Zorginhoudelijke kenmerken van aanvragen en HiT toetsingen, HiT groepen.

	HiT groepen					
	<i>controle</i>		<i>achteraf</i>		<i>prestatie</i>	
	<i>voor</i>	<i>na</i>	<i>voor</i>	<i>na</i>	<i>voor</i>	<i>na</i>
<i>Geïndiceerde functies</i>						
PV geïndiceerd	0,57	0,57	0,58	0,57	0,65	0,64
VP geïndiceerd	0,29	0,28	0,30	0,29	0,35	0,33
BGG geïndiceerd	0,10	0,10	0,11	0,11	0,09	0,09
BGI geïndiceerd	0,16	0,15	0,14	0,14	0,14	0,14
BHG of BHI geïndiceerd	0,03	0,03	0,03	0,02	0,02	0,02
ZZP geïndiceerd	0,24	0,25	0,24	0,25	0,17	0,18
<i>Klassen & type ZZP pakket*</i>						
GGZ	0,14	0,14	0,14	0,13	0,20	0,17
VG	0,22	0,28	0,15	0,22	0,13	0,16
V&V	0,59	0,53	0,67	0,62	0,63	0,62
PV klasse (1 - 9)	3,35	3,34	3,28	3,30	3,32	3,33
VP klasse (0 - 8)	1,83	1,82	1,81	1,81	1,88	1,86
BGG klasse (1 - 10)	5,17	5,18	5,16	5,12	5,07	5,10
BGI klasse (1 - 9)	3,74	3,76	3,57	3,60	4,07	4,17
BHG klasse (1 - 10)	6,47	5,94	5,74	5,73	5,61	5,55
<i>Cliëntinformatie</i>						
Vrouw	0,61	0,61	0,62	0,61	0,63	0,62
Leeftijd	70,7	70,7	72,0	71,9	73,4	73,8
Observaties	74.103	58.363	83.182	66.202	88.576	68.073
<i>HiT toetsing - Reden niet conform**</i>						
Afwijkende grondslag	0,08	0,08	0,11	0,11	0,09	0,13
Afwijkende leveringsvoorwaarden	0,29	0,17	0,25	0,08	0,23	0,15
Afwijkende functies	0,23	0,30	0,27	0,24	0,27	0,26
Afwijkende klasse (te hoog)	0,31	0,31	0,36	0,39	0,30	0,36
Afwijkende klasse (te laag)	0,05	0,08	0,05	0,02	0,05	0,06
Afwijkende verbijzonderingen	0,22	0,16	0,17	0,00	0,22	0,13
Afwijkende beperkingen	0,48	0,46	0,45	0,00	0,45	0,31
Niet integraal	0,10	0,12	0,08	0,06	0,12	0,08
Observaties	467	477	516	500	412	398

Gemiddelden over alle soorten aanvragen.

* Alleen fracties ZZP V&V, GGZ en VG geïndiceerde pakketten zijn weergegeven; de fracties LG en ZG zijn zeer klein (respectievelijk rond 0,03 en <0,01).

** Fractie van niet conform getoetste HiT aanvragen waarvoor een bepaalde reden voor niet conform toetsingsresultaat wordt gerapporteerd. Voor een niet conform getoetste HiT aanvragen kunnen meerdere redenen worden opgegeven.

Tabel 4.4: Zorginhoudelijke kenmerken van aanvragen, SIP groepen.

	SIP groepen							
	<i>controle</i>		<i>laag</i>		<i>hoog</i>		<i>prestatie</i>	
	<i>voor</i>	<i>na</i>	<i>voor</i>	<i>na</i>	<i>voor</i>	<i>na</i>	<i>voor</i>	<i>na</i>
<i>Geïndiceerde functies</i>								
PV geïndiceerd	0,72	0,73	0,70	0,70	0,70	0,71	0,71	0,72
VP geïndiceerd	0,38	0,37	0,39	0,37	0,39	0,38	0,40	0,38
BGG geïndiceerd	0,07	0,07	0,07	0,06	0,07	0,08	0,07	0,07
BGI geïndiceerd	0,12	0,12	0,10	0,10	0,10	0,10	0,12	0,11
BHG of BHI geïndiceerd	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01
ZZP geïndiceerd	0,15	0,14	0,17	0,19	0,16	0,15	0,15	0,14
<i>Klassen & type ZZP pakket*</i>								
V&V	0,96	0,97	0,96	0,97	0,97	0,97	0,95	0,96
PV klasse (1 - 9)	3,34	3,35	3,30	3,35	3,35	3,38	3,37	3,37
VP klasse (0 - 8)	1,87	1,86	1,91	1,90	1,87	1,87	1,96	1,96
BGG klasse (1 - 10)	5,25	5,17	5,12	5,11	5,26	5,33	5,05	5,11
BGI klasse (1 - 9)	5,16	5,05	5,61	5,91	5,48	5,53	5,48	5,60
BHG klasse (1 - 10)	5,13	5,04	5,06	5,09	5,14	5,15	4,97	4,85
<i>Clïëntinformatie</i>								
Vrouw	0,63	0,63	0,62	0,62	0,64	0,63	0,63	0,63
Leeftijd	76,1	76,5	75,5	76,1	76,3	76,5	75,8	76,0
Observaties	72.193	61.361	65.910	47.516	77.583	59.347	70.373	57.205

Gemiddelden over alle soorten aanvragen.

* Alleen fracties ZZP V&V, GGZ en VG geïndiceerde pakketten zijn weergegeven; de fracties LG en ZG zijn zeer klein (respectievelijk rond 0,03 en <0,01).

** Fractie van niet conform getoetste HiT aanvragen waarvoor een bepaalde reden voor niet conform toetsingsresultaat wordt gerapporteerd. Voor een niet conform getoetste HiT aanvragen kunnen meerdere redenen worden opgegeven.

aangevraagde klasse komt daarentegen veel minder vaak voor (slechts 6% van de gevallen). Tevens is in veel gevallen sprake van afwijkende beperkingen, voornamelijk in de controlegroep. Tenslotte is een afwijkende functie in 27% van de gevallen een reden voor een niet conform toetsingsresultaat.¹⁴

4.2.3 Aantallen aanvragen, toetsingen en fractie conform

Tabel 4.5 geeft informatie over het gemiddeld aantal ingediende aanvragen, uitgesplitst naar soort, het aantal toetsingen en de fractie conform getoetste aanvragen in de verschillende experimentele groepen en uitgesplitst in voor/tijdens het experiment. Allereerst is te zien dat het gemiddeld aantal HiT aanvragen per zorgaanbieder per periode in alle HiT groepen (iets) lager is in de periode voorafgaand aan het experiment. De toename in het aantal HiT aanvragen tijdens het experiment is procentueel het grootste in de HiT achteraf groep. Wanneer we het aantal HiT aanvragen uitsplitsen naar type, zien we dat de grootste categorie HiT type 1 ('aanpassen indicatie extramurale zorg') is, gevolgd door

¹⁴De registratie van de reden voor niet conform toetsing bij de achteraf getoetste HiT aanvragen lijkt op een andere manier te hebben plaatsgevonden. Waar de fracties waarin de diverse redenen voorkomen in de periode voor het experiment nog redelijk vergelijkbaar zijn tussen de verschillende groepen, zien we dat in de periode van het experiment een aantal redenen niet langer voorkomt voor achteraf toetsingen. Dit is bijvoorbeeld het geval voor 'afwijkende verbijzonderingen' en 'afwijkende beperkingen'.

HiT type 2 (‘gelijkblijvend zorgzwaartepakket verlengen’) en type 4 (‘aanpassen zorgzwaartepakket’) aanvragen. De procentuele toename is verreweg het grootst voor HiT type 2 aanvragen (tussen de 73% toename in de HiT prestatie groep en 88% toename in de HiT controle groep). In alle HiT groepen is het gemiddeld aantal HiT toetsingen hoger in de experiment periode. Het gaat om een procentuele toename van 24% in de HiT prestatie groep en een toename van 26% (48%) in de HiT controle (achteraf) groep. Dit is het directe gevolg van de toename in het aantal aanvragen, aangezien het gehanteerde toetsingspercentage gelijk is gebleven. Tenslotte zien we voor de HiT achteraf groep een behoorlijke stijging in de gemiddelde fractie conform getoetste aanvragen gedurende het experiment, het gaat om een toename van 2,7 procentpunt, van 83% naar 85,7% conform. Daarmee is de gemiddelde fractie conform in de experiment periode ook hoger dan in de HiT controle en HiT prestatie groep.

Tabel 4.5: Aanvragen, toetsingen en fractie conform getoetst.

HiT groepen								
	<i>controle</i>		<i>achteraf</i>		<i>prestatie</i>			
	<i>voor</i>	<i>na</i>	<i>voor</i>	<i>na</i>	<i>voor</i>	<i>na</i>		
HiT aanvragen	6,76	8,39	7,12	9,79	6,59	7,45		
HiT Type 1	4,87	5,46	5,22	7,04	5,00	5,49		
HiT Type 2	0,85	1,59	0,76	1,41	0,56	0,97		
HiT Type 3	0,24	0,28	0,30	0,27	0,21	0,17		
HiT Type 4	0,73	1,05	0,79	1,07	0,76	0,82		
HiT Type 6	0,08	0,00	0,06	0,00	0,05	0,00		
HiT toetsingen	1,54	1,94	1,67	2,47	1,41	1,76		
Fractie conform	0,82	0,82	0,83	0,86	0,83	0,83		
Observaties	1.818	1.414	1.782	1.386	1.782	1.386		
# <i>ZABs</i>	101	101	99	99	99	99		
# <i>perioden</i>	18	14	18	14	18	14		
SIP groepen								
	<i>controle</i>		<i>laag</i>		<i>hoog</i>		<i>prestatie</i>	
	<i>voor</i>	<i>na</i>	<i>voor</i>	<i>na</i>	<i>voor</i>	<i>na</i>	<i>voor</i>	<i>na</i>
SIP aanvragen	37,36	40,78	36,04	34,35	37,91	37,85	34,25	37,68
SIP laag risico	30,48	33,69	28,96	26,61	30,98	30,24	25,55	27,92
SIP hoog risico	6,75	6,98	7,01	7,69	6,84	7,55	8,57	9,64
SIP toetsingen	2,79	3,05	2,76	1,28	2,90	4,85	2,67	4,70
Fractie conform	0,87	0,85	0,85	0,86	0,87	0,86	0,88	0,86
Observaties	1.008	784	990	770	1.044	812	1.026	798
# <i>ZABs</i>	56	56	55	55	58	58	57	57
# <i>perioden</i>	18	14	18	14	18	14	18	14

Gemiddelden over periode - zorgaanbieder observaties. Een periode is twee weken.

De kolommen getiteld *voor* zijn gemiddelden over de periode vóór de start van het experiment (9 januari 2012 tot 17 september 2012). De kolommen getiteld *na* zijn gebaseerd op de periode na invoering van variatie in toetsingspraktijk (17 september 2012 tot en met 7 april 2013).

In het tweede deel van Tabel 4.5 zijn soortgelijke cijfers voor de vier SIP groepen weergegeven. Deze statistieken laten zien dat het gemiddeld aantal SIP aanvragen per periode per zorgaanbieder in de controlegroep en de SIP prestatiegroep iets hoger ligt tijdens het

experiment, terwijl er sprake is van een daling in de laag en hoog toetsingspercentage groepen. Het gaat hier echter om vrij kleine wijzigingen van 3% tot 6%. De uitsplitsing naar laag en hoog risico SIP protocollen laat zien dat de meeste SIP aanvragen in de laag risico categorie vallen.

De variatie in toetsingspraktijk voor zorgaanbieders in de SIP groepen heeft directe gevolgen voor het aantal uitgevoerde SIP toetsingen gedurende het experiment. In de SIP laag toetsingspercentage groep, is het toetsingspercentage omlaag bijgesteld, wat zich vertaalt in een 55% afname van het aantal getoetste aanvragen. De bijstelling omhoog van het toetsingspercentage in de SIP hoog toetsingspercentage groep, zien we terug in een 62% toename van het aantal getoetste aanvragen. In de SIP prestatie groep is het toetsingspercentage voor iedere zorgaanbieder periodiek bijgesteld op basis van prestatie. Het gemiddelde aantal toetsingen per zorgaanbieder in deze groep ligt aanzienlijk hoger gedurende het experiment. Dit duidt op een toename in het gemiddelde toetsingspercentage. De ontwikkeling in prestatieafhankelijke toetsingspercentages komt aan bod in sectie 5.1.

Tenslotte zien we een daling in de fractie conform getoetste aanvragen in de controle-groep, de groep met het hoge toetsingspercentage en de SIP prestatie groep. In de laag toetsingspercentage groep blijft het gemiddelde conformpercentage ongeveer gelijk. De fractie conform ligt in het algemeen iets hoger voor SIP aanvragen (in de SIP groepen) dan voor HiT aanvragen (in de HiT groepen).

5 Beschrijvende analyses

In deze sectie worden beschrijvende analyses gepresenteerd. Sectie 5.1 bevat informatie over de wijzigingen in de prestatieafhankelijke toetsingspraktijk die periodiek zijn doorgevoerd.¹⁵ Sectie 5.1.1 doet dit voor de HiT prestatie groep, sectie 5.1.2 doet dit voor de SIP prestatie groep. De aanpassingen in de toetsingspraktijk geven een beeld van het gemiddelde prestatieniveau van zorgaanbieders in deze prestatieafhankelijke toetsingsgroepen. Bovendien geeft het inzicht in de mate waarin zorgaanbieders variëren in hun prestatieniveau en daarmee de mate waarin prestatiegerelateerde toetsingspraktijk kan worden benut om de middelen beschikbaar voor toetsing daar in te zetten waar ruimte is voor verbetering van het aanvraaggedrag. In totaal zijn er 12 keer aanpassingen doorgevoerd in het toetsingsmoment en het toetsingspercentage. Sectie 5.2 gaat vervolgens in op de relevante uitkomstvariabelen. De figuren die in deze sectie worden gepresenteerd, geven een eerste beeld van de effecten van de diverse vormen van toetsingspraktijk. De formele econometrische analyses volgen in sectie 6.

5.1 Periodieke aanpassingen prestatieafhankelijke toetsingspraktijk

5.1.1 Aanpassingen in het toetsingsmoment (HiT prestatie groep)

Tabel 5.1 laat per aanpassingsronde zien welk percentage van de zorgaanbieders in de HiT prestatie groep (95 zorgaanbieders in totaal) vooraf wordt getoetst. Bovendien wordt weergegeven hoeveel zorgaanbieders van vooraf naar achteraf en vice versa bewegen.

In Tabel 5.1 is te zien dat het percentage zorgaanbieders in de HiT prestatie groep dat achteraf wordt getoetst gedurende het experiment is toegenomen. Er zijn dus steeds meer zorgaanbieders die boven de grens van 92% conform scoren. Het aantal zorgaanbieders op achteraf toetsing was een aantal rondes vrij stabiel (ronde 5 tot en met 9), maar is in de laatste aanpassingsrondes verder toegenomen. Na de laatste aanpassingsronde werd 51,6% van de zorgaanbieders achteraf getoetst. Echter, het feit dat nog steeds 48,4% van de zorgaanbieders vooraf wordt getoetst, geeft aan dat bijna de helft van de zorgaanbieders nog beneden de norm van 92% presteerde. De laatste twee kolommen van Tabel 5.1 geven aan dat er wat fluctuaties zijn in het aantal veranderingen van toetsingsmoment per aanpassingsronde. Er zijn telkens zowel zorgaanbieders die van vooraf naar achteraf gaan, als zorgaanbieders die van achteraf naar vooraf toetsing gaan.

¹⁵In de HiT prestatiegroep is een aantal (4 van de 99) zorgaanbieders geweest waarvoor geen toetsingen zijn uitgevoerd gedurende het experiment. Deze zorgaanbieders maken bijna allemaal deel uit van de uitgevallen zorgaanbieders zoals gedefinieerd in Appendix E. Voor hen is het toetsingsmoment daarom gedurende het hele experiment vooraf geweest. Omdat dit niet het gevolg is van behaalde conformpercentages bij toetsingen, zijn deze zorgaanbieders in de beschrijvingen in deze sectie buiten beschouwing gelaten.

Tabel 5.1: Periodieke aanpassingen in het toetsingsmoment.

Aanpassingsronde	% vooraf	vooraf → achteraf		achteraf → vooraf	
		aantal	%	aantal	%
Ronde 0	100,0%	-		-	
Ronde 1 <i>15 oktober 2012</i>	57,9%	40	(42,1%)	0	(0,0%)
Ronde 2 <i>5 november 2012</i>	61,1%	9	(9,5%)	12	(12,6%)
Ronde 3 <i>19 november 2012</i>	57,8%	5	(5,3%)	1	(1,1%)
Ronde 4 <i>3 december 2012</i>	57,8%	6	(6,3%)	6	(6,3%)
Ronde 5 <i>17 december 2012</i>	52,6%	6	(6,3%)	2	(2,1%)
Ronde 6 <i>31 december 2012</i>	53,7%	6	(6,3%)	7	(7,4%)
Ronde 7 <i>14 januari 2013</i>	53,7%	3	(3,2%)	3	(3,2%)
Ronde 8 <i>28 januari 2013</i>	55,8%	7	(7,4%)	9	(9,5%)
Ronde 9 <i>11 februari 2013</i>	55,8%	5	(5,3%)	5	(5,3%)
Ronde 10 <i>25 februari 2013</i>	50,5%	8	(8,4%)	3	(3,2%)
Ronde 11 <i>11 maart 2013</i>	51,6%	3	(3,2%)	4	(4,2%)
Ronde 12 <i>25 maart 2013</i>	48,4%	10	(10,5%)	7	(7,4%)

Conclusie (1): *Gedurende het experiment zijn er steeds meer zorgaanbieders in de HiT prestatie groep boven de 92% conform-norm gaan presteren. Dit heeft zich vertaald in een toenemend aantal achteraf getoetste zorgaanbieders.*

Naast de in Tabel 5.1 gepresenteerde informatie is het ook interessant om te weten of een groep zorgaanbieders balanceert op de grens van de 92%-norm en daardoor het toetsingsmoment voor deze zorgaanbieders telkens wordt aangepast, of dat bepaalde zorgaanbieders consequent boven (onder) de 92%-norm presteren. Tabel 5.2 geeft inzicht in de verdeling van de wijzigingen in het toetsingsmoment over zorgaanbieders. Het toetsingsmoment kan voor een zorgaanbieder maximaal twaalf keer gewijzigd zijn. De tabel laat zien dat het grootste aantal wijzigingen vijf is: er zijn twee zorgaanbieders voor wie het toetsingsmoment vijf maal is aangepast. De observatie dat zorgaanbieders niet telkens wisselen tussen vooraf en achteraf toetsing geeft aan dat het gebruik van een prestatieafhankelijk toetsingsmoment maakt dat we onderscheid kunnen maken tussen zorgaanbieders die voldoende accuraat zijn in het indienen van aanvragen en zorgaanbieders die vaak beneden de norm presteren. Bovendien geeft de tabel weer dat voor 12 zorgaanbieders het toetsingsmoment geen enkele keer is gewijzigd, zij zijn altijd vooraf getoetst. Dat betekent dat 12% van de zorgaanbieders in de HiT prestatie groep consequent onder de 92%-norm heeft gepresteerd.

Conclusie (2): *Ongeveer 12% van de zorgaanbieders in de HiT prestatie groep presteert*

Tabel 5.2: Aantal wijzigingen in toetsingsmoment.

Aantal wijzigingen	Aantal (%) zorgaanbieders
0	12 (12,6%)
1	32 (33,7%)
2	28 (29,5%)
3	15 (15,8%)
4	6 (6,3%)
5	2 (2,1%)

consequent beneden de vastgestelde norm van 92% conform getoetste aanvragen. Er is geen sprake van zorgaanbieders die steeds rond de grens van 92% conform bewegen en daardoor telkens wijzigingen in het toetsingsmoment ondervinden. Dit betekent dat het gebruik van een prestatieafhankelijk toetsingsmoment het mogelijk maakt te differentiëren tussen zorgaanbieders die voldoende accuraat zijn in het indienen van aanvragen en zorgaanbieders waarvoor dat niet geldt.

Tenslotte geeft Tabel 5.3 informatie over het aantal rondes waarin zorgaanbieders vooraf zijn getoetst.¹⁶ De gegevens in deze tabel duiden erop dat 11 zorgaanbieders direct aan het begin van het experiment over zijn gegaan op achteraf toetsing en daarna geen wijzigingen in het toetsingsmoment hebben ondervonden. Zij presteren dus consequent boven de 92%-norm. Daarentegen zijn er veel zorgaanbieders die consequent beneden deze norm presteren: 12 zorgaanbieders zijn in alle rondes vooraf getoetst. Van de zorgaanbieders tussen deze twee uitersten in, presteert het grootste gedeelte meestal beneden de grens van 92% conform toetsingen. 43 zorgaanbieders zijn in 7 tot 11 van de rondes in het experiment vooraf getoetst, terwijl maar 29 zorgaanbieders tussen de 1 en 6 rondes vooraf zijn getoetst. Gemiddeld genomen is het prestatieniveau in de HiT prestatie groep gedurende het experiment dus beneden de doelstelling van 92% geweest.

Conclusie (3): *Ongeveer 11% van de zorgaanbieders in de HiT prestatie groep presteert consequent boven de 92%-norm en is vanaf de ingangsdatum van de eerste aanpassingen achteraf getoetst. Het merendeel van de zorgaanbieders presteert echter in de meeste rondes beneden de doelstelling van 92%. Het gebruik van een prestatieafhankelijk toetsingsmoment kan helpen de ‘goede’ en de ‘slechte’ zorgaanbieders te identificeren.*

5.1.2 Aanpassingen in het toetsingspercentage (SIP prestatie groep)

Tabel 5.5 geeft voor iedere aanpassingsronde het gemiddelde gehanteerde toetsingspercentage voor laag en hoog risico SIP aanvragen voor zorgaanbieders in de SIP prestatie groep weer. De tabel laat zien dat het gemiddelde toetsingspercentage in bijna iedere ronde is gestegen. Het gemiddelde lage (hoge) toetsingspercentage is per ronde met ge-

¹⁶Hierin is de eerste periode, van 17 september 2012 tot 15 oktober 2012, niet meegenomen als ronde vooraf getoetst: in deze periode werden alle zorgaanbieders nog vooraf getoetst.

Tabel 5.3: Aantal keer toetsingsmoment vooraf.

Aantal rondes vooraf getoetst	Aantal (%) zorgaanbieders
0	11 (11,6%)
1	9 (9,5%)
2	2 (2,1%)
3	5 (5,3%)
4	6 (6,3%)
5	5 (5,3%)
6	2 (2,1%)
7	7 (7,4%)
8	11 (11,6%)
9	8 (8,4%)
10	7 (7,4%)
11	10 (10,5%)
12	12 (12,6%)

middeld 0,63 (0,34) procentpunt toegenomen.¹⁷

Tabel 5.4: Periodieke aanpassingen in toetsingspercentage.

Aanpassingsronde	Gemiddelde	
	<i>laag</i>	<i>hoog</i>
Ronde 0	6,0 (0,0)	16,0 (0,0)
Ronde 1	5,8 (1,1)	15,8 (1,1)
Ronde 2	7,6 (3,4)	17,5 (3,1)
Ronde 3	8,2 (4,3)	17,9 (3,5)
Ronde 4	8,5 (5,4)	18,0 (4,3)
Ronde 5	9,0 (5,8)	18,3 (4,5)
Ronde 6	9,9 (7,3)	18,6 (5,2)
Ronde 7	10,5 (7,5)	18,9 (5,1)
Ronde 8	11,4 (7,9)	19,4 (5,2)
Ronde 9	12,1 (8,1)	19,8 (5,2)
Ronde 10	12,7 (8,3)	20,1 (5,4)
Ronde 11	12,9 (8,8)	20,0 (5,5)
Ronde 12	13,5 (9,4)	20,1 (5,5)

Standaarddeviatie in haakjes.

Tabel 5.5 illustreert de variatie in toetsingspercentages in meer detail. In deze tabel is per ronde aangegeven welk percentage van alle zorgaanbieders in de SIP prestatie groep een afname of toename in het lage óf het hoge toetsingspercentage heeft ervaren en voor welk percentage zorgaanbieders de toetsingspercentages gelijk zijn gebleven. Deze cijfers illustreren dat in iedere ronde het toetsingspercentage voor iets meer dan de helft van de zorgaanbieders verandert. In de loop van het experiment is voor een steeds kleiner aantal zorgaanbieders een beweging omlaag te zien, terwijl voor een steeds groter percentage zorgaanbieders het toetsingspercentage toeneemt. Bovendien is het percentage zorgaanbieders gerapporteerd voor wie het hoge toetsingspercentage gelijk is aan het

¹⁷Het gemiddelde toetsingspercentage dat wordt gehanteerd voor laag risico SIP aanvragen is sneller gestegen dan dat voor hoog risico SIPs. Dit is het gevolg van het maximum toetsingspercentage dat wordt gebruikt.

maximum van 26% en voor welk percentage zorgaanbieders het lage toetsingspercentage het minimum van 2% heeft bereikt. Het percentage zorgaanbieders met het hoge risico toetsingspercentage op het maximum van 26% is toegenomen tot 39% in de laatste ronde. Anderzijds geldt in de laatste ronde voor slechts 5% van de zorgaanbieders dat het minimum van 2% is bereikt voor het laag risico toetsingspercentage. Dit laatste percentage is aan het begin van het experiment wat hoger geweest (variërend van 11% tot 7% in rondes 4 tot en met 7). Voor geen van de zorgaanbieders voor wie het laag risico SIP toetsingspercentage het minimum bereikt, komt ook het toetsingspercentage voor hoog risico SIP aanvragen op het minimum terecht. Anderzijds geldt dat, in de laatste aanpassingsronde, voor 28% van de zorgaanbieders het laag en hoog risico toetsingspercentage het maximum van 26% hebben bereikt.

Tabel 5.5: Variatie in het toetsingspercentage.

Aanpassingsronde	Omlaag	Gelijk	Omhoog	Op maximum	Op minimum
Ronde 1	37%	46%	18%	0%	0%
Ronde 2	21%	25%	54%	2%	4%
Ronde 3	28%	44%	28%	5%	2%
Ronde 4	33%	44%	23%	12%	11%
Ronde 5	28%	47%	25%	12%	7%
Ronde 6	30%	37%	33%	21%	8%
Ronde 7	18%	54%	28%	23%	7%
Ronde 8	19%	51%	30%	28%	5%
Ronde 9	16%	53%	32%	32%	5%
Ronde 10	25%	44%	32%	37%	4%
Ronde 11	21%	51%	28%	35%	5%
Ronde 12	18%	54%	28%	39%	5%

Gerapporteerd zijn percentages van alle 57 zorgaanbieders in de SIP prestatie groep.

Conclusie (4): *Het gemiddelde toetsingspercentage in de SIP prestatie groep is tijdens het experiment steeds verder toegenomen. Een groep zorgaanbieders van toenemende omvang presteert zodanig beneden de norm van 93,75% conform dat het maximale toetsingspercentage van 26% wordt bereikt. Er is slechts een kleine groep zorgaanbieders die het minimum van 2% bereiken als gevolg van prestatie boven de norm.*

5.2 Effecten van variatie in toetsingspraktijk: grafische analyses

Doel van het experiment is om te kijken wat de effecten van verschillende vormen van toetsingspraktijk zijn. Relevante uitkomsten zijn in dit geval het aantal aanvragen, het aantal toetsingen en de fractie conform getoetste aanvragen. De volgende subsecties geven per uitkomstvariabele een (grafische) beschrijving van de ontwikkeling over tijd in de variabele, zowel voor als tijdens het experiment, uitgesplitst voor de verschillende experiment groepen.¹⁸ Als periode is twee weken gebruikt; dat komt overeen met de

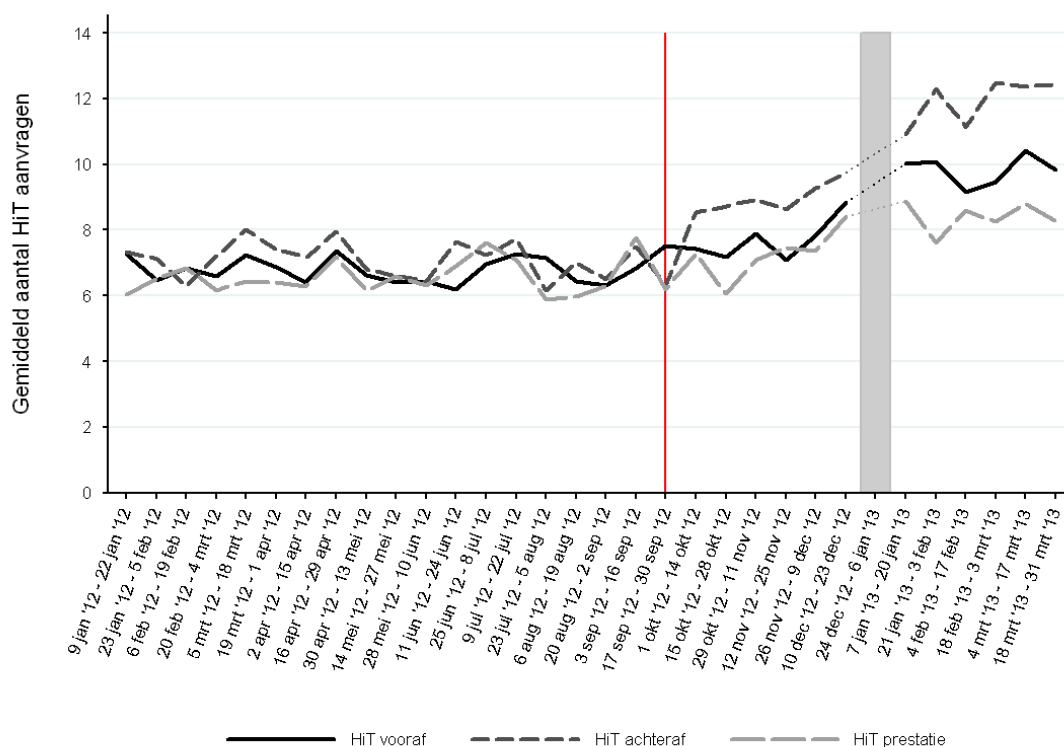
¹⁸Zoals aangegeven in Appendix E zijn er een aantal zorgaanbieders die in de loop van het experiment inactief zijn geworden. In de gerapporteerde figuren zijn de gegevens voor deze zorgaanbieders wel

periode van de aanpassingen in de groepen met een prestatieafhankelijk toetsingsmoment of toetsingspercentage.

5.2.1 Aantal aanvragen

Figuur 5.1 laat allereerst de ontwikkeling in het gemiddeld aantal HiT aanvragen per zorgaanbieder in de verschillende HiT groepen zien. De verticale rode lijn geeft het moment waarop het experiment is gestart (17 september 2012). De grijze band in deze, en volgende, figuren bevat de periode rond Kerst en Oud en Nieuw (24 december 2012 - 6 januari 2013). In deze periode zijn er veel minder aanvragen ingediend, met als gevolg minder toetsingen. Dit veroorzaakt een veel lager niveau van de grafiek in deze periode, wat afleidt van de ontwikkeling waarin we daadwerkelijk geïnteresseerd zijn. Om dit laatste beter zichtbaar te maken, is in de figuren voor aantal aanvragen en aantal toetsingen een geïnterpoleerde waarde voor deze periode berekend en weergegeven.

Figuur 5.1: Gemiddeld aantal HiT aanvragen, per HiT groep.



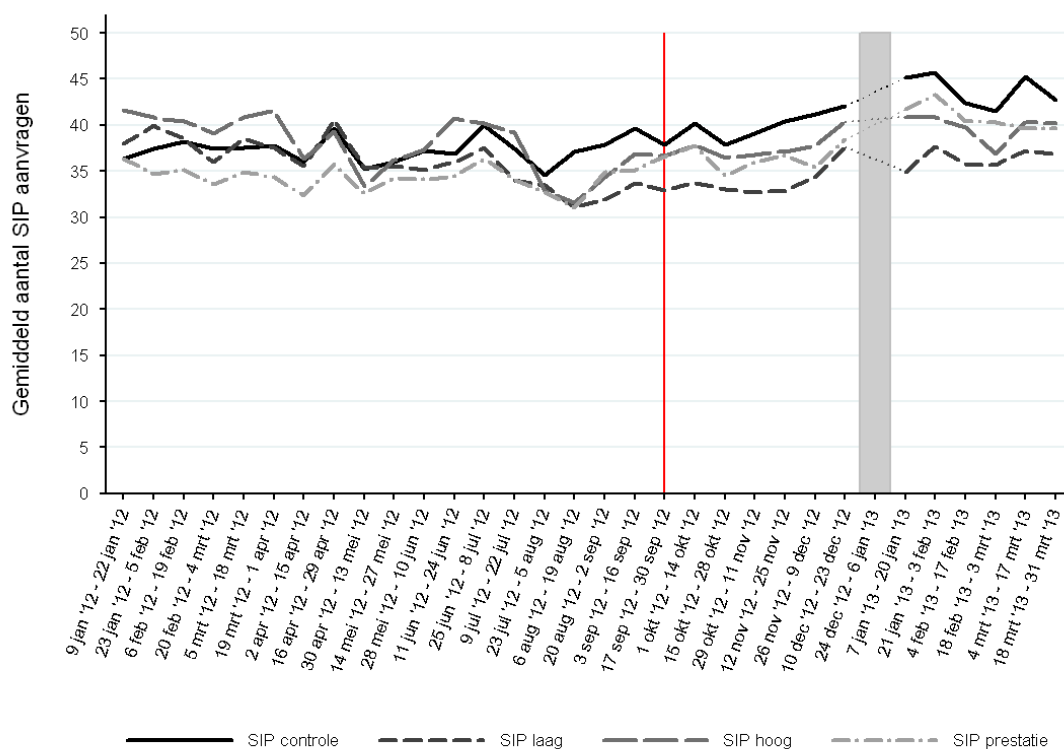
Uit Figuur 5.1 blijkt dat sinds de start van het experiment het gemiddeld aantal aanvragen in de HiT achteraf groep is toegenomen, ook ten opzichte van het aantal aanvragen in de controlegroep en de HiT prestatie groep. Achteraf toetsing, en dus het ontbreken van bijstellingsmogelijkheden wanneer een aanvraag niet conform blijkt te zijn, lijkt gepaard

meegenomen. Het weglaten van deze zorgaanbieders leidt niet tot verschillen in de resultaten. In de regressieanalyses in sectie 6.3 zullen ter controle ook modellen worden geschat op gegevens waaruit deze zorgaanbieders zijn verwijderd.

te gaan met een toename in het aantal ingediende aanvragen. De aanwezige dreiging van potentiële vooraf toetsing in de HiT prestatie groep lijkt een effect op het aantal aanvragen zoals geobserveerd in de HiT achteraf groep tegen te werken. Het gemiddeld aantal aanvragen in de HiT prestatie groep blijft namelijk gelijk lopen met het aantal aanvragen dat wordt ingediend door zorgaanbieders in de controlegroep.

In Figuur 5.2 is het gemiddeld aantal aanvragen per zorgaanbieder in de diverse SIP groepen weergegeven. In de SIP groepen zijn nauwelijks veranderingen in het aantal aanvragen op te merken sinds de start van het experiment. Het aantal aanvragen ontwikkelt zich in alle SIP groepen op een soortgelijke manier sinds de start van het experiment, alhoewel de spreiding tussen groepen iets groter lijkt te worden vanaf het begin 2013.

Figuur 5.2: Gemiddeld aantal SIP aanvragen, per SIP groep.



Conclusie (5a): *Het gemiddeld aantal HiT aanvragen per zorgaanbieder in de HiT achteraf groep is sinds de start van het experiment toegenomen. Er is tevens een licht stijgende trend in het aantal aanvragen in de controlegroep en de HiT prestatie groep, maar daar wijkt de trend in de HiT achteraf groep duidelijk van af.*

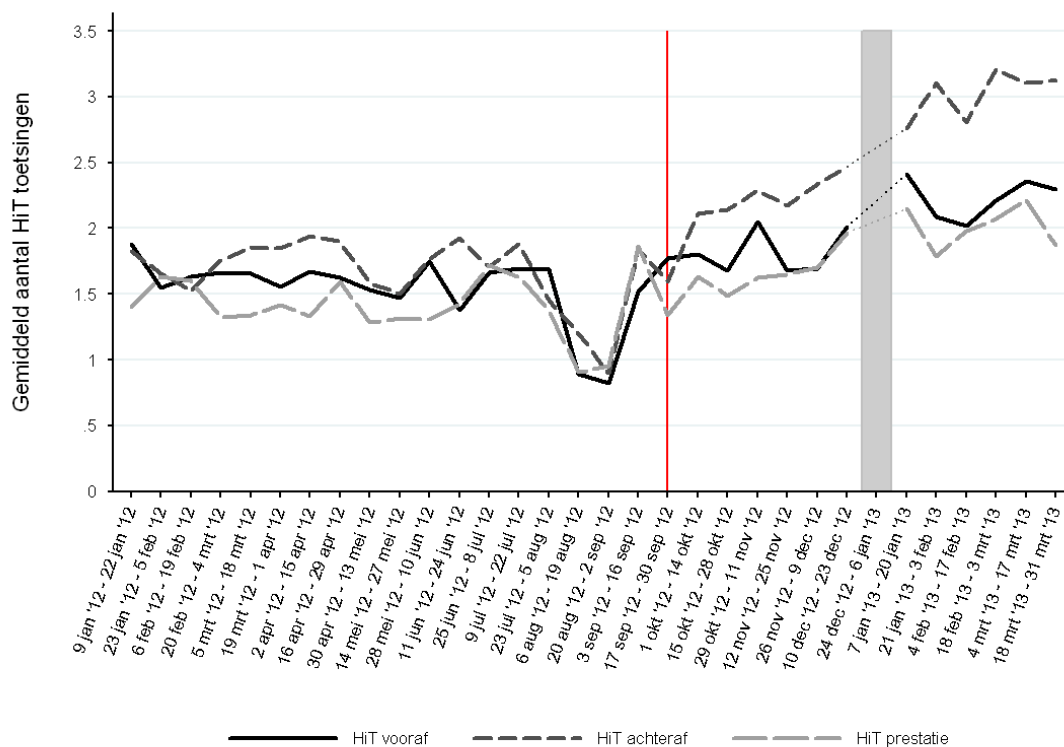
Conclusie (5b) *Voor geen enkele vorm van variatie in het toetsingspercentage is er een duidelijk effect op het aantal ingediende SIP aanvragen te bemerken.*

5.2.2 Aantal toetsingen

Figuur 5.3 en Figuur 5.4 geven een beeld van de ontwikkelingen in het gemiddeld aantal getoetste aanvragen per zorgaanbieder in de verschillende experiment groepen. Het aantal toetsingen wordt direct bepaald door enerzijds het aantal ingediende aanvragen en anderzijds het gehanteerde toetsingspercentage. Het toetsingspercentage bij HiT aanvragen is ongewijzigd in het experiment, maar Figuur 5.1 heeft laten zien dat het aantal aanvragen is toegenomen. Als gevolg daarvan is ook een toename te zien in het aantal toetsingen, vooral in de HiT achteraf groep.

Figuur 5.4 laat zien dat het aantal SIP toetsingen in de SIP controlegroep niet is gewijzigd sinds de start van het experiment. Daarentegen is het aantal toetsingen in de SIP laag toetsingspercentage groep afgenomen en in de hoog toetsingspercentage groep toegenomen, als direct gevolg van de veranderingen in het toetsingspercentage. Zoals eerder is weergegeven in Tabel 5.5, is het toetsingspercentage in de SIP prestatie groep gedurende het experiment toegenomen. Dit vertaalt zich in een geleidelijke toename van het aantal toetsingen in deze groep. Aan het eind van de looptijd van het experiment is het aantal toetsingen in deze laatste groep zelfs aanzienlijk hoger dan het aantal toetsingen in de hoog toetsingspercentage groep.

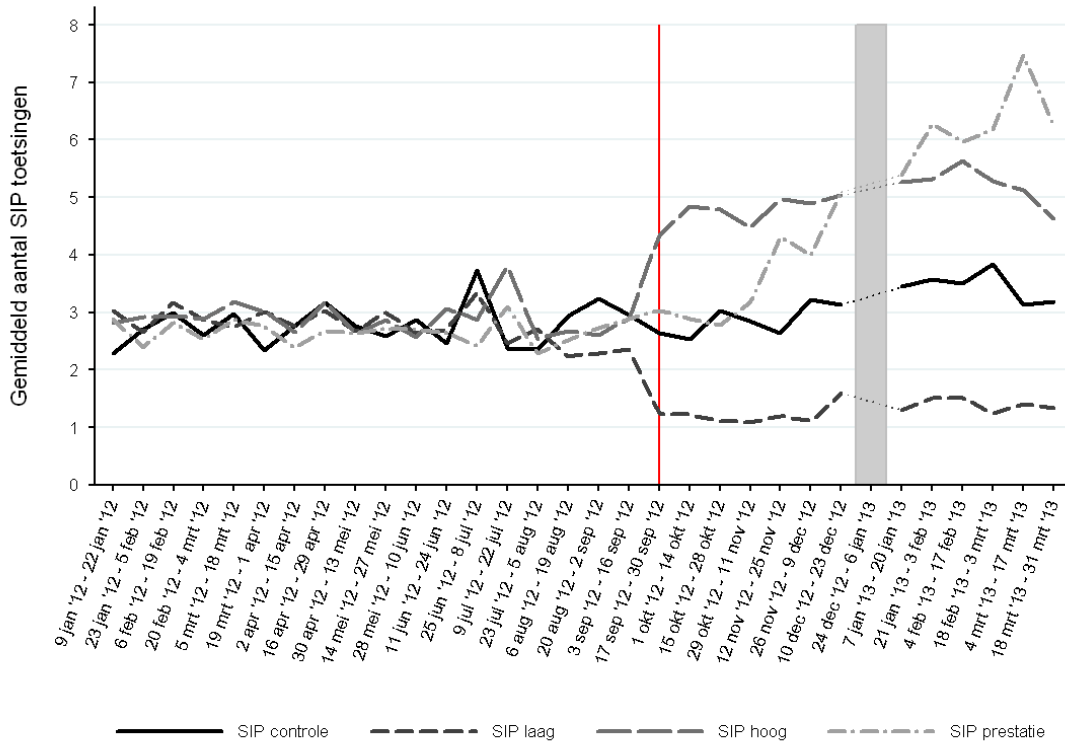
Figuur 5.3: Gemiddeld aantal HiT toetsingen, per HiT groep.



Conclusie (6): *Het patroon in het gemiddeld aantal toetsingen in HiT groepen vloeit voort uit de eerder beschreven ontwikkeling in het aantal aanvragen: het aantal toetsingen*

in de HiT achteraf groep neemt toe. De geleidelijke toename van het gemiddelde toetsingspercentage in de SIP prestatie groep vertaalt zich in een stijgende trend in het aantal toetsingen in deze groep.

Figuur 5.4: Gemiddeld aantal SIP toetsingen, per SIP groep.



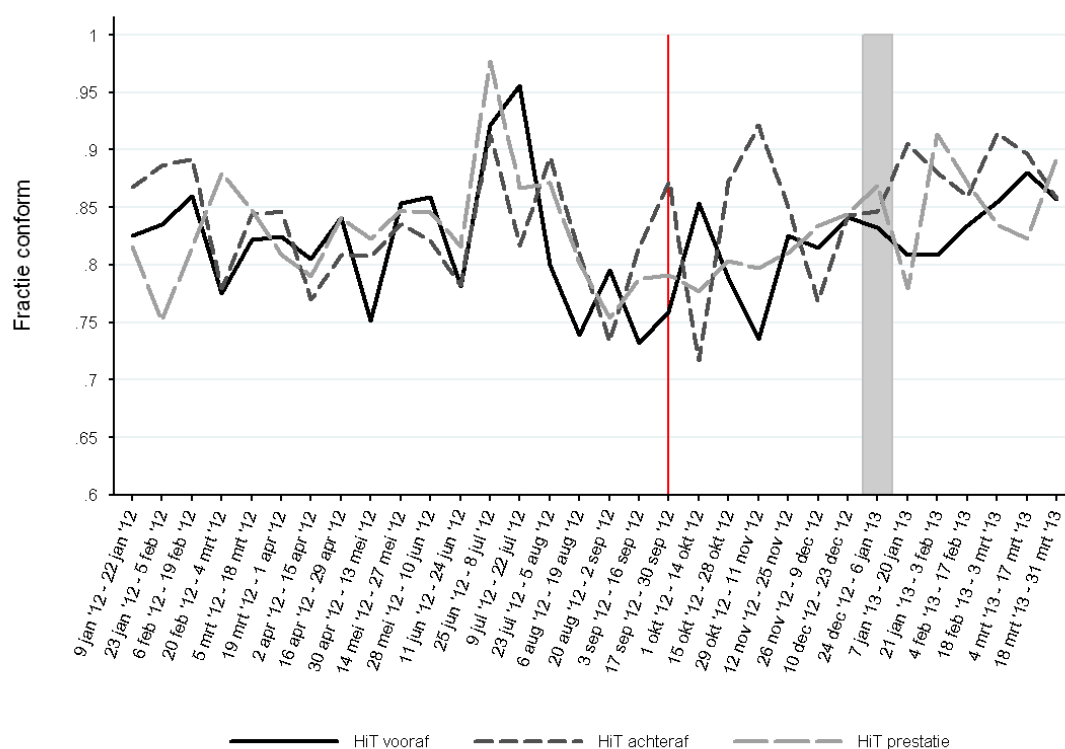
5.2.3 Fractie conform getoetste aanvragen

Figuur 5.5 en Figuur 5.6 geven inzicht in de ontwikkeling in de gemiddelde fractie conform getoetste aanvragen over zorgaanbieders per groep.¹⁹ Uit Figuur 5.5 kan worden opgemaakt dat het conformpercentage in de drie HiT groepen redelijk gelijk op blijft lopen, ook na de start van het experiment. Opvallend is wel dat vanaf de start van het experiment de fractie conform getoetste aanvragen een wat stijgende trend vertoont, terwijl dat voor aanvang van het experiment niet het geval was. Omdat dit in alle groepen optreedt, kan dit echter niet direct worden toegewezen aan een specifieke vorm van toetsingspraktijk. Figuur 5.6 laat geen effect zien van de verschillende vormen van variatie in het toetsingspercentage op het behaalde conformpercentage. De ontwikkeling in de fractie conform getoetste aanvragen voor en tijdens het experiment is nagenoeg hetzelfde en er zijn geen duidelijke afwijkingen tussen groepen.

¹⁹In het berekenen van het periodieke gemiddelde conformpercentage zijn alleen die zorgaanbieders meegenomen voor wie in de betreffende periode een aanvraag is getoetst. Dit betekent dat het aantal observaties waarop de gemiddelden zijn gebaseerd, per periode en per groep varieert.

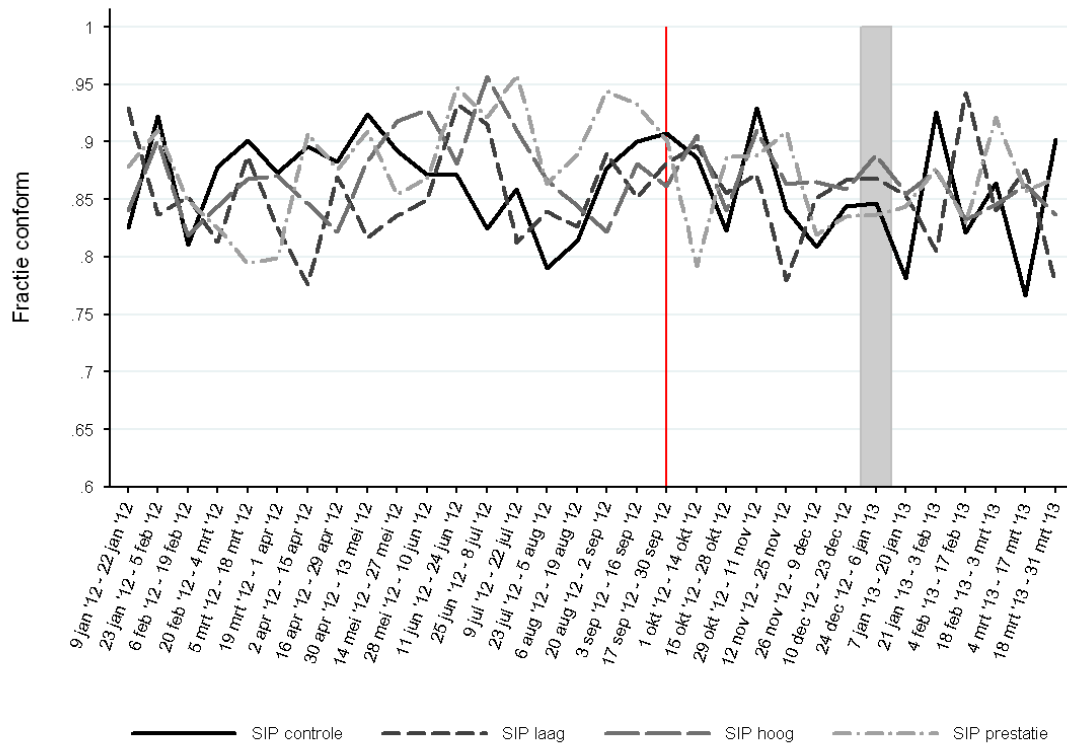
Conclusie (7): *Het hanteren van prestatieafhankelijke toetsingspercentages resulteert niet in selectie van slechtere SIP aanvragen voor toetsing. Bovendien zien we geen effect van de verschillende vormen van variatie in toetsingsmoment en toetsingspercentage op de behaalde gemiddelde fractie conform getoetste aanvragen.*

Figuur 5.5: Gemiddelde fractie conform getoetste aanvragen, per HiT groep.



Uit de bovenstaande figuren blijkt al dat er in alle groepen veel variatie is in de gemiddelde fractie conform getoetste aanvragen. Deze ruis in het conformpercentage kan er voor zorgen dat kleine effecten van de verschillende vormen van variatie in toetsingspraktijk niet worden opgemerkt in statistische analyses. Omdat vóór de aanvang van het experiment er wat verschillen waren in conformpercentages voor diverse typen HiT aanvragen, splitsen we de ontwikkeling in de gemiddelde fractie conform getoetste aanvragen ook uit naar type HiT aanvragen. Figuren van de ontwikkeling in het gemiddelde conformpercentage per type HiT zijn terug te vinden in Appendix F. Dit geeft geen nieuwe inzichten ten opzichte van de voorgaande resultaten.

Figuur 5.6: Gemiddelde fractie conform getoetste aanvragen, per SIP groep.



6 Regressieanalyses

In de voorgaande sectie zijn beschrijvende analyses gepresenteerd. Om vast te stellen of er significante effecten zijn van de overgang op een andere vorm van toetsingspraktijk op de kwantiteit en/of kwaliteit van aanvragen, worden in deze sectie de resultaten van verscheidene regressieanalyses gepresenteerd. In sectie 6.1 worden enkele basisanalyses gepresenteerd. Allereerst wordt een basismodel geschat waarin simpelweg wordt gekeken of er sprake is van een effect van de variatie in toetsingspraktijk op het aantal aanvragen, het aantal toetsingen of de fractie conform getoetste aanvragen. Vervolgens wordt gekeken naar de aanwezigheid van substitutie-effecten, dat wil zeggen: heeft de variatie in toetsingspraktijk voor HiT of SIP aanvragen gevolgen voor het soort aanvragen (de aanvragenmix) dat door zorgaanbieders wordt ingediend. Vervolgens wordt in sectie 6.2 onderzocht of sprake is van heterogene effecten, dat wil zeggen verschil in effect van de variatie in de toetsingspraktijk op diverse soorten zorgaanbieders, bijvoorbeeld naar omvang van de zorgaanbieder in termen van het aantal aanvragen en naar aanvraagmix (de verhouding waarin diverse soorten aanvragen worden ingediend). Tenslotte wordt in sectie 6.3 gekeken of de resultaten robuust zijn, onder meer voor het verwijderen van inactieve zorgaanbieders uit de data.

6.1 Basisanalyses

Om de omvang en significantie van effecten te bepalen, wordt een multivariaat regressiemodel geschat waarin rekening wordt gehouden met systematische verschillen tussen zorgaanbieders en tussen tijdsperiodes.²⁰ We schatten aparte modellen voor het gedeelte van het experiment waarin het toetsingsmoment wordt gevarieerd en het gedeelte van het experiment waarin de toetsingspercentages worden aangepast. Het model voor het experiment met variatie in het toetsingsmoment is weergegeven in vergelijking (1).

$$Y_{i,t} = \alpha_i + \gamma_t + \delta_{achteraf} T_{i,t}^{achteraf} + \delta_{prestatie} T_{i,t}^{prestatie} + \varepsilon_{i,t} \quad (1)$$

waar i een index is voor zorgaanbieders en t een index voor tijd (gemeten in eenheden van twee weken).²¹ De term α_i is opgenomen om rekening te houden met systemati-

²⁰We maken gebruik van de zogenaamde *difference-in-differences* methode. Dit is een methode om het effect van een bepaalde ingreep, in dit geval variatie in het toetsingsmoment of het toetsingspercentage, te bepalen, door een vergelijking te maken tussen de ontwikkeling, vóór en na die ingreep, in een controlegroep en een groep die de ingreep ook daadwerkelijk ondervindt. We formuleren daarvoor een *panel fixed effects model* om rekening te kunnen houden met systematische verschillen tussen zorgaanbieders en tussen tijdsperiodes.

²¹Een soortgelijk model wordt geschat voor het experiment met variatie in het toetsingspercentage:

$$Y_{i,t} = \alpha_i + \gamma_t + \delta_{laag} T_{i,t}^{laag} + \delta_{hoog} T_{i,t}^{hoog} + \delta_{prestatie} T_{i,t}^{prestatie} + \varepsilon_{i,t} \quad (2)$$

sche verschillen tussen zorgaanbieders die constant zijn over tijd, bijvoorbeeld kenmerken van een zorgaanbieder als de gemiddelde grootte in termen van het aantal ingediende aanvragen (een zogenaamd zorgaanbieder *fixed effect*). De term γ_t is opgenomen om te corrigeren voor systematische verschillen over tijd, zoals de eerder genoemde scherpe afname in het aantal aanvragen in de periode rond Kerst en Oud en Nieuw (een zogenaamd tijd *fixed effect*). De variabele $T_{i,t}^{achteraf}$ ($T_{i,t}^{prestatie}$) geeft aan of zorgaanbieder i op tijdstip t in de achteraf (prestatie) groep zat.²² De parameters waarin we geïnteresseerd zijn, zijn de δ 's. Deze parameters meten het effect van variatie in toetsingspraktijk: $\delta_{achteraf}$ meet het effect van achteraf toetsing ten opzichte van vooraf toetsing en $\delta_{prestatie}$ geeft het effect van het gebruik van een prestatieafhankelijk toetsingsmoment. Tenslotte is $Y_{i,t}$ de uitkomstvariabele, dat wil zeggen de variabele waarvoor we willen kijken of deze wordt beïnvloed door de variatie in toetsingspraktijk. We bekijken verschillende relevante uitkomstvariabelen, te weten het aantal aanvragen, het aantal toetsingen en de fractie conform getoetste aanvragen.

6.1.1 Aantal aanvragen

Allereerst nemen we het aantal ingediende aanvragen als uitkomstvariabele. We normaliseren het aantal ingediende aanvragen zodanig dat de geschatte effecten dienen te worden geïnterpreteerd als een effect op het aantal ingediende aanvragen per week. Om een indicatie te geven van de economische significantie van het effect, is tevens het gemiddelde in de pre-experiment periode in de controlegroep gerapporteerd.

Tabel 6.1 geeft resultaten voor variatie in het toetsingsmoment (panel A) en het toetsingspercentage (panel B).²³ De resultaten in panel A laten een positief, maar insignificant, effect zien van de overgang op achteraf toetsing en een negatief effect van de overgang op het gebruik van een prestatieafhankelijk toetsingsmoment. Indien we de effectgrootte vergelijken met het gemiddeld aantal ingediende aanvragen in de controlegroep voor aanvang van het experiment, zien we dat het geschatte effect vrij groot is: een procentuele toename van 18% (gebaseerd op de cijfers in de tweede kolom) in het gemiddeld aantal aanvragen dat per week wordt ingediend. We hebben tevens getoetst of het effect van achteraf toetsing significant verschillend is van het effect van het gebruik van een prestatieafhankelijk toetsingsmoment. Deze toets wijst uit dat de effecten op het aantal aanvragen voor deze twee groepen inderdaad significant verschillend zijn.

²² $T_{i,t}^{achteraf}$ ($T_{i,t}^{prestatie}$) is gelijk aan 1 voor een zorgaanbieder in achteraf (prestatie), in een periode t gedurende het experiment (vanaf 17 september 2012). In andere gevallen is deze indicator gelijk aan 0. Dat betekent dus dat voor zorgaanbieders in de controlegroep beide indicatoren altijd gelijk aan 0 zijn.

²³Voor de regressieanalyses zijn zorgaanbieders die zijn gedefinieerd als inactief op basis van de definitie gepresenteerd in Appendix E niet meegenomen in de periodes aan het begin en het eind van de observatieperiode waarin ze inactief waren. We hebben de basisanalyses in sectie 6.1 ook uitgevoerd op een *balanced panel* waarin het aantal aanvragen en toetsingen voor deze zorgaanbieders op 0 is gezet in de periodes waarin ze inactief waren. In dat geval vinden we een significant effect (zij het slechts bij een 10% significantieniveau) op het aantal HiT aanvragen ingediend door zorgaanbieders in de achteraf toetsing groep.

Om meer in detail te bekijken waar het effect op het aantal ingediende aanvragen vandaan komt, kunnen we een uitsplitsing maken naar type HiT aanvraag. Het overgrote deel (ruim 70%) van de ingediende HiT aanvragen is type 1. Schattingsresultaten voor het effect op het aantal HiT type 1 aanvragen en het aantal andere type HiT aanvragen zijn gerapporteerd in de derde en vierde kolom van Tabel 6.1. We zien dat er sprake is van een positief effect van de overgang op achteraf toetsing op het aantal ingediende HiT type 1 aanvragen, terwijl er een negatief effect is op het aantal ander soort HiT aanvragen dat wordt ingediend. Het positieve effect op het aantal HiT type 1 aanvragen is vrij groot: het betreft een toename van 29% ten opzichte van het gemiddelde in de controlegroep voor aanvang van het experiment. Bovendien is dit verschil significant bij gebruik van een significantieniveau van 10%. Voor de overgang op het gebruik van een prestatieafhankelijk toetsingsmoment zien we een significant negatief effect op het aantal HiT aanvragen anders dan type 1 dat wordt ingediend. Het gaat hier om een 34% afname ten opzichte van het gemiddelde in de controlegroep voor aanvang van het experiment. Wel dient hierbij te worden opgemerkt dat het gemiddelde aantal ander soort HiT aanvragen dat werd ingediend door zorgaanbieders in de groep met prestatieafhankelijk toetsingsmoment ook voor aanvang van het experiment al een stuk lager lag dan in de controlegroep.²⁴

Tabel 6.1: Schattingsresultaten voor het effect van variatie in de toetsingspraktijk op het aantal aanvragen.

Panel A: Variatie in het toetsingsmoment.						
	# HiT		# HiT type 1		# ander soort HiT	
HiT achteraf	0,624	(0,407)	0,709*	(0,366)	-0,085	(0,185)
HiT prestatie	-0,325	(0,210)	0,001	(0,140)	-0,326**	(0,158)
gem. controlegr.	3,381	(0,092)	2,435	(0,086)	0,946	(0,037)
R^2	0,752		0,801		0,608	
aantal obs.	9.286		9.286		9.286	
aantal ZABs	299		299		299	
Panel B: Variatie in het toetsingspercentage.						
	# SIP					
SIP laag	-1,291	(1,116)				
SIP hoog	-1,480	(1,127)				
SIP prestatie	-0,614	(0,972)				
gem. controlegr.	18,681	(0,622)				
R^2	0,921					
aantal obs.	7.067					
aantal ZABs	226					

Cluster-robuste standaardfouten tussen haakjes. Het gerapporteerde gemiddelde in de controlegroep is het gemiddelde per week, gemeten over de periode voor de start van het experiment. * significant op significantieniveau van 10%, ** significant op significantieniveau van 5%, *** significant op significantieniveau van 1%.

Voor iedere 2-weeks periode is een dummy opgenomen (*time fixed effects*). De geschatte coëfficiënt voor deze tijdsdummies zijn niet gerapporteerd.

²⁴Het gemiddelde voor aanvang van het experiment was 0,792 met een standaardfout van 0,039.

Panel B van Tabel 6.1 geeft informatie over het effect van variatie in het toetsingspercentage op het aantal ingediende SIP aanvragen. Alle geschatte effecten zijn negatief, maar we zien dat geen van de vormen van variatie in het toetsingspercentage een significant effect heeft op het aantal ingediende SIP aanvragen.

De schattingen in Tabel 6.1 geven het gemiddelde effect over de gehele experiment periode. We zagen echter eerder in figuren 5.1 en 5.2 in sectie 5 dat er een verschil in trend lijkt te zijn tussen de periode september tot december 2012 en de periode januari tot april 2013. Dit kan bijvoorbeeld te maken hebben met een leereffect: het kost tijd voordat zorgaanbieders precies door hebben hoe de toetsingspraktijk in de nieuwe situatie wordt bepaald. Bovendien was sprake van minder frequente aanpassingen aan het begin van het experiment (zie sectie 3). Om die reden splitsen we het *treatment* effect uit naar een effect in 2012 en een effect in 2013. De resultaten in Tabel 6.2 laten zien dat er alleen een significant effect is van de overgang op achteraf toetsing in 2013. In 2012 is er een klein positief effect, maar dat effect is niet significant verschillend van nul. Daarentegen is er voor de experiment periode in 2013 een groot positief en significant (zij het bij gebruik van een significantieniveau van 10%) effect op het aantal ingediende HiT aanvragen. De omvang van het effect is groter dan eerder werd gerapporteerd in Tabel 6.1, dat heeft te maken met het feit dat die regressieresultaten het effect baseren op het gemiddeld aantal aanvragen over de hele periode van het experiment, terwijl het aantal aanvragen in 2013 hoger ligt dan daarvoor. De resultaten in de onderstaande tabel laten zien dat (in 2013) sprake is van een stijging van 31% in het gemiddeld aantal aanvragen dat per week wordt ingediend. Bovendien zien we dat het effect van overgang op het gebruik van een prestatieafhankelijk toetsingsmoment in geen van de periodes significant is. De omvang van de effecten neemt wel toe over tijd (-0,52 in 2013 ten opzichte van slechts -0,18 in 2012) maar kan in geen van de gevallen worden onderscheiden van het ontbreken van een effect.

De uitsplitsing van het *treatment effect* in de periode 2012 en 2013 geeft voor de variatie in het toetsingspercentage een zelfde beeld als de resultaten in Tabel 6.1. De effecten van de diverse vormen van variatie in het toetsingspercentage zijn allen negatief (daling in het aantal ingediende aanvragen), maar kunnen statistisch gezien niet worden onderscheiden van nul.

Conclusie (8): *De econometrische analyses laten een (insignificant) positief effect van achteraf toetsing op het aantal ingediende HiT aanvragen zien. Er is sprake van een significant verschil in het effect voor de groepen met achteraf toetsing en prestatieafhankelijk toetsingsmoment. Wanneer we in de periode van het experiment nader onderscheid maken tussen het effect in 2012 en het effect in 2013, vinden we voor 2013 een groter en significant effect op het aantal ingediende aanvragen in de achteraf toetsing groep. De omvang van dit effect is groot: we vinden een 31% toename in het aantal ingediende HiT aanvragen. We vinden geen significante effecten van variatie in toetsingspercentage op het aantal ingediende SIP aanvragen.*

Tabel 6.2: Schattingsresultaten voor het effect van variatie in de toetsingspraktijk op het aantal aanvragen; uitsplitsing van de experiment periode.

	Variatie in toetsingsmoment		Variatie in toetsingspercentage	
	# HiT		# SIP	
HiT achteraf, 2012	0,309	(0,288)		
HiT achteraf, 2013	1,056*	(0,621)		
HiT prestatie, 2012	-0,178	(0,160)		
HiT prestatie, 2013	-0,519	(0,347)		
SIP laag, 2012			-0,970	(1,030)
SIP laag, 2013			-1,720	(1,313)
SIP hoog, 2012			-1,407	(1,102)
SIP hoog, 2013			-1,577	(1,270)
SIP prestatie, 2012			-0,636	(0,894)
SIP prestatie, 2013			-0,582	(1,188)
gem. controlegr.	3,381	(0,092)	18,681	(0,622)
R^2	0,753		0,921	
aantal obs.	9.286		7.067	
aantal ZABs	299		226	

Cluster-robuste standaardfouten tussen haakjes. Het gerapporteerde gemiddelde in de controlegroep is het gemiddelde per week, gemeten over de periode voor de start van het experiment. * significant op significantieniveau van 10%, ** significant op significantieniveau van 5%, *** significant op significantieniveau van 1%.

Voor iedere 2-weeks periode is een dummy opgenomen (*time fixed effects*). De geschatte coëfficiënt voor deze tijdsdummies zijn niet gerapporteerd.

6.1.2 Aantal toetsingen

Naast een effect op het aantal aanvragen, zoals onderzocht in de voorgaande sectie, valt ook een effect op het aantal toetsingen te verwachten. Deels is dit mechanisch het gevolg van de opgelegde variatie in de toetsingspraktijk, maar voor de zorgaanbieders die deel uit maken van een groep waarin het toetsingsregime afhangt van prestatie, kan het effect op het aantal toetsingen interessante inzichten geven.

Tabel 6.3 geeft een eerste beeld; de eerste kolom beschrijft de effecten van variatie in het toetsingsmoment op het aantal toetsingen, terwijl de tweede kolom de effecten van diverse vormen van variatie in het toetsingspercentage op het aantal SIP toetsingen weergeeft. We zien een significant positief effect op het aantal getoetste HiT aanvragen in de achteraf toetsing groep. Dit is een direct gevolg van de gevonden toename in het aantal ingediende HiT aanvragen, aangezien het gebruikte toetsingspercentage gelijk is gebleven. Ten opzichte van het gemiddeld aantal toetsingen per week in de controlegroep voor aanvang van het experiment gaat het om een 29% toename. We vinden geen significant effect op het aantal toetsingen in de groep met een prestatieafhankelijk toetsingsmoment. Uit een toets voor verschil in effect van achteraf toetsing ten opzichte van prestatieafhankelijk toetsingsmoment blijkt dat sprake is van een significant verschil tussen deze twee groepen.

De tweede kolom van Tabel 6.3 laat zien dat er sprake is van significante effecten op het

aantal toetsingen voor alle experimentele SIP groepen. Natuurlijk zijn de effecten voor de groepen met laag en hoog toetsingspercentage het mechanische gevolg van de opgelegde aanpassing in het toetsingspercentage. We zien een significante 57% afname in het aantal toetsingen in de groep waarvoor het toetsingspercentage is verlaagd en een significante 62% toename in het aantal toetsingen in de groep waarvoor het toetsingspercentage is verhoogd. Tenslotte zien we ook een significante toename van het aantal getoetste SIP aanvragen voor zorgaanbieders in de groep met prestatieafhankelijk toetsingspercentage. Dit is een gevolg van de geleidelijke toename in het gemiddelde toetsingspercentage in deze groep, zoals is gerapporteerd in Tabel 5.5. De stijging in het aantal toetsingen in deze groep is bijna even groot als in de groep met het hoge toetsingspercentage, ondanks de geleidelijke stijging in het toetsingspercentage in de prestatie groep. Dit wordt mede veroorzaakt doordat voor 28% van de zorgaanbieders het laag risico SIP toetsingspercentage het maximum van 26% heeft bereikt, terwijl het gehanteerde toetsingspercentage voor dit type SIP in de groep met een hoog toetsingspercentage slechts 10% is.

Tabel 6.3: Schattingsresultaten voor het effect van variatie in de toetsingspraktijk op het aantal toetsingen.

	Toetsingsmoment		Toetsingspercentage	
	# HiT toetsingen		# SIP toetsingen	
HiT achteraf	0,221**	(0,106)		
HiT prestatie	-0,016	(0,063)		
SIP laag			-0,799***	(0,123)
SIP hoog			0,861***	(0,165)
SIP prestatie			0,835***	(0,200)
gem. controlegr.	0,769	(0,025)	1,394	(0,052)
R^2	0,602		0,672	
aantal obs.	9.286		7.067	
aantal ZABs	299		226	

Cluster-robuste standaardfouten tussen haakjes. Het gerapporteerde gemiddelde in de controlegroep is het gemiddelde per week, gemeten over de periode voor de start van het experiment. * significant op significantieniveau van 10%, ** significant op significantieniveau van 5%, *** significant op significantieniveau van 1%.

Voor iedere 2-weeks periode is een dummy opgenomen (*time fixed effects*). De geschatte coëfficiënt voor deze tijdsdummies zijn niet gerapporteerd.

Net als in Tabel 6.2 kunnen we ook voor het aantal toetsingen een model schatten waarin de experiment periode wordt gesplitst in 2012 en 2013. Uit de figuren voor het aantal toetsingen (Figuur 5.3 en Figuur 5.4 in sectie 5.2.2) komt namelijk een soortgelijk patroon naar voren als voor het aantal aanvragen, met een bijbehorend verschil in de trend in 2012 en in 2013. De resultaten in Tabel 6.4 geven aan dat de toename in het aantal toetsingen in de achteraf toetsing groep alleen significant is in 2013, net als het effect op het aantal aanvragen. Wanneer we in de tweede kolom kijken naar het effect van variatie in het toetsingspercentage op het aantal toetsingen, zien we dat de effecten in alle periodes significant zijn, maar dat de effecten in absolute waarde groter zijn in 2013 dan in 2012. Dit verschil is significant voor de SIP prestatie groep, wat duidt op de

geobserveerde geleidelijke toename in het toetsingspercentage in deze groep. Het effect in 2013 in deze prestatie groep is ook groot in omvang: ten opzichte van het gemiddeld aantal toetsingen in de controlegroep voor aanvang van het experiment is sprake van een verdubbeling in het aantal toetsingen.

Tabel 6.4: Schattingsresultaten voor het effect van variatie in de toetsingspraktijk op het aantal toetsingen; uitsplitsing van de experiment periode.

	Toetsingsmoment		Toetsingspercentage	
	# HiT toetsingen		# SIP toetsingen	
HiT achteraf, 2012	0,113	(0,080)		
HiT achteraf, 2013	0,370**	(0,160)		
HiT prestatie, 2012	-0,011	(0,056)		
HiT prestatie, 2013	-0,020	(0,095)		
SIP laag, 2012			-0,695***	(0,126)
SIP laag, 2013			-0,937***	(0,149)
SIP hoog, 2012			0,853***	(0,154)
SIP hoog, 2013			0,871***	(0,211)
SIP prestatie, 2012			0,411***	(0,124)
SIP prestatie, 2013			1,406***	(0,351)
gem. controlegr.	0,769	(0,025)	1,394	(0,052)
R^2	0,603		0,678	
aantal obs.	9.286		7.067	
aantal ZABs	299		226	

Cluster-robuuste standaardfouten tussen haakjes. Het gerapporteerde gemiddelde in de controlegroep is het gemiddelde per week, gemeten over de periode voor de start van het experiment. * significant op significantieniveau van 10%, ** significant op significantieniveau van 5%, *** significant op significantieniveau van 1%.

Voor iedere 2-weeks periode is een dummy opgenomen (*time fixed effects*). De geschatte coëfficiënt voor deze tijdsdummies zijn niet gerapporteerd.

Conclusie (9): *Als gevolg van de toename in het aantal ingediende HiT aanvragen, vinden we ook een toename in het aantal uitgevoerde toetsingen in de achteraf toetsing groep. De effecten van variatie in toetsingspercentage op het aantal getoetste SIP aanvragen zijn een direct (mechanisch) gevolg van de opgelegde variatie. Interessant is de grote en significante toename in het aantal SIP toetsingen in de prestatie groep.*

6.1.3 Fractie conform getoetste aanvragen

In het voorgaande hebben we bekeken of sprake is van effecten op kwantiteit. Er kan ook worden onderzocht of er een effect op kwaliteit bestaat. Kwaliteit kan hierbij worden gemeten als de fractie conform getoetste aanvragen. De resultaten voor het schatten van de modellen in vergelijkingen (1) en (2) met als uitkomstvariabele de fractie conform getoetste aanvragen, zijn weergegeven in Tabel 6.5.²⁵

²⁵Deze resultaten zijn op een kleiner aantal observaties gebaseerd, dit omdat in periodes waarin een zorgaanbieder geen toetsingen heeft gehad, het conformpercentage niet gedefinieerd is.

Tabel 6.5: Schattingsresultaten voor het effect van variatie in de toetsingspraktijk op de fractie conform getoetste aanvragen.

	Variatie in toetsingsmoment		Variatie in toetsingspercentage	
	fractie conform getoetste HiTs		fractie conform getoetste SIPs	
HiT achteraf	0,027	(0,020)		
HiT prestatie	0,002	(0,020)		
SIP laag			-0,001	(0,022)
SIP hoog			0,000	(0,021)
SIP prestatie			-0,000	(0,020)
gem. controlegr.	0,823	(0,009)	0,868	(0,010)
R^2	0,105		0,150	
aantal obs.	5.727		4.857	
aantal ZABs	299		224	

Cluster-robuste standaardfouten tussen haakjes. Het gerapporteerde gemiddelde in de controlegroep is het gemiddelde per week, gemeten over de periode voor de start van het experiment. * significant op significantieniveau van 10%, ** significant op significantieniveau van 5%, *** significant op significantieniveau van 1%.

Voor iedere 2-weeks periode is een dummy opgenomen (*time fixed effects*). De geschatte coëfficiënt voor deze tijdsdummies zijn niet gerapporteerd.

De resultaten in Tabel 6.5 laten zien dat er geen sprake is van een significant effect op de fractie conform getoetste aanvragen. De geschatte effecten zijn allemaal erg klein en insignificant. Alleen het effect op de fractie conform getoetste aanvragen in de achteraf toetsing groep is iets groter (2,7 procentpunt toename ten opzichte van een conformpercentage van 82,3% voor aanvang van het experiment in de controlegroep), maar de standaardfout is eveneens groot. Een toets voor verschil in effect voor de achteraf en prestatie groep wijst uit dat geen sprake is van een significant verschil tussen deze groepen. Het is opmerkelijk dat, ondanks de toename in het aantal ingediende HiT aanvragen (zie sectie 6.1.1), geen sprake is van een (negatief) effect op de fractie conform getoetste aanvragen.²⁶

Conclusie (10): *Zowel voor variatie in het toetsingspercentage als in het toetsingsmoment vinden we geen significante effecten op de behaalde fractie conform getoetste aanvragen.*

6.1.4 Substitutie-effecten

De regressieresultaten in sectie 6.1.1 laten een positief effect van de overgang op achteraf toetsing op het aantal ingediende HiT aanvragen zien. Mogelijkerwijs is hier sprake van substitutie: omdat het toetsingsregime voor HiT aanvragen minder strikt is geworden, kan het zijn dat zorgaanbieders, waar mogelijk, kiezen om een HiT aanvraag in te dienen. Om dit te testen hebben we het model geschat met het aantal ingediende spoedaanvragen, indicatiemeldingen en reguliere AWBZ aanvragen als uitkomstmaten. We hebben

²⁶Ook voor deze uitkomstvariabele hebben we een model geschat waarin de experiment periode wordt uitgesplitst in 2012 en 2013. Deze uitsplitsing leidt niet tot kwalitatief afwijkende resultaten: ook in dat geval is geen van de effecten significant en de omvang van de effecten is in de meeste gevallen klein.

hierbij de experiment periode gesplitst in 2012 en 2013. We kunnen dan bijvoorbeeld zien of voor de achteraf toetsing groep gedurende het experiment het aantal spoedaanvragen dat wordt ingediend significant afwijkt van het aantal spoedaanvragen dat door zorgaanbieders in de controlegroep wordt ingediend. De resultaten zijn weergegeven in Tabel 6.6.

Tabel 6.6: Schattingsresultaten voor het effect van variatie in de toetsingspraktijk op het aantal indicatiemeldingen, spoedaanvragen en reguliere AWBZ aanvragen.

Panel A: Variatie in het toetsingsmoment.								
	# SIP		# IM		# spoed		# regulier	
HiT achteraf, 2012	-0,112	(0,437)	0,109	(0,094)	-0,430	(0,299)	-0,025	(0,220)
HiT achteraf, 2013	-0,341	(0,521)	0,298	(0,188)	-0,121	(0,370)	-0,693**	(0,287)
HiT prestatie, 2012	0,036	(0,514)	0,109	(0,124)	0,549	(0,345)	0,171	(0,198)
HiT prestatie, 2013	-0,430	(0,664)	0,175	(0,201)	0,597	(0,569)	-0,248	(0,248)
gem. controlegr.	6,253	(0,334)	1,633	(0,074)	4,181	(0,181)	4,017	(0,097)
R^2	0,948		0,849		0,911		0,812	
aantal obs.	9.286		9.286		9.286		9.286	
aantal ZABs	299		299		299		299	

Panel B: Variatie in het toetsingspercentage.								
	# HIT		# IM		# spoed		# regulier	
SIP laag, 2012	-0,141	(0,394)	0,070	(0,166)	-1,082*	(0,648)	-0,177	(0,353)
SIP laag, 2013	0,114	(0,755)	-0,543*	(0,302)	-1,958*	(0,999)	-0,479	(0,458)
SIP hoog, 2012	0,262	(0,339)	-0,071	(0,192)	-1,270*	(0,709)	-0,218	(0,397)
SIP hoog, 2013	0,705	(0,759)	-0,504	(0,362)	-2,104*	(1,071)	-0,561	(0,525)
SIP prestatie, 2012	-0,089	(0,257)	-0,029	(0,156)	-0,443	(0,695)	-0,281	(0,363)
SIP prestatie, 2013	-0,219	(0,491)	-0,435	(0,329)	-0,675	(1,153)	-0,303	(0,440)
gem. controlegr.	2,376	(0,136)	1,698	(0,102)	6,628	(0,334)	5,049	(0,192)
R^2	0,831		0,850		0,905		0,813	
aantal obs.	7.067		7.067		7.067		7.067	
aantal ZABs	226		226		226		226	

Cluster-robuste standaardfouten tussen haakjes. Het gerapporteerde gemiddelde in de controlegroep is het gemiddelde per week, gemeten over de periode voor de start van het experiment. * significant op significantieniveau van 10%, ** significant op significantieniveau van 5%, *** significant op significantieniveau van 1%.

Panel A van Tabel 6.6 laat zien dat we geen significant effect vinden van de overgang op achteraf toetsing op het aantal ingediende SIP aanvragen, het aantal ingediende indicatiemeldingen en het aantal ingediende spoedaanvragen. Voor de spoedaanvragen is de richting van het effect voor de achteraf toetsing groep wel zoals bij substitutie te verwachten zou zijn: het aantal spoedaanvragen neemt af. Anderzijds zien we een toename van het aantal spoedaanvragen in de prestatie groep, waar we een afname in het aantal HiT aanvragen vonden. Ook hier komt de richting van het effect dus overeen met hetgeen we zouden verwachten in geval van substitutie, maar opnieuw is dit effect insignificant. We vinden een significante afname in het aantal reguliere AWBZ aanvragen. Maar met het weergegeven aantal schattingen valt statistisch gezien te verwachten dat een significant effect wordt gevonden. De omvang van de coëfficiënt kan als volgt worden geïnterpreteerd: gemiddeld worden er per week 0,7 reguliere AWBZ aanvragen minder ingediend in de experiment periode in 2013 in de achteraf toetsing groep dan in de controlegroep. Op een gemiddelde in de controlegroep van 4,017 voor aanvang van het experiment is dit een afname van 17%. De gevonden toename in het aantal ingediende HiT aanvragen was gemiddeld 1,06 aanvragen per week in 2013. Wanneer we kijken naar het aantal ingediende HiTs, indicatiemeldingen, spoedaanvragen en reguliere aanvragen door zorgaanbieders in de diverse SIP groepen zien we nauwelijks effecten. Wel neemt

in de groepen met een exogeen laag en een exogeen hoog toetsingspercentage het aantal ingediende spoedaanvragen af.

Conclusie (11): *De resultaten laten enige substitutie zien van HiT aanvragen met spoed-aanvragen en reguliere aanvragen in de achteraf toetsing groep. Voor zorgaanbieders met een exogeen laag en exogeen hoog toetsingspercentage vinden we afnames in het aantal ingediende spoedaanvragen, maar omdat we ook een afname in het aantal SIP aanvragen zagen is dit niet de richting die we zouden verwachten in geval van substitutie-effecten.*

6.2 Heterogene effecten

In sectie 6.1 is gekeken naar gemiddelden over alle zorgaanbieders. Er kan echter sprake zijn van heterogene effecten, dat wil zeggen dat verschillende zorgaanbieders verschillend reageren op veranderingen in de toetsingspraktijk. Dit onderzoeken we in deze sectie.

6.2.1 Grootte van de zorgaanbieder in termen van aantal aanvragen

Allereerst maken we onderscheid tussen het effect voor kleine en grote zorgaanbieders. Daarnaast maken we een meer gedetailleerde uitsplitsing in kleine, middelgrote en grote zorgaanbieders. De definities die voor deze twee analyses worden bekeken, zijn respectievelijk:

1. Uitsplitsing in kleine en grote zorgaanbieders. Deze definitie splitst de zorgaanbieders in twee groepen van ongeveer gelijke omvang. De grens ligt hierbij op gemiddeld per periode van twee weken minder dan 5 (22) ingediende HiT (SIP) aanvragen.²⁷
2. Uitsplitsing in kleine, middelgrote en grote zorgaanbieders. Een zorgaanbieder in een HiT (SIP) groep is geclassificeerd als klein indien deze gemiddeld per twee weken, gemeten over de periode voorafgaand aan het experiment, minder dan 2 (6) HiT (SIP) aanvragen heeft ingediend. Middelgrote zorgaanbieders in een HiT (SIP) groep hebben gemiddeld per twee weken minstens 2 (6) maar minder dan 7 (45) HiT (SIP) aanvragen ingediend. Zorgaanbieders met gemiddeld minstens 7 (45) HiT (SIP) aanvragen worden tenslotte gecategoriseerd als grote zorgaanbieder. Beschrijvende statistieken over welke fractie zorgaanbieders aan de hand van deze definitie in iedere klasse valt, zijn te vinden in Tabel 4.2.

Tabellen 6.7, 6.8 en 6.9 laten de schattingsresultaten zien voor de diverse uitkomstvariabelen. De eerste kolom bevat telkens de resultaten uit de voorgaande basisanalyses. De tweede kolom bevat de resultaten bij gebruik van de eerste definitie (uitsplitsing in kleine en grote zorgaanbieders), terwijl de derde kolom de resultaten bevat bij gebruik van de tweede definitie met een uitsplitsing in kleine, middelgrote en grote zorgaanbieders. We

²⁷Op basis van deze definitie is 58% van de zorgaanbieders in één van de HiT groepen en 45% van de zorgaanbieders in één van de SIP groepen geclassificeerd als klein. De fractie kleine zorgaanbieders verschilt enigszins per groep: het gaat om 54% in de HiT controlegroep en 59% in de achteraf en prestatie groep. Voor SIP zorgaanbieders zijn de percentages: 39% in de controlegroep, 42% (50%) in de groep met laag (hoog) toetsingspercentage en 49% in de groep met prestatieafhankelijk toetsingspercentage.

gebruiken indicator variabelen die aangeven of een zorgaanbieder in de categorie klein, middelgroot of groot valt. Deze indicator interacteren we met de tijd *fixed effects* en de indicatoren voor de diverse *treatments*. Dat betekent dat we toestaan dat de algemene tijdstrend voor kleine, middelgrote en grote zorgaanbieders verschilt, ongeacht de variatie in het toetsingsmoment.

De resultaten voor het aantal ingediende aanvragen in Tabel 6.7 laten zien dat de effecten voor de grote zorgaanbieders groter zijn. In absolute termen is dat een logisch gevolg van de gehanteerde classificatie, maar ook relatief gezien zijn de effecten voor grote zorgaanbieders groter: waar het effect van de overgang op achteraf toetsing voor kleine zorgaanbieders slechts een 6,7% toename is, zien we een geschatte 26,9% toename voor grote zorgaanbieders. Voor de achteraf toetsing groep is alleen het effect voor de grote zorgaanbieders significant. Voor de diverse SIP groepen zien we soortgelijke resultaten: in absolute termen zijn de effecten voor de grote zorgaanbieders logischerwijs aanzienlijk groter. De meeste effecten zijn echter onnauwkeurig geschat en daarom zijn de conclusies vergelijkbaar met die voor de basisanalyses. Ook de resultaten bij gebruik van het aantal toetsingen als uitkomstvariabele, zoals weergegeven in Tabel 6.8 zijn vergelijkbaar. Tenslotte geeft Tabel 6.9 de resultaten weer bij gebruik van de fractie conform getoetste aanvragen als uitkomstvariabele. Hieruit komt naar voren dat er eveneens geen sprake lijkt te zijn van heterogene effecten op de fractie conform getoetste aanvragen. We vinden geen significante effecten van de variatie in toetsingspraktijk op de fractie conform getoetste aanvragen.

Conclusie (12): *De effecten op het aantal aanvragen en het aantal toetsingen zijn logischerwijs groter in absolute termen voor de grotere zorgaanbieders. De resultaten duiden niet op de aanwezigheid van heterogene effecten op het aantal aanvragen, het aantal toetsingen en de fractie conform getoetste aanvragen voor kleine, middelgrote en grote zorgaanbieders.*

Tabel 6.7: Heterogene effecten: kleine versus grote zorgaanbieders; aantal aanvragen.

Panel A: Variatie in het toetsingsmoment.						
	Baseline		Definitie (1)		Definitie (2)	
HiT achteraf	0,624	(0,407)				
HiT prestatie	-0,325	(0,210)				
HiT achteraf × klein			0,099	(0,151)	0,037	(0,191)
HiT achteraf × middelgroot					-0,007	(0,168)
HiT achteraf × groot			1,507*	(0,895)	1,920*	(1,146)
HiT prestatie × klein			-0,154	(0,144)	-0,035	(0,110)
HiT prestatie × middelgroot					-0,087	(0,204)
HiT prestatie × groot			-0,427	(0,408)	-0,693	(0,595)
gem. controlegr. (alle)	3,381	(0,092)				
gem. controlegr. (klein, def. (1))			1,473	(0,041)		
gem. controlegr. (groot, def. (1))			5,594	(0,163)		
gem. controlegr. (klein, def. (2))					0,688	(0,049)
gem. controlegr. (middelgr., def. (2))					2,104	(0,046)
gem. controlegr. (groot, def. (2))					7,337	(0,230)
R^2	0,752		0,762		0,768	
aantal obs.	9.286		9.286		9.286	
aantal ZABs	299		299		299	

Panel B: Variatie in het toetsingspercentage.						
	Baseline		Definitie (1)		Definitie (2)	
SIP laag	-1,291	(1,116)				
SIP hoog	-1,480	(1,127)				
SIP prestatie	-0,614	(0,972)				
SIP laag × klein			0,450	(0,318)	-0,014	(0,190)
SIP laag × middelgroot					0,433	(0,967)
SIP laag × groot			-2,481	(1,850)	-5,537*	(3,266)
SIP hoog × klein			0,207	(0,891)	-0,305	(0,191)
SIP hoog × middelgroot					0,957	(1,110)
SIP hoog × groot			-2,764	(1,800)	-5,680*	(3,184)
SIP prestatie × klein			0,270	(0,693)	-0,374	(0,327)
SIP prestatie × middelgroot					0,386	(0,728)
SIP prestatie × groot			-1,179	(1,527)	-2,739	(2,890)
gem. controlegr. (alle)	18,681	(0,622)				
gem. controlegr. (klein, def. (1))			3,222	(0,131)		
gem. controlegr. (groot, def. (1))			28,549	(0,790)		
gem. controlegr. (klein, def. (2))					1,649	(0,089)
gem. controlegr. (middelgr., def. (2))					11,668	(0,299)
gem. controlegr. (groot, def. (2))					42,932	(1,162)
R^2	0,921		0,924		0,925	
aantal obs.	7.067		7.067		7.067	
aantal ZABs	226		226		226	

De definitie van kleine en grote zorgaanbieder verschilt in kolommen 2 en 3. Daarom moet niet de absolute omvang van de coëfficiënten van de interactietermen tussen deze kolommen worden vergeleken.

Cluster-robuste standaardfouten tussen haakjes. Het gerapporteerde gemiddelde in de controlegroep is het gemiddelde per week, gemeten over de periode voor de start van het experiment. * significant op significantieniveau van 10%, ** significant op significantieniveau van 5%, *** significant op significantieniveau van 1%.

Tabel 6.8: Heterogene effecten: kleine versus grote zorgaanbieders; aantal toetsingen.

Panel A: Variatie in het toetsingsmoment.						
	Baseline		Definitie (1)		Definitie (2)	
HiT achteraf	0,221**	(0,106)				
HiT prestatie	-0,016	(0,063)				
HiT achteraf × klein			0,084*	(0,046)	0,068	(0,059)
HiT achteraf × middelgroot					0,036	(0,053)
HiT achteraf × groot			0,455**	(0,231)	0,596**	(0,292)
HiT prestatie × klein			-0,024	(0,040)	-0,010	(0,033)
HiT prestatie × middelgroot					0,032	(0,061)
HiT prestatie × groot			0,032	(0,125)	-0,009	(0,179)
gem. controlegr. (alle)	0,769	(0,025)				
gem. controlegr. (klein, def. (1))			0,343	(0,015)		
gem. controlegr. (groot, def. (1))			1,262	(0,045)		
gem. controlegr. (klein, def. (2))					0,165	(0,021)
gem. controlegr. (middelgr., def. (2))					0,471	(0,017)
gem. controlegr. (groot, def. (2))					1,679	(0,064)
R^2	0,602		0,616		0,623	
aantal obs.	9.286		9.286		9.286	
aantal ZABs	299		299		299	

Panel B: Variatie in het toetsingspercentage.						
	Baseline		Definitie (1)		Definitie (2)	
SIP laag	-0,799***	(0,123)				
SIP hoog	0,861***	(0,165)				
SIP prestatie	0,835***	(0,200)				
SIP laag × klein			-0,218***	(0,049)	-0,102***	(0,032)
SIP laag × middelgroot					-0,401***	(0,079)
SIP laag × groot			-1,223***	(0,178)	-2,068***	(0,250)
SIP hoog × klein			0,253	(0,181)	-0,024	(0,024)
SIP hoog × middelgroot					0,902***	(0,246)
SIP hoog × groot			1,458***	(0,201)	1,393***	(0,292)
SIP prestatie × klein			0,259*	(0,145)	0,005	(0,064)
SIP prestatie × middelgroot					0,417***	(0,158)
SIP prestatie × groot			1,342***	(0,315)	1,699***	(0,434)
gem. controlegr. (alle)	1,394	(0,052)				
gem. controlegr. (klein, def. (1))			0,210	(0,019)		
gem. controlegr. (groot, def. (1))			2,149	(0,069)		
gem. controlegr. (klein, def. (2))					0,096	(0,017)
gem. controlegr. (middelgr., def. (2))					0,895	(0,040)
gem. controlegr. (groot, def. (2))					3,177	(0,106)
R^2	0,672		0,691		0,703	
aantal obs.	7.067		7.067		7.067	
aantal ZABs	226		226		226	

De definitie van kleine en grote zorgaanbieder verschilt in kolommen 2 en 3. Daarom moet niet de absolute omvang van de coëfficiënten van de interactietermen tussen deze kolommen worden vergeleken.

Cluster-robuste standaardfouten tussen haakjes. Het gerapporteerde gemiddelde in de controlegroep is het gemiddelde per week, gemeten over de periode voor de start van het experiment. * significant op significantieniveau van 10%, ** significant op significantieniveau van 5%, *** significant op significantieniveau van 1%.

Tabel 6.9: Heterogene effecten: kleine versus grote zorgaanbieders; conformfractie.

Panel A: Variatie in het toetsingsmoment.						
	Baseline		Definitie (1)		Definitie (2)	
HiT achteraf	0,027	(0,020)				
HiT prestatie	0,002	(0,020)				
HiT achteraf × klein			0,010	(0,036)	0,085	(0,101)
HiT achteraf × middelgroot					0,005	(0,032)
HiT achteraf × groot			0,034	(0,024)	0,042	(0,027)
HiT prestatie × klein			-0,008	(0,037)	-0,040	(0,093)
HiT prestatie × middelgroot					0,015	(0,033)
HiT prestatie × groot			0,007	(0,024)	0,007	(0,026)
gem. controlegr. (alle)	0,823	(0,009)				
gem. controlegr. (klein, def. (1))			0,827	(0,016)		
gem. controlegr. (groot, def. (1))			0,820	(0,011)		
gem. controlegr. (klein, def. (2))					0,889	(0,038)
gem. controlegr. (middelgr., def. (2))					0,813	(0,014)
gem. controlegr. (groot, def. (2))					0,827	(0,013)
R^2	0,105		0,111		0,119	
aantal obs.	5.727		5.727		5.727	
aantal ZABs	299		299		299	
Panel B: Variatie in het toetsingspercentage.						
	Baseline		Definitie (1)		Definitie (2)	
SIP laag	-0,001	(0,022)				
SIP hoog	0,000	(0,021)				
SIP prestatie	-0,000	(0,020)				
SIP laag × klein			-0,087	(0,073)	0,014	(0,156)
SIP laag × middelgroot					-0,021	(0,034)
SIP laag × groot			0,017	(0,022)	0,030	(0,022)
SIP hoog × klein			-0,034	(0,074)	0,076	(0,169)
SIP hoog × middelgroot					-0,014	(0,035)
SIP hoog × groot			0,005	(0,018)	-0,000	(0,018)
SIP prestatie × klein			-0,039	(0,065)	0,350*	(0,209)
SIP prestatie × middelgroot					-0,010	(0,031)
SIP prestatie × groot			0,006	(0,017)	-0,006	(0,016)
gem. controlegr. (alle)	0,868	(0,010)				
gem. controlegr. (klein, def. (1))			0,768	(0,037)		
gem. controlegr. (groot, def. (1))			0,889	(0,009)		
gem. controlegr. (klein, def. (2))					0,682	(0,081)
gem. controlegr. (middelgr., def. (2))					0,870	(0,015)
gem. controlegr. (groot, def. (2))					0,887	(0,011)
R^2	0,150		0,159		0,179	
aantal obs.	4.857		4.857		4.857	
aantal ZABs	224		224		224	

De definitie van kleine en grote zorgaanbieder verschilt in kolommen 2 en 3.

Cluster-robuste standaardfouten tussen haakjes. Het gerapporteerde gemiddelde in de controlegroep is het gemiddelde per week, gemeten over de periode voor de start van het experiment. * significant op significantieniveau van 10%, ** significant op significantieniveau van 5%, *** significant op significantieniveau van 1%.

6.2.2 Aanvraagmix van de zorgaanbieder

De aanvraagmix, de samenstelling (in termen van het soort aanvragen) van de ingediende aanvragen, kan een rol spelen in de mate waarin variatie in toetsingsmoment en toetsingspercentage invloed heeft op het gedrag van zorgaanbieders. Dit kan bijvoorbeeld te maken hebben met de ervaring die een zorgaanbieder heeft met een bepaald soort aanvraag, maar ook met de mate waarin een zorgaanbieder last heeft van de variatie in de toetsingspraktijk. We schatten soortgelijke modellen als in de voorgaande sectie, maar maken nu een onderscheid tussen zorgaanbieders naar aandeel van HiT/SIP aanvragen in het totaal aantal ingediende aanvragen. We gebruiken de volgende definitie om dit onderscheid te maken.

- Zorgaanbieders in één van de HiT groepen worden geclassificeerd als zorgaanbieder met een klein HiT aandeel wanneer de fractie HiT aanvragen in het totaal van ingediende aanvragen in de periode voor de start van het experiment hoogstens 0,25 is. Zorgaanbieders met een grotere fractie HiT aanvragen vallen in de groep met een groot HiT aandeel. Volgens deze definitie heeft 48% van de HiT zorgaanbieders een klein aandeel HiT aanvragen.
- Zorgaanbieders in één van de SIP groepen worden geclassificeerd als zorgaanbieder met een klein SIP aandeel wanneer de fractie SIP aanvragen in het totaal van ingediende aanvragen in de periode voor aanvang van het experiment hoogstens 0,55 is. Zorgaanbieders met een grotere fractie SIP aanvragen vallen in de groep met een groot SIP aandeel. Volgens deze definitie heeft 50% van de SIP zorgaanbieders een klein aandeel SIP aanvragen.

Tabel 6.10 geeft de resultaten weer voor de diverse uitkomstvariabelen. De resultaten van de basisanalyses zijn niet in deze tabel weergegeven, maar kunnen in sectie 6.1 worden gevonden. Allereerst blijkt uit de tabel dat er een significant positief effect is op het aantal HiT aanvragen voor zorgaanbieders met een klein aandeel HiT aanvragen in de HiT achteraf toetsing groep. Ten opzichte van het gemiddelde in de controlegroep van zorgaanbieders met een klein aandeel HiT aanvragen gaat het om een stijging van ruim 40%. Daarentegen vinden we een insignificant (negatief) effect voor zorgaanbieders met een groot aandeel HiT aanvragen. Het effect op zorgaanbieders in de achteraf toetsing groep met een klein aandeel HiT aanvragen en een groot aandeel HiT aanvragen is echter niet significant verschillend. Voor de prestatie groep vinden we daarentegen een significant negatief effect op het aantal aanvragen ingediend door zorgaanbieders met een groot aandeel HiT aanvragen. Het betreft hier een daling van bijna 22% ten opzichte van zorgaanbieders met een groot aandeel HiT aanvragen in de controlegroep. Indien we tenslotte kijken naar de effecten op het aantal toetsingen in de SIP prestatie groep is het interessant om op te merken dat het grootste effect (zowel in absolute termen als in procentuele termen) wordt geschat voor de zorgaanbieders met een klein aandeel SIP aanvragen. Dit geeft aan dat het toetsingspercentage voor zorgaanbieders met een klein aandeel SIP aanvragen sterker stijgt dan voor zorgaanbieders met een groot aandeel SIP aanvragen. We hebben getoetst of er een verschil is in het effect op de fractie conform getoetste aanvragen, maar vinden geen significante heterogeniteit in effecten.

Conclusie (13): *We vinden een verschil in het effect van de overgang op achteraf toetsing op het aantal ingediende aanvragen door zorgaanbieders met een klein aandeel HiT aanvragen en zorgaanbieders met een groot aandeel HiT aanvragen. Voor de eerste groep vinden we een significant positief effect op het aantal ingediende aanvragen. Voor de andere uitkomstvariabelen vinden we geen verschillen in effecten.*

Tabel 6.10: Heterogene effecten: aanvraagmix.

Panel A: Variatie in het toetsingsmoment.						
	# aanvragen		# toetsingen		fractie conform	
HiT achteraf × klein HiT aandeel	1,550**	(0,727)	0,410**	(0,191)	0,067**	(0,028)
HiT achteraf × groot HiT aandeel	-0,244	(0,396)	0,047	(0,106)	-0,009	(0,027)
HiT prestatie × klein HiT aandeel	-0,035	(0,275)	0,068	(0,082)	0,007	(0,030)
HiT prestatie × groot HiT aandeel	-0,667**	(0,315)	-0,111	(0,097)	-0,002	(0,027)
gem. controlegr. (klein HiT aandeel)	3,631	(0,139)	0,829	(0,037)	0,809	(0,013)
gem. controlegr. (groot HiT aandeel)	3,040	(0,107)	0,686	(0,030)	0,841	(0,014)
R^2	0,755		0,605		0,112	
aantal obs.	5.727		9.286		5.727	
aantal ZABs	299		299		299	
Panel B: Variatie in het toetsingspercentage.						
	# aanvragen		# toetsingen		fractie conform	
SIP laag × klein SIP aandeel	-1,857	(1,686)	-0,408***	(0,119)	0,022	(0,049)
SIP laag × groot SIP aandeel	-0,684	(1,354)	-1,137***	(0,186)	-0,020	(0,021)
SIP hoog × klein SIP aandeel	-3,259	(1,989)	0,487**	(0,200)	0,032	(0,041)
SIP hoog × groot SIP aandeel	0,105	(1,036)	1,161***	(0,241)	-0,021	(0,021)
SIP prestatie × klein SIP aandeel	-0,881	(1,734)	1,004***	(0,284)	0,033	(0,040)
SIP prestatie × groot SIP aandeel	-0,377	(0,707)	0,624**	(0,272)	-0,022	(0,016)
gem. controlegr. (klein SIP aandeel)	14,020	(0,763)	0,824	(0,053)	0,842	(0,019)
gem. controlegr. (groot SIP aandeel)	23,990	(0,952)	2,043	(0,085)	0,887	(0,011)
R^2	0,923		0,679		0,156	
aantal obs.	7.067		7.067		4.857	
aantal ZABs	226		226		224	

Cluster-robuste standaardfouten tussen haakjes. Het gerapporteerde gemiddelde in de controlegroep is het gemiddelde per week, gemeten over de periode voor de start van het experiment. * significant op significantieniveau van 10%, ** significant op significantieniveau van 5%, *** significant op significantieniveau van 1%.

6.2.3 Ziekenhuizen ten opzichte van ander soort zorgaanbieders

Er kan eveneens worden onderzocht of er verschillen bestaan in de manier waarop ziekenhuizen reageren op variatie in het toetsingspercentage en de manier waarop ander soort zorgaanbieders (bijvoorbeeld thuiszorginstellingen) op soortgelijke variatie reageren.²⁸ Ziekenhuizen dienen vaak SIP aanvragen in voordat de patiënt uit het ziekenhuis wordt ontslagen, om te zorgen dat er voldoende zorg inzetbaar is wanneer de patiënt het ziekenhuis eenmaal heeft verlaten. Daarom kan het zijn dat de invloed van het toetsingspercentage op het gedrag van ziekenhuizen kleiner is dan voor andere zorgaanbieders. We schatten soortgelijke modellen als in de voorgaande secties, maar we maken nu een onderscheid tussen ziekenhuizen en andere zorgaanbieders. De resultaten zijn opgenomen in Tabel 6.11. Panel A van deze tabel geeft de schattingen weer voor het effect op het aantal ingediende aanvragen. We zien hier dat het effect voor ziekenhuizen in iedere SIP groep groter is in absolute waarde dan het effect voor andere zorgaanbieders. De meeste effecten zijn insignificant, maar we vinden wel een groot significant negatief effect op het aantal aanvragen ingediend door niet-ziekenhuizen in de hoog toetsingspercentage groep. Dit effect heeft de richting die we zouden verwachten. Panel B laat vervolgens de effecten zien op het aantal uitgevoerde toetsingen. Interessant is het verschil tussen ziekenhuizen en ander soort zorgaanbieders in de prestatiegroep. Hier zien we een relatief klein positief effect voor het aantal toetsingen voor ziekenhuizen (ten opzichte van ziekenhuizen in de controlegroep gaat het hier om een toename van 29%). Dit is het gevolg van de eerder gerapporteerde stijging in het toetsingspercentage in de SIP groep met het prestatieafhankelijke toetsingspercentage. Het effect voor niet-ziekenhuizen is echter veel groter, zowel in absolute omvang als procentueel: ten opzichte van ander soort zorgaanbieders in de controlegroep zien we een geschatte toename van 91%. We hebben getest of het verschil in effect significant is, maar kunnen de nulhypothese van gelijke effecten (in absolute omvang) niet verwerpen. Als we tenslotte kijken naar de geschatte effecten op de fractie conform getoetste aanvragen, zoals gepresenteerd in panel C, zien we dat in de meeste groepen geen sprake is van significante effecten. We zien telkens dat het geschatte effecten voor ziekenhuizen negatief is en voor ander soort zorgaanbieders positief, maar de effecten zijn onnauwkeurig geschat en we kunnen de nulhypothese van gelijke effecten voor ziekenhuizen en ander soort zorgaanbieders (in iedere SIP groep) niet verwerpen.

Conclusie (14): *We vinden verschillen in het effect op het aantal ingediende aanvragen door ziekenhuizen en andere zorgaanbieders in de groep met het hoge toetsingspercentage. Het aantal ingediende aanvragen neemt toe voor ziekenhuizen, maar daalt voor ander soort zorgaanbieders. Er lijkt echter geen sprake te zijn van verschillen tussen ziekenhuizen en andere zorgaanbieders in de groep met het prestatieafhankelijke toetsingspercentage. Bovendien vinden we geen verschillen in het effect op de fractie conform getoetste aanvragen.*

²⁸Dit is alleen relevant voor het gedeelte van het experiment waarin het toetsingspercentage wordt gevarieerd, omdat alleen daar ziekenhuizen deelnamen aan het experiment.

Tabel 6.11: Heterogene effecten: ziekenhuis versus ander soort zorgaanbieders.

Panel A: Aantal SIP aanvragen				
	Baseline		Ziekenhuis / ander soort ZABs	
SIP laag	-1,577	(1,040)		
SIP hoog	-1,464	(1,136)		
SIP prestatie	-0,571	(0,976)		
SIP laag × ziekenhuis			-0,617	(0,756)
SIP laag × ander soort ZAB			-2,088	(1,628)
SIP hoog × ziekenhuis			1,423	(1,208)
SIP hoog × ander soort ZAB			-3,237**	(1,582)
SIP prestatie × ziekenhuis			0,086	(0,723)
SIP prestatie × ander soort ZAB			-0,851	(1,457)
gem. controlegr. (alle)	18,814	(0,632)		
gem. controlegr. (ziekenhuis)			19,682	(0,623)
gem. controlegr. (ander soort ZAB)			18,388	(0,891)
R^2	0,924		0,925	
aantal obs.	6.826		6.826	
aantal ZABs	216		216	
Panel B: Aantal SIP toetsingen				
SIP laag	-0,795***	(0,128)		
SIP hoog	0,879***	(0,167)		
SIP prestatie	0,840***	(0,204)		
SIP laag × ziekenhuis			-0,930***	(0,203)
SIP laag × ander soort ZAB			-0,684***	(0,174)
SIP hoog × ziekenhuis			1,511***	(0,258)
SIP hoog × ander soort ZAB			0,454**	(0,186)
SIP prestatie × ziekenhuis			0,589*	(0,304)
SIP prestatie × ander soort ZAB			1,015***	(0,274)
gem. controlegr. (alle)	1,407	(0,053)		
gem. controlegr. (ziekenhuis)			2,005	(0,089)
gem. controlegr. (ander soort ZAB)			1,114	(0,063)
R^2	0,671		0,679	
aantal obs.	6.826		6.826	
aantal ZABs	216		216	
Panel C: Fractie conform getoetste SIP aanvragen				
SIP laag	-0,009	(0,022)		
SIP hoog	-0,003	(0,021)		
SIP prestatie	-0,004	(0,019)		
SIP laag × ziekenhuis			-0,045*	(0,024)
SIP laag × ander soort ZAB			0,021	(0,039)
SIP hoog × ziekenhuis			-0,018	(0,025)
SIP hoog × ander soort ZAB			0,005	(0,031)
SIP prestatie × ziekenhuis			-0,036*	(0,019)
SIP prestatie × ander soort ZAB			0,021	(0,031)
gem. controlegr. (alle)	0,866	(0,010)		
gem. controlegr. (ziekenhuis)			0,892	(0,013)
gem. controlegr. (ander soort ZAB)			0,847	(0,015)
R^2	0,150		0,156	
aantal obs.	4.704		4.704	
aantal ZABs	214		214	

Cluster-robuste standaardfouten tussen haakjes. Het gerapporteerde gemiddelde in de controle-groep is het gemiddelde per week, gemeten over de periode voor de start van het experiment. * significant op significantieniveau van 10%, ** significant op significantieniveau van 5%, *** significant op significantieniveau van 1%.

6.3 Robuustheidsanalyses

In sectie 6.1 zijn resultaten van de basisanalyses getoond. Hierin zijn zorgaanbieders die zijn gedefinieerd als uitgevallen nog wel meegenomen (in die periodes waarin ze nog niet inactief waren). Wanneer de regressies worden gedaan op data waaruit de inactief geworden zorgaanbieders volledig zijn verwijderd, zijn de resultaten zoals weergegeven in Tabel 6.12. Wanneer we deze resultaten vergelijken met de resultaten voor het *unbalanced panel* in Tabellen 6.1, 6.3 en 6.5, kunnen we concluderen dat de resultaten robuust zijn. De omvang van de effecten is vergelijkbaar, de gerapporteerde gemiddelde aantallen aanvragen en toetsingen in de controlegroep voor aanvang van het experiment zijn logischerwijs wat hoger wanneer de inactief geworden zorgaanbieders eerst worden verwijderd. De resultaten lijken echter niet erg gevoelig te zijn voor de manier waarop rekening wordt gehouden met de inactief geworden zorgaanbieders.

Tabel 6.12: Robuustheidsanalyses: verwijderen van inactieve zorgaanbieders.

Panel A: Variatie in het toetsingsmoment.										
	# HiT aanvragen		# HiT Type 1		# ander soort HiT		# HiT toetsingen		fractie conform HiT	
HiT achteraf	0,639	(0,415)	0,717*	(0,373)	-0,077	(0,189)	0,224**	(0,108)	0,024	(0,021)
HiT prestatie	-0,339	(0,213)	-0,018	(0,142)	-0,321**	(0,161)	-0,023	(0,064)	0,002	(0,020)
gem. controlegr.	3,525	(0,100)	2,541	(0,094)	0,984	(0,040)	0,798	(0,027)	0,824	(0,010)
R^2	0,754		0,804		0,608		0,605		0,103	
aantal obs.	8.952		8.952		8.952		8.952		5.545	
aantal ZABs	281		281		281		281		281	

Panel B: Variatie in het toetsingspercentage.						
	# SIP aanvragen		# SIP toetsingen		fractie conform SIP	
SIP laag	-1,301	(1,128)	-0,799***	(0,124)	0,003	(0,022)
SIP hoog	-0,990	(1,045)	0,894***	(0,168)	0,006	(0,021)
SIP prestatie	-0,578	(0,988)	0,851***	(0,203)	0,003	(0,019)
gem. controlegr.	19,286	(0,646)	1,440	(0,054)	0,869	(0,010)
R^2	0,927		0,667		0,151	
aantal obs.	6.830		6.830		4.721	
aantal ZABs	214		214		212	

Cluster-robuste standaardfouten tussen haakjes. Het gerapporteerde gemiddelde in de controlegroep is het gemiddelde per week, gemeten over de periode voor de start van het experiment. * significant op significantieniveau van 10%, ** significant op significantieniveau van 5%, *** significant op significantieniveau van 1%.

Voor iedere 2-weeks periode is een dummy opgenomen (*time fixed effects*). De geschatte coëfficiënt voor deze tijdsdummies zijn niet gerapporteerd.

In alle voorgaande analyses hebben we gecorrigeerd voor algemene periode effecten (tijd *fixed effects*) door voor ieder periode van twee weken een dummy variabele op te nemen.²⁹ Andere mogelijkheden voor het corrigeren voor tijdstrends zijn het gebruik van polynomen in tijd of het opnemen van kwartaaldummies. De resultaten van deze alternatieve methoden zijn opgenomen in Tabellen 6.13 en 6.14. Als we deze resultaten vergelijken met die in sectie 6.1 valt op dat, wanneer we een 2^e -graads polynoom gebruiken om voor tijdstrends te corrigeren, er een significant negatief effect wordt gevonden op het aantal HiT aanvragen dat wordt ingediend door zorgaanbieders in de groep waarvoor het gehanteerde toetsingsmoment afhangt van prestatie. In vergelijking met de resultaten in Tabel 6.1 is het effect een fractie groter en de standaardfout kleiner. Dit is het gevolg van een onderschatting van het aantal ingediende HiT aanvragen in 2013. Een betere overeenkomst met de geobserveerde data wordt verkregen wanneer we, naast de polynoom in tijd of naast de kwartaaldummies, een uitsplitsing maken in de periodes 2012 en 2013, zoals eerder is gedaan in sectie 6.1. De resultaten van deze regressies zijn te vinden in Appendix G. Op basis van deze laatste resultaten komen we niet tot kwalitatief andere conclusies. We zien dat, bij gebruik van een 2^e -graads polynoom in tijd, het positieve effect op de fractie conform getoetste aanvragen in de achteraf toetsing groep net significant wordt (bij gebruik van een 10% significantieniveau). Bovendien vinden we, zowel bij gebruik van kwartaaldummies als een 2^e -graads polynoom, een significant effect op het aantal toetsingen in 2012 in de achteraf toetsing groep. Dat terwijl we in de basisanalyses alleen een significant effect vonden voor 2013.

Een laatste robuustheidsanalyse kijkt naar de invloed van het niveau van aggregatie. In de eerdere analyses hebben we de data geaggregeerd naar periodes van twee weken. Dit beperkt de ruis ten opzichte van aggregatie per week. Als robuustheidsanalyse hebben we de data geaggregeerd naar periodes van vier weken (en vervolgens genormaliseerd om vergelijkbare effecten op gemiddelde aantallen per week te kunnen rapporteren) en gecorrigeerd voor algemene tijdstrends door dummies per periode van vier weken op te nemen. De resultaten van deze analyses zijn te vinden in Tabel 6.15 en komen met overeen met de eerder gepresenteerde resultaten.

Conclusie (15): *De robuustheidsanalyses laten zien dat de resultaten robuust zijn voor: (1) het volledig buiten beschouwing laten van uitgevallen zorgaanbieders, (2) het op alternatieve wijze corrigeren voor de algemene trend over tijd in de uitkomstvariabelen, en (3) het reduceren van de ruis in de periodieke cijfers door de data te aggregeren op langere periodes.*

²⁹Een dummy variabele voor periode t is gelijk aan 1 als het een observatie voor zorgaanbieder i in periode t betreft en is in andere periodes gelijk aan 0.

Tabel 6.13: Robuustheidsanalyses: polynoom in tijd.

Panel A: Variatie in het toetsingsmoment.						
	# HiT aanvragen		# HiT toetsingen		fractie conform HiT	
HiT achteraf	0,546	(0,334)	0,242***	(0,089)	0,012	(0,019)
HiT prestatie	-0,404**	(0,169)	0,004	(0,054)	-0,012	(0,018)
gem. controlegr.	3,381	(0,092)	0,769	(0,025)	0,823	(0,009)
R^2	0,750		0,597		0,092	
aantal obs.	9.286		9.286		5.727	
aantal ZABs	299		299		299	
Panel B: Variatie in het toetsingspercentage.						
	# SIP aanvragen		# SIP toetsingen		fractie conform SIP	
SIP laag	-0,653	(0,796)	-0,948***	(0,120)	-0,007	(0,019)
SIP hoog	-0,838	(0,855)	0,711***	(0,158)	-0,005	(0,017)
SIP prestatie	0,028	(0,660)	0,685***	(0,165)	-0,005	(0,015)
gem. controlegr.	18,681	(0,622)	1,394	(0,052)	0,868	(0,010)
R^2	0,920		0,668		0,142	
aantal obs.	7.067		7.067		4.857	
aantal ZABs	226		226		224	

Cluster-robuste standaardfouten tussen haakjes. Het gerapporteerde gemiddelde in de controlegroep is het gemiddelde per week, gemeten over de periode voor de start van het experiment. * significant op significantieniveau van 10%, ** significant op significantieniveau van 5%, *** significant op significantieniveau van 1%.

Om te corrigeren voor algemene tijdstrends is een 2^e-graads polynoom in tijd en een dummy voor de kerstperiode opgenomen. De bijbehorende coëfficiënten zijn niet in de tabel weergegeven.

Tabel 6.14: Robuustheidsanalyses: kwartaaldummies.

Panel A: Variatie in het toetsingsmoment.						
	# HiT aanvragen		# HiT toetsingen		fractie conform HiT	
HiT achteraf	0,578	(0,362)	0,228**	(0,095)	0,024	(0,019)
HiT prestatie	-0,371**	(0,177)	-0,009	(0,056)	-0,001	(0,019)
gem. controlegr.	3,381	(0,092)	0,769	(0,025)	0,823	(0,009)
R^2	0,751		0,598		0,093	
aantal obs.	9.286		9.286		5.727	
aantal ZABs	299		299		299	
Panel B: Variatie in het toetsingspercentage.						
	# SIP aanvragen		# SIP toetsingen		fractie conform SIP	
SIP laag	-0,854	(0,921)	-0,858***	(0,118)	0,005	(0,020)
SIP hoog	-1,050	(0,944)	0,801***	(0,157)	0,007	(0,018)
SIP prestatie	-0,174	(0,781)	0,775***	(0,183)	0,007	(0,017)
gem. controlegr.	18,681	(0,622)	1,394	(0,052)	0,868	(0,010)
R^2	0,920		0,669		0,143	
aantal obs.	7.067		7.067		4.857	
aantal ZABs	226		226		224	

Cluster-robuste standaardfouten tussen haakjes. Het gerapporteerde gemiddelde in de controlegroep is het gemiddelde per week, gemeten over de periode voor de start van het experiment. * significant op significantieniveau van 10%, ** significant op significantieniveau van 5%, *** significant op significantieniveau van 1%.

Om te corrigeren voor algemene tijdstrends zijn kwartaaldummies opgenomen. Bovendien is een aparte dummy voor de kerstperiode opgenomen. De geschatte coëfficiënten voor deze dummy variabelen zijn niet gerapporteerd in de tabel.

Tabel 6.15: Robuustheidsanalyses: aggregatie per 4 weken.

Panel A: Variatie in het toetsingsmoment.											
	# HiT aanvragen				# HiT toetsingen				fractie conform		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	
HiT achteraf	0,611	(0,407)			0,218**	(0,107)			0,022	(0,021)	
HiT prestatie	-0,320	(0,210)			-0,015	(0,063)			-0,006	(0,022)	
HiT achteraf, 2012			0,290	(0,289)			0,108	(0,080)			0,024 (0,027)
HiT achteraf, 2013			1,052*	(0,620)			0,368**	(0,160)			0,020 (0,024)
HiT prestatie, 2012			-0,178	(0,159)			-0,011	(0,056)			0,000 (0,028)
HiT prestatie, 2013			-0,507	(0,346)			-0,018	(0,095)			-0,013 (0,026)
gem. controlegr.	3,719	(0,108)	3,719	(0,108)	0,853	(0,026)	0,853	(0,026)	0,825	(0,007)	0,825 (0,007)
R^2	0,814		0,815		0,711		0,712		0,138		0,138
aantal obs.	4.650		4.650		4.650		4.650		3.636		3.636
aantal ZABs	299		299		299		299		299		299

Panel B: Variatie in het toetsingspercentage.											
	# SIP aanvragen				# SIP toetsingen				fractie conform		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	
SIP laag	-1,291	(1,116)			-0,799***	(0,123)			-0,004	(0,029)	
SIP hoog	-1,593	(1,174)			0,850***	(0,166)			0,007	(0,027)	
SIP prestatie	-0,590	(0,974)			0,837***	(0,200)			0,010	(0,025)	
SIP laag, 2012			-0,970	(1,030)			-0,695***	(0,126)			-0,005 (0,036)
SIP laag, 2013			-1,720	(1,313)			-0,937***	(0,149)			-0,002 (0,038)
SIP hoog, 2012			-1,395	(1,103)			0,853***	(0,154)			0,008 (0,030)
SIP hoog, 2013			-1,861	(1,365)			0,842***	(0,214)			0,007 (0,036)
SIP prestatie, 2012			-0,613	(0,895)			0,414***	(0,124)			-0,009 (0,029)
SIP prestatie, 2013			-0,558	(1,190)			1,408***	(0,351)			0,034 (0,033)
gem. controlegr.	19,404	(0,699)	19,404	(0,699)	1,449	(0,053)	1,449	(0,053)	0,852	(0,009)	0,852 (0,009)
R^2	0,934		0,934		0,757		0,764		0,231		0,231
aantal obs.	3.538		3.538		3.538		3.538		2.792		2.792
aantal ZABs	226		226		226		226		224		224

Cluster-robuste standaardfouten tussen haakjes. Het gerapporteerde gemiddelde in de controlegroep is het gemiddelde per week, gemeten over de periode voor de start van het experiment.
 * significant op significantieniveau van 10%, ** significant op significantieniveau van 5%, *** significant op significantieniveau van 1%.

7 Conclusies

In deze eindrapportage van het Proeftuin onderzoek ‘Meten is Weten’ zijn de effecten van variatie in de gehanteerde toetsingspraktijk op ontwikkelingen in aantallen aanvragen, aantallen toetsingen en fractie conform getoetste aanvragen bekeken. In het uitgevoerde experiment hebben geselecteerde zorgaanbieders gedurende een half jaar te maken gehad met variatie in het toetsingsmoment en/of het toetsingspercentage. Hierbij is zowel gekeken naar exogene als endogene (prestatieafhankelijke) variatie in één van deze beleidsinstrumenten. In deze sectie worden de belangrijkste resultaten van dit onderzoek besproken.

HiT aanvragen worden normaliter vooraf getoetst. In het veldexperiment hebben we gekeken naar de effecten van achteraf toetsing, waarbij zorgaanbieders niet hoeven te wachten met het inzetten van zorg, en het gebruik van een prestatieafhankelijk moment van toetsing. De belangrijkste conclusies zijn:

1. Er zijn geen significante verschillen in (gemiddelde) kwantitatieve én kwalitatieve prestaties van zorgaanbieders in de vooraf toetsing groep en zorgaanbieders in de groep met een prestatieafhankelijk toetsingsmoment.
2. Het aantal zorgaanbieders in de prestatie groep dat vooraf werd getoetst, is gedurende het experiment steeds verder afgenomen. Door gebruik te maken van een prestatieafhankelijk toetsingsmoment kunnen die zorgaanbieders die (consequent) goed presteren worden geselecteerd voor achteraf toetsing. Zij kunnen de aangevraagde zorg dan snel inzetten. Ten opzichte van vooraf toetsing gaat het gebruik van een prestatieafhankelijk toetsingsmoment dus gepaard met een selectie-effect, maar worden geen effecten op aanvraaggedrag van zorgaanbieders gevonden.
3. Daarentegen blijkt uit de resultaten een toename in het aantal aanvragen ingediend door zorgaanbieders in de achteraf toetsing groep. Dit kan niet (geheel) worden verklaard door substitutie in het soort aanvragen en heeft daarmee directe invloed op de grote geldstroom (AWBZ uitgaven). Dit effect op aantal aanvragen ontbreekt in de prestatie groep en duidt daarmee op een opvoedende werking van de mogelijkheid tot overgang op vooraf toetsing bij ondermaatse prestaties.
4. Deze effecten zijn niet verschillend voor kleine en grote zorgaanbieders. We vinden echter wel verschillen in effecten voor zorgaanbieders met een klein aandeel HiT aanvragen in het totale aantal aanvragen en zorgaanbieders met een groot aandeel HiT aanvragen in de achteraf toetsing groep. Zorgaanbieders met een initieel klein aandeel HiT aanvragen zijn voornamelijk verantwoordelijk voor de stijging in het aantal HiT aanvragen.

SIP aanvragen worden achteraf getoetst, er is dus geen mogelijkheid voor correctie in geval van een niet conform toetsing. In het veldexperiment hebben we onderzocht wat het effect is van het verlagen van het toetsingspercentage, het verhogen van het toetsingspercentage,

of het aanpassen van het toetsingspercentage op basis van de resultaten van recente toetsingen. De belangrijkste conclusies van dit deel van het veldexperiment zijn:

1. Geen enkele vorm van variatie in het toetsingspercentage (laag, hoog of prestatieafhankelijk) heeft een statistisch significant effect op kwantiteit (aantal aanvragen) noch op kwaliteit (conformpercentage).
2. De resultaten hebben daarentegen wel de (mechanische) effecten op het aantal toetsingen geïllustreerd. Vooral het effect op het aantal toetsingen in de prestatie groep is daarbij interessant: voor deze groep is een toename in het aantal toetsingen gevonden die, gemiddeld over de experiment periode, ongeveer even groot is als de gemiddelde toename in de hoog toetsingspercentage groep. Dit is het gevolg geweest van een geleidelijke toename van het toetsingspercentage gedurende het experiment, wegens prestaties die gemiddeld beneden de vastgestelde norm lagen.
3. We vinden verschillen in effecten in de hoog toetsingspercentage groep wanneer we onderscheid maken tussen ziekenhuizen en andere zorgaanbieders: voor ziekenhuizen vinden we een groter positief effect op aantal aanvragen en aantal toetsingen.

Als we de resultaten van de twee onderdelen van het experiment in combinatie bekijken, volgen daaruit enkele algemene lessen. Het variëren van het toetsingspercentage in een regime van achteraf toetsing, dus waarbij geen correctiemogelijkheden aanwezig zijn, heeft nauwelijks invloed op het gedrag van zorgaanbieders. De dimensie vooraf/achteraf lijkt belangrijk te zijn. Alleen achteraf toetsen is geen restricterende maatregel, maar van de combinatie met vooraf toetsing bij ondermaatse prestaties gaat wel een beperkende werking uit.

A Simulatiestudie looptijd experiment

Het belang van een simulatiestudie is om te bepalen hoeveel data nodig is om statistisch een betrouwbaar antwoord te geven. Bij veel empirisch onderzoek wordt ernaar gestreefd om een power van minimaal 0,6 a 0,7 te krijgen. Om vast te stellen hoe lang het experiment dient te lopen om statistisch gegronde conclusies te kunnen trekken, hebben we voor aanvang van het experiment een simulatiestudie uitgevoerd, een zogenaamde power-analyse. In zo'n simulatiestudie wordt gekeken hoe groot de kans is dat bij een bepaalde steekproefomvang een gegeven effect van de wijziging in toetsingspraktijk gevonden wordt. Op basis van recente gegevens over het aantal HiT en SIP aanvragen in de periode voorafgaand aan het experiment hebben we deze steekproefomvang vertaald naar looptijd van het onderzoek. De onderstaande tabellen geven een ondergrens aan de power, omdat de gebruikte statistische modellen conservatief zijn en minder gegevens gebruiken dan beschikbaar. De tabellen kunnen daarom gebruikt worden bij het bepalen van de schaal en looptijd van elk experiment.

Bij het HiT gedeelte van het onderzoek zijn er 315 zorgaanbieders die voldoende groot zijn om mee te doen. Deze zorgaanbieders worden onderverdeeld in drie groepen. Gemiddeld doet een zorgaanbieder per maand 15 aanvragen, waarvan 25% getoetst wordt. We gaan uit van een conformpercentage van 85%. In de tabel hieronder laten we zien wat de power is als de andere wijze van toetsing een effect heeft op het conformpercentage van 2,5%, 5%, 7,5% of 10% punt. We laten dit zien voor verschillende looptijden van het onderzoek.

Tabel A.1: Resultaten van de simulatiestudie voor het HiT gedeelte van het onderzoek.

Looptijd	2,5%	5%	7,5%	10%
1 maand	0,10	0,25	0,55	0,85
2 maanden	0,15	0,55	0,89	1,00
3 maanden	0,23	0,74	0,98	1,00
4 maanden	0,29	0,84	0,99	1,00
5 maanden	0,34	0,88	1,00	1,00
6 maanden	0,39	0,93	1,00	1,00

Gerapporteerd is de kans op het vinden van een significant effect als het werkelijke effect van een wijziging in het toetsingsmoment gelijk is aan 2,5%, 5%, 7,5% of 10% punt, voor diverse looptijden van het onderzoek.

Bij de SIPs zijn momenteel twee toetsingsprotocollen, normale toetsing (6%) en striktere toetsing (16%). We gaan er vanuit dat er 240 zorgaanbieders zijn die verdeeld worden over vier groepen, dus dat betekent 60 aanbieders per groep. Voor Sips die vallen onder de normale toetsing hebben de aanbieders gemiddeld 16 aanvragen per maand. Voor Sips die vallen onder de striktere toetsing is dat 44. We gaan uit van een conformpercentage van 80%. Evenals bij de HiTs, geeft de onderstaande tabel aan hoe groot de kans is dat we een significant effect vinden als het werkelijke effect van een verhoging van het

toetsingspercentage (naar 10% en 25%) gelijk is aan 2,5%, 5%, 7,5% of 10% punt. We laten dit wederom zien voor verschillende looptijden van het onderzoek.

Tabel A.2: Resultaten van de simulatiestudie voor het SIP gedeelte van het onderzoek.

Looptijd	2,5%	5%	7,5%	10%
<i>Normale toetsing</i>				
1 maand	0,08	0,13	0,21	0,37
2 maanden	0,07	0,14	0,35	0,56
3 maanden	0,08	0,20	0,47	0,71
4 maanden	0,10	0,27	0,56	0,87
5 maanden	0,11	0,33	0,67	0,94
6 maanden	0,15	0,43	0,79	0,96
<i>Strikte toetsing</i>				
1 maand	0,14	0,53	0,87	0,99
2 maanden	0,28	0,84	0,99	1,00
3 maanden	0,42	0,95	1,00	1,00
4 maanden	0,55	0,98	1,00	1,00
5 maanden	0,65	1,00	1,00	1,00
6 maanden	0,73	1,00	1,00	1,00

Weergegeven is de kans op het vinden van een significant effect als het werkelijke effect van een verhoging van het toetsingspercentage (naar 10% en 25%) gelijk is aan 2,5%, 5%, 7,5% of 10% punt, voor diverse looptijden van het onderzoek.

Tegen het einde van het experiment hebben we de resultaten van de simulatiestudie naast de verzamelde data gelegd om te kunnen bepalen of er voldoende data is verzameld om het experiment te beëindigen. Hiervoor moet de grootte van zorgaanbieders, in termen van aantal aanvragen, in de data over de looptijd van het experiment worden vergeleken met de benodigde aantallen zorgaanbieders en grootte van zorgaanbieders volgens de poweranalyse. Hieruit is gebleken dat inderdaad voldoende gegevens zijn verzameld.

B Indeling van zorgaanbieders over de experimentele groepen

Tabel B.1 bevat de indeling van zorgaanbieders die zijn geselecteerd voor de participatie in het eerste onderdeel van het experiment over de drie experimentele groepen.

Tabel B.1: Indeling van zorgaanbieders in HiT groepen.

HiT controle		HiT achteraf		HiT prestatie	
<i>Naam</i>	<i>Code</i>	<i>Naam</i>	<i>Code</i>	<i>Naam</i>	<i>Code</i>
SDW Zorg	12658	Stichting de Tussenvoorziening	17543	Stichting Zozijn	14222
Reinaerde loket	11682	Carint Reggeland Groep	11650	St Thuiszorg en Maatschappelijk werk Rivierenland	10365
Pluryn loc Werkenrode	12752	Stichting Gehandicaptenzorg	13913	Thuiszorg Groningen loc Mercator	13059
Woonzorg Unie Veluwe	10354	RIBW Midden Brabant	13764	SIG	13995
Frion	14174	Zorggroep Oude en Nieuwe Land	10206	Bavo Europoort	11451
Labyrint Zorg	16762	Arienszorgpalet	12035	St Zorgzaam Zeeuws Vlaanderen	11742
JP van den Bent Stichting (ERAI)	14427	Leekerweide	10202	Beeuwkes Thuiszorg Ede	12754
Mediant GGZ	13798	Goede Ree	10175	Thebe Zuid Oost - Service Centrum	10477
Leger des Heils GWW Den Haag	11493	Sjaloomzorg	13451	Robert Coppes Stichting 73732816	14163
Esdege Reigersdaal afd De Ronde	14403	Allerzorg 751270	13844	Promens Care onderdeel RIBW Drenthe	11822
Amaris Zorggroep De Ark	16983	Arkin loc Domselaerstraat	13257	Catharina Stichting	10562
Actiezorg	10133	Emergis	12309	Verian	11695
Stichting Groenhuysen t groene thuis	12302	Stichting De Zorgboog (ERAI)	10248	Moveoo	12297
Daelzicht	14085	Bartimeus	11804	Zorggroep Ter Weel	11761
Orbis Thuiszorg	18612	Zorggroep Apeldoorn loc Randerode	11927	Zorggroep Almere	11659
Stichting Samen Zorgen	13793	Cicerozorggroep	14952	Odibaan	10338
Stichting Coloriet	10200	RIBW Heuvelland en Maasvallei	12180	St Gerlachus	10471
Sherpa	13614	Zorginstellingen Pieter van Foreest	11791	St Joris	12228
GGZ Breburg Groep Midden Brabant	13290	Stichting De Driestroom	13610	Valkenhof	10559
Novadic Kentron	14393	sno Traverse	10303	Het Friese Land (ERAI)	12110
Kwintes	10322	Zorggroep Thuis	16952	GGZ Noord Holland Noord	10178
Limor	11951	Zorggroep Thuis	11656	Amstelring	11906
Yulius	10387	Evean zorgconsulenten IM Amsterdam	11760	Stichting Icare Drenthe	17432

Vervolgd op de volgende pagina

Tabel B.1 – vervolg HiT controle		HiT achteraf		HiT prestatie	
Naam	Code	Naam	Code	Naam	Code
Neos	11683	Abrona (ERAI)	14475	Interakt Contour 60605924	12664
s Heeren Loo Midden Nederland	12519	Schreuderhuizen	13734	Tragel Zorg	15762
ActiVite	10036	Nieuw Woelwijck	12442	RIBW Fonteyenburg	11774
s Heerenloo Rivierenland	12500	Stichting De Boei	10110	Directzorg Nederland BV	13171
Livio Zorg aan Huis	11726	Stichting Azora	11789	De Bijnkershoek	10564
thuiszorg Pantein (ERAI)	11847	Stichting Verpleeghuis Het Parkhuis	10351	Woonzorgnet BV	13803
De Boskamp woonzorg unie Veluwe	18752	De Hartekamp Groep	10118	Amerpoort	11541
IrisZorg Cluster Noord	12327	Zienn opvang en ondersteuning	11871	SVVE de Archipel	10103
Pameijer	11859	Zuidzorg	10126	GGNet	13643
Stichting Het Gors	13868	RIBW Zwolle eo	11950	SVRZ loc De Vurssche	12234
Van Neijnselgroep	10452	RIBW Arnhem en Veluwe Vallei	13682	Quintes Thuiszorg	12541
Stichting Zorgstroom	10446	Stichting werkt voor Ouderen	11681	Zorgbedrijf Parnassia	14198
IPSE de Bruggen	12242	Stichting Zorgcentra Zuidwest Drenthe	10314	Zuwe Thuiszorg (ERAI)	10260
Thuiszorg Noord West Twente	10941	Thebe Thuiszorg Midden Brabant BV	10296	Stichting Palet loc Swettehiem (ERAI)	12083
Stichting De Wijngaerd	11692	Respect Zorggroep Scheveningen	11620	RST Zorgverleners	11714
Prinsentichting	16922	Zonnehuisgroep Noord	16273	SMO Breda e.o.	11790
DSV wonen welzijn zorg	11482	Gemiva-SVG Groep	10187	Allevo	12040
Siza Dorp Groep	13745	Stichting Anton Constandse	12718	Stichting Thuiszorg De Versterking	14483
Zorggroep Drenthe De Vijverhof	10911	ZINN	10209	Thuiszorg De Friese Wouden (ERAI)	11910
Careyn	16882	Stichting De Wielborgh	11071	Zideris	14208
Mosae thuiszorg	11618	Leger des Heils MC Rotterdam	12395	Stichting Prisma (ERAI)	12198
Thuiszorg van Oranje	10435	UMCG thuis	11651	Interakt Contour 65656922	12638
De Zorgzaak	10161	Clementia Zorgverleners	14853	RIBW Kennemerland Amstelland en de Meerlanden	11472
St Talant	12153	Eleos	11615	Zorggroep Meander	10167
Zonnehuisgroep Amstelland locatie Menno Simons	10871	Sensis Loc Vlasborch	13407	Delta Psychiatrisch Centrum	12895
Avondrust	12058	Emeritus BV	13004	Stichting Marga Klompe	11873
Stichting RIMO	10113	WWZ Marienstaete Valent (ERAI)	13136	Proteion Thuiszorg Noord Limburg	10305
St Zorgcentra de Kempen	10107	ARBE Dienstverlening	12352	Proteion Roncalli	12120
Stichting Zorgspectrum	12183	Stichting Zorgcentra Betuweland	10457	Orbis zorgconcern	10491
Rivierduinen	11592	Zorggroep Elde	10364	Beter Thuis Wonen	10139
Stichting Amarant	12865	Stichting De Trans	11874	Stichting ViVa Zorggroep	11920
Zorggroep St Maarten	11724	Stichting Zorg Ondersteuning Nederland	12536	VVT Zorg B.V.	11693
Vitalis Zorggroep locatie Wissehaege	12722	Boogh	14242	Stichting Icare	10535
Het Schild	14064	Zorgcentra Pro Seniore	11982	TZ Zuid-Kennemerland - Zorgbalans	10042
Altrecht GGZ	10462	Stichting Centrum Vaartserijn	16633	Trimenzo zorg aan huis	16672

Vervolgd op de volgende pagina

Tabel B.1 – vervolg		HiT controle		HiT achteraf		HiT prestatie	
Naam	Code	Naam	Code	Naam	Code	Naam	Code
Oranjehaeve	12041	Laurens		11866	Stichting Land van Horne regio Zuid Oost Brabant		11755
Meerkanten GGZ Flevo Veluwe	11845	Leger des Heils GWC Amsterdam		11637	Stichting Elisabeth		12130
Stichting Thuiszorg Diensten Centrale	10115	Zorgcombinatie Zwolle		10034	Inovum		11490
Stichting Buitenamstel Geestgronden	12659	Iriszorg		10301	Evean zorgconsulenten EM Zaanstreek Waterland		11746
Buurtzorg Nederland (ERAI)	10342	Woon Zorgcentrum Haaglanden (ERAI)		17442	St Humanitas DMH		11492
Stichting PerspeKtief	11351	Stichting Thuiszorgservice Groningen		16022	Zuidoostzorg		12015
Stichting Estinea	12296	Stichting Quarijn		10403	SBWU		13236
Vivent Thuiszorg	10254	Rivas Zorggroep		10444	Lentis		11815
Radar	10791	Stichting Baalderborggroep		14162	Voorzet Wonen		15092
Stichting De Linde	13609	Zorggroep Kinderman		14205	Thebe Mark en Maasmond		10571
Stichting De Gelderhorst	14139	Savant Zorg		10106	PrivaZorg		10261
Cello 2	12485	Stichting Zorgcombinatie Interzorg afdeling PCS Th		12039	Vivantes Zorggroep		10383
Stichting Maatschappelijke Opvang Helmond	11664	Reinier van Arkelgroep		12693	Haagse Wijk en WoonZorg (ERAI)		13158
Omring Thuiszorg West Friesland	11539	Dichterbij		10215	NSWAC		12917
Vitalis Zorggroep locatie Brunswijck Peppelrode VP	11593	Stichting Triade		12191	Stichting ORO		13454
Hoog Anstel	11783	Severinus		12318	GGZ Regio Breda		12145
Zorggroep De Leiboom	11891	Sovak		12107	HVO Querido		12499
RIBW ZWWF	10124	Woonzorgconcern IJsselheem		10172	Bouman GGZ		13274
Zorgwaard	17302	Zorgcentrum De Posten		12534	Woonzorgcentrum Groot Hoogwaak		10479
Leger des Heils CWZW Flevoland	10881	GGZ Drenthe		12101	GGZ Westelijk Noord Brabant		12782
Stichting Waardeburgh locatie Overslydrecht	12634	Verpleeghuis Innersdijk		12214	ZorgSter		17412
Stichting Land van Horne regio Noord en Midden Lim	11670	NOVO		13002	De Bleerinck		14179
Zorggroep Charim	11710	Profila Zorg		14426	Thuiszorg Sensire		13167
Vierstroomzorgring	11509	Stichting Reakt		10331	De Pieter Raat Stichting		10382
RIBW Nijmegen & Rivierenlnd	10487	Stichting Humanitas		16912	Zorgcirkel		10158
Odion	10631	Stichting RIBW Zorg en Welzijn		12985	Carinova loc Raalte		10032
GroenekruisDomicura	10449	Zorgcentrum Geldershof		12869	De Mieden		12851
De Noorderbrug 66660202	13311	Mensana RIBW Noord en Midden Limburg		11926	Maasveld		14414
Careander	12393	Stichting Vilente (ERAI)		12081	St Burgemeester van Julsinghatehuis		13008
Beweging 3.0	11488	Stichting Fatima		14457	Woonzorggroep Wilgaerden Avondlicht		11813
Amsterdam Thuiszorg	10681	Driezorg Rivierenhof		11875	MensEnZorg		11924
Zorggroep Noord en Midden Limburg loc Zorg aan Hui	11985	Viataal		13672	Stichting NOVIZorg		14243
Agathos Thuiszorg	11576	Thuiszorg West Brabant		10464	Stichting De Binnenvest		10255
Internos katholieke thuiszorg	10455	Vivium zorggroep TGV RIBW		12641	Stichting Thuiszorg Breda		10318
Opella (ERAI)	10320	Leger des Heils CWZW Zuid West Nederland		11221	St Ons Tweede Thuis		14374

Vervolg op de volgende pagina

Tabel B.1 – vervolg HiT controle			HiT achteraf			HiT prestatie		
Naam	Code	Naam	Code	Naam	Code	Naam	Code	
Promens Care onderdeel de Leite	12047	Stichting Ruitersbos	11687	Mondriaan			10558	
Zrgp Noord en Midden Limburg loc Tml	12166	Topaz	14415	Clara Feyeona Heem			12478	
Lunet Zorg 30300156	13503	GGZ Friesland	14345	Cordaan RIBW			11992	
Circonflex	11680	Stichting Philadelphia Zorg	17632	Steinmetz de Compaan			13620	
Leger des Heils CWZW Noord	12411	TriviumMeulenbeltZorg	12026	Thuiszorg Rotterdam BV			10337	
Stichting Woonzorgcentra Flevoland	10312	Libertas Leiden	11251	Stichting Wassenaarse Zorgverlening			12001	
Thuiszorg Groot Gelre BV	11641	ZZG zorggroep	11776	St Verpleeghuis Eikendonk			10611	
Diafaan	11545	Stichting Oosterlengte	10771	ASVZ			12712	
GGzE	12232	Stichting Thuiszorg Midden Gelderland	10266	Stichting Onder Een Dak			13618	
Stichting Brentano Amstelveen loc Zorgbemiddeling	12440	Woonzorggroep Wilgaerden Bosmanstichting	11759	De Schutse 420141			13809	
Omring Thuiszorg Kop van Noord Holland	11525	Aveant (ERAI)	10359	Woon Zorgcentrum De Breedonk			12095	
Syndion	10211	t Heem	10375	Interzorg Thuiszorg			12216	

Tabel B.2 bevat de indeling van zorgaanbieders die zijn geselecteerd voor deelname in het tweede gedeelte van het experiment over de vier experimentele groepen in dat deel van het onderzoek. De namen van zorgaanbieders zoals weergegeven in de tabel kunnen wijzigen over tijd, de namen zoals deze voor aanvang van het onderzoek waren geregistreerd zijn hier gerapporteerd. In beide tabellen zijn de zorgaanbieders die deelnemen aan ERAI rood gemaakt.

Tabel B.2: Indeling van zorgaanbieders in SIP groepen.

SIP controle		SIP laag		SIP hoog		SIP prestatie	
Naam	Code	Naam	Code	Naam	Code	Naam	Code
Rinette Zorg	10151	Maasstad Ziekenhuis	12117	St. Antonius Ziekenhuis	13088	Thuiszorg Noord West Twente	10941
Gelre ziekenhuizen locatie Zutphen	11531	Opella (ERAI)	10320	Thuiszorg de Zellingen	10184	Stichting Hilverzorg	12250
PrivaZorg	10261	Sint Franciscus Gasthuis	11480	Carint Reggeland Groep	11650	Lievensberg Ziekenhuis	13744
Zuidzorg	10126	Elisabethziekenhuis Tilburg	11704	Orbis Medisch Centrum	18532	GroenekruisDomicura	10449
Het Van Weel - Bethesda Ziekenhuis	11923	Haga Ziekenhuis locatie Leyweg (ERAI)	12027	Evean zorgconsulenten EM Amsterdam	11988	Stichting Thuiszorg Diakonie	12292
Amaris Zorggroep De Ark	16983	Ropcke Zweers Ziekenhuis	12064	Diaconessenhuis Leiden	10140	NKI AVL	12082

Vervolgd op de volgende pagina

SIP controle		SIP laag		SIP hoog		SIP prestatie	
Naam	Code	Naam	Code	Naam	Code	Naam	Code
Maartenskliniek	12386	Vitras CMD (ERAI)	10198	Haagse Wijk en WoonZorg (ERAI)	13158	BrabantZorg	11474
Ziekenhuis Rivierenland Tiel	11863	Maxima Medisch Centrum	12498	Stichting Icare Flevoland (ERAI)	17862	Stichting De Zorgboog (ERAI)	10248
Leveste locatie Scheper Ziekenhuis	10546	Franciscus Ziekenhuis	11961	Spaarne Ziekenhuis	14168	ZorgSaam Ziekenhuis	12121
Medisch Centrum Alkmaar	10203	Buurtzorg Nederland (ERAI)	10342	Vivent Thuiszorg	10254	St Jans Gasthuis Weert	11877
UMC St Radboud	11792	Vlietland Ziekenhuis	11291	Rijnland Zorggroep locatie Verzorging en Verpleging	17542	Steinmetz de Compaan, Rubroek en Provenier	10601
Stichting ViVa Zorggroep	11920	Carinova loc Raalte	10032	Stichting De Bilthuysen	10183	Stichting Thuiszorg Breda	10318
MC Haaglanden locatie Westeinde (ERAI)	11572	Evean zorgconsulenten IM Amsterdam	11760	ZAB Zorg	17552	Universitair Medisch Centrum Groningen UMCG	13124
Sint Lucas Andreas Ziekenhuis	15402	Zonnehuisgroep Noord	16273	ZGT Hengelo	13604	Verpleeghuis Anholt	11547
Thuiszorg van Oranje	10435	Verpleeghuis Innersdijk	12214	Groene Hart Ziekenhuis	10292	Havenziekenhuis	11844
Aafje	13010	Zorggroep Thuis	11656	Laurens	11866	St Residentiele en Ambulante Zorg	10343
Pasana	11916	Zuidoostzorg	12015	Thuiszorg Sensire	13167	ZKH Gelderse Vallei	11686
Clementia Zorgverleners	14853	Zorggroep St Maarten	11724	Ambulante Thuiszorg	12568	Academisch Ziekenhuis Maastricht	10334
Zorgcombinatie Nieuwe Maas	11857	Stichting Thuiszorg Diensten Centrale	10115	Erasmus MC	12099	Avondrust	12058
Zorggroep Elde	10364	Isala Klinieken	10271	Ruwaard van Putten ZKH	11484	Elkerliek ziekenhuis	10194
Actiezorg	10133	Catharina-Ziekenhuis	10378	Zorgcentrum Schoonoord	10341	Centrum voor reuma en revalidatie Rotterdam	14112
St Zorgsaam Zeeuws Vlaanderen	11742	Rode Kruis Ziekenhuis	11802	Zorgcentrum Reeuwijk	11101	Leids Universitair Medisch Centrum	10281
Leger des Heils GWC Amsterdam	11637	Atrium medisch centrum	11669	Zaans Medisch Centrum	11784	Zorggroep Almere	11659
Stichting Icare Zwolle	17534	Alerimus	11486	Omring Transfer Gemini Ziekenhuis	12595	Diakonessenhuis Utrecht	11588
Zorggroep Pasana	13567	St Zorgcentra de Kempen	10107	Thuiszorg West Brabant	10464	Thuiszorg Rotterdam BV	10337
Reinier de Graafgroep	12438	MC Groep	14287	Stichting Zorgstroom	10446	Omring Thuiszorg West Friesland	11539
Vivium zorggroep TGV - VenV teams	12629	Transferpunt AMC	11935	De Bleerinc	14179	Ziekenhuis Rijnstate	11948
Stichting Zorggarant Noord-Nederland	11973	Driezorg Rivierenhof	11875	St Joris	12228	Het Friese Land (ERAI)	12110
Stichting Icare Drenthe	17432	Sensire Verpleeghuis Den Ooiman	10346	Stg De Markenlanden	11907	Omring Thuiszorg Kop van Noord Holland	11525
Hoofdkantoor Osiragroep	11621	Meander Medisch Centrum	14372	Ikazia Ziekenhuis	10154	Zorgcombinatie Noorderboog ziekenhuis Diaconessenh	11931
Proteion Thuiszorg Noord Limburg	10305	Zorggroep Thuis	16952	Stichting Thuiszorg Midden Gelderland	10266	WWZ Marienstaete Valent (ERAI)	13136
Zorggroep Meander	10167	Orbis zorgconcern	10491	Ziekenhuis De Tjongerschans	11990	Stichting Land van Horne regio Noord en Midden Lim	11670
BovenIJziekenhuis	12691	Ziekenhuis Nij Smellinghe	12174	Amsterdam Thuiszorg	10681	Allerzorg Hilversum Gooi eo	16972
Albert Schweitzer Ziekenhuis	10391	Leger des Heils CWZW Midden NL	11518	Medisch Centrum Leeuwarden	11954	TZ Zuid-Kennemerland - Zorgbalans	10042

Vervolgd op de volgende pagina

SIP controle		SIP laag		SIP hoog		SIP prestatie	
Naam	Code	Naam	Code	Naam	Code	Naam	Code
Noorderbreedte (ERAI)	11551	Kennemer Gasthuis loket nazorg	13653	Evean zorgconsulenten EM Zaanstreek Waterland	11746	Thuiszorg De Friese Wouden (ERAI)	11910
IJssellandziekenhuis	12708	Gelre Ziekenhuizen locatie Apeldoorn	11622	UMC Utrecht Divisie Hart en Longen - Longafdeling	18092	Martini Ziekenhuis	10142
Amstelring	11906	Haga Ziekenhuis locatie Sportlaan (ERAI)	12074	St.Thuiszorgcentrale IVT	10249	Stichting Thuiszorgservice Twente	15312
Zrgp Noord en Midden Limburg loc Tml	12166	Tergooiziekenhuizen locatie Blaricum en locatie Hi	10891	Ziekenhuis St Jansdal	12105	Antonius Ziekenhuis	12531
Maasziekenhuis Boxmeer	10480	Slotervaart Ziekenhuis	13433	Ziekenhuis Rijnstate locatie Zevenaar	12252	Stichting Humanitas	16912
ActiVite	10036	Zorgcirkel	10158	Allevo	12040	Evean Thuiszorg	10132
Vierstroomzorgring	11509	Livio Zorg aan Huis	11726	Woon Zorgcentrum Haaglanden (ERAI)	17442	Stichting Azora	11789
Centraal Bureau Cordaan	11933	Actiefzorg	10196	Viecuri Medisch Centrum voor Noord Limburg	10220	Ziekenhuis Bernhoven	12774
Amphia Ziekenhuis	14063	Stichting Icare Ede (ERAI)	18152	Wilhelmina Ziekenhuis	10169	Streekziekenhuis Koningin Beatrix	12745
Zorggroep Oude en Nieuwe Land	10206	Transfer Waterland ZHS - Zorgcirkel	17034	ZZG zorggroep	11776	St Annaziekenhuis	12236
Refaja Ziekenhuis	10163	thuiszorg Pantein (ERAI)	11847	Transfer Zaans Medisch Centrum - Evean	18232	ZGT	16953
Stichting Innoforte	11684	SVVE de Archipel	10103	Stichting Thuiszorgservice Drenthe	15542	Thuiszorg Groot Gelre BV	11641
De Zorgzaak	10161	UMC Utrecht	10291	St Lucas Ziekenhuis	10368	Curadomi organisatie voor thuiszorg	11768
Thuiszorg INIS	12003	St Thuiszorg en Maatschappelijk werk Rivierenland	10365	Vitalis Zorggroep locatie Wissehaege	12722	VU medisch Centrum	10533
Verpleeghuis NEBO	11599	Woonzorgcentrum Groot Hoogwaak	10479	Deventer Ziekenhuis	11626	Canisius Wilhelmina Ziekenhuis	11731
Ziekenhuis Walcheren (ERAI)	10951	Careyn	16882	Zorggroep Noord en Midden Limburg loc Zorg aan Hui	11985	Orbis Thuiszorg	18612
Respect Zorggroep Scheveningen	11620	Stichting SHDH	12291	Allerzorg 75751058	11876	Beweging 3.0	11488
Stichting Palet loc Swettehiem (ERAI)	12083	Stichting Thuiszorgservice Groningen	16022	Zorggroep de Vechtstreek	10450	Thuiszorg Service Nederland	11629
Libertas Leiden	11251	Medisch Spectrum Twente	11872	Oncologisch Zorgbelang Limburg	14942	Zuwe Thuiszorg (ERAI)	10260
Laurentiusziekenhuis	10104	ZINN	10209	Zorgkompas (ERAI)	12387	Stichting Flevohuis	10538
Florence Thuiszorg (ERAI)	11001	Lange Land Ziekenhuis	11521	Albert van Koningsbruggen	10256	Stichting werkt voor Ouderen	11681
Thebe Mark en Maasmond	10571	Zuwe Hofpoort ziekenhuis	11411	Voorhoevehuis	10377	Beter Thuis Wonen	10139
Stichting Thuiszorg Zuidwest Friesland	12143	Het Hoge Veer	11764	Stichting Jeroen Bosch Ziekenhuis	11473	Stichting Zorgspectrum	12183
Zorggroep Noordwest-Veluwe	10160	RST Zorgverleners	11714	Rivas Zorggroep	10444	Stichting Tante Louise Vivensis ABG	12333
Bethesda ziekenhuis	10651	Slingeland Ziekenhuis	13086	Delfzicht	11869	Savant Zorg	10106
Bocholtz	11816	Thebe Thuiszorg Midden Brabant BV	10296	Duinrust	13319	Stichting Thuiszorgservice Gelderland	16472

Vervolg op de volgende pagina

SIP controle		SIP laag		SIP hoog		SIP prestatie	
<i>Naam</i>	<i>Code</i>	<i>Naam</i>	<i>Code</i>	<i>Naam</i>	<i>Code</i>	<i>Naam</i>	<i>Code</i>
Aveant (ERAI)	10359	Stichting Icare	10535	Arienszorgpalet	12035	Ziekenhuis Bronovo (ERAI)	12371
MC Haaglanden locatie Antoniushove (ERAI)	11546	Stichting Zorgcentra Rivierenland	10250	Evita Particuliere Zorg BV	11581	Tweesteden Ziekenhuis	10101
Westfriesgasthuis	10224	Stichting de Wever	16942	Stg Zorgcontact	11648	Verian	11695

C HiT typen en SIP protocolnummers

Tabel C.1 bevat een overzicht van alle voorkomende HiT typen en SIP protocolnummers in de data met daarbij een omschrijving van de betreffende zorg. Voor SIPs is bovendien de categorie aanduiding weergegeven op basis waarvan het lage of het hoge toetsingspercentage wordt gebruikt.

Belangrijk om op te merken is dat de inhoud van HiT typen en SIP nummers per jaar kan wijzigen. Met ingang van 2013 zijn HiT type 4 en de inhoud van de volgende SIP protocolnummers gewijzigd: 3, 4, 6, 9, 11, 13, 15, 20, 21, 29, 31, 70. Daarnaast is SIP 34 komen te vervallen. Ook de toegekende categorie aan een SIP nummer kan variëren.

Tabel C.1: Overzicht HiT typen en SIP protocolnummers.

HiT type	Omschrijving
Type 1	Aanpassen indicatie extramurale zorg.
Type 2	Gelijkblijvend ZZP verlengen.
Type 3	Aanpassen dagbesteding binnen ZZP.
Type 4	Aanpassen ZZP & aanpassen functiegerichte verblijfsindicatie naar ZZP.
Type 5	Pilot aanpassen ZZP.
Type 6	VB-Idu omzetten.
Onbekend	Onbekend.

SIP nummer	Omschrijving	Categorie
3	Zorg na ziekenhuisopname in verband met een operatie heup/knie (zonder verblijf).	laag
4	Wondverzorging.	laag
6	Ambulante Compressie Therapie (ACT).	laag
7	Steunkousen aan- en/of uittrekken.	laag
9	(Hulp bij) het toedienen van injecties.	laag
10	(Hulp bij) ooggel, oog-, oor- of neusdruppels/-zalf en/of orale medicatie (pillen/poeders/drankjes) en/of aanbrengen medicinale pleister.	laag
11	Pijnbestrijding subcutaan door middel van pomp.	laag
12	Sondevoeding toedienen.	laag
13	Stomaverzorging (AP of UP), verwisselen, huid- en hygiënische verzorging.	laag
14	Klysma (microlax) toedienen.	laag
15	Katheteriseren eenmalig.	laag
16	Blaasspoelen via bestaande katheter.	laag
18	Wassen en aan-/uitkleden, kortdurende persoonlijke verzorging (≥ 3 maanden).	laag
19	Hulp bij het douchen/wassen/kleden tot en met 3 keer per week voor ouderen > 75 jaar.	laag
20	Bloedsuikerdagcurve prikken.	laag
21	Verwisselen/inbrengen van een sonde.	laag
24	Palliatief terminale zorg (levensverwachting < 3 maanden).	laag
25	Gespecialiseerde intramurale zorg t.b.v. cliënten met de ziekte van Huntington.	laag
27	Behandeling kinderen met een dominante grondslag verstandelijke handicap.	laag
28	Persoonlijke verzorging (tot en met klasse 5) na ontslag uit ziekenhuis (geldigheid maximaal 6 maanden).	hoog
29	Verpleging (tot en met klasse 5) na ontslag uit ziekenhuis (geldigheid maximaal 6 maanden).	hoog
31	Tijdelijk verblijf in een intramurale setting ten behoeve van herstel na een ziekenhuisopname.	laag
34	ZZP VV 09a: geriatrische revalidatiezorg in een instelling.	hoog
35	ZZP VV 09b: herstelgerichte behandeling met verpleging en verzorging in een instelling.	hoog
70	ZZP GGZ B03, B04, B05, B06, B07: voortgezet tijdelijk verblijf met behandeling na 365 dagen opname op grond van de zorgverzekeringswet. Alleen toepasbaar als herstel verwacht wordt binnen twee jaar en gepaard gaat met ontslag uit de instelling.	laag

D Gedetailleerde beschrijving van de dataset

De definitieve dataset bevat informatie over alle aanvragen ingediend in de periode van 1 januari 2012 tot en met 7 april 2013. Voor iedere geregistreerde aanvraag observeren we, voor zover relevant, de volgende gegevens:

1. Informatie over de zorgaanbieder die de aanvraag heeft ingediend:
 - (a) Experimentele groep waarin de zorgaanbieder is ingedeeld
 - (b) Naam en AF-code van de zorgaanbieder
 - (c) Doelgroep. Dit geeft aan of het een ziekenhuis of ander soort zorginstelling betreft
 - (d) Deskundigheid. Dit geeft aan of de instelling voornamelijk gericht is op verpleging & verzorging, gehandicaptenzorg of geestelijke gezondheidszorg;
2. Aanvraagdatum en afsluitdatum van de aanvraag, en, waar relevant aanvraagdatum en afsluitdatum van het toetsingstraject;
3. Soort aanvraag: HiT, SIP, IM, spoedaanvraag, AWBZ aanvraag, ERAI aanvraag of overige aanvraag;
4. HiT type of SIP protocolnummer in geval van een HiT of SIP aanvraag;
5. Informatie over de cliënt waarvoor de aangevraagde zorg is bestemd:
 - (a) 4 cijfers van de postcode
 - (b) Leeftijd op het moment van ingang van de indicatie
 - (c) Geslacht;
6. Grondslag voor AWBZ zorg. Dit is een van de volgende categorieën: somatische aandoening/beperking, psychogeriatrische aandoening/beperking, psychiatische aandoening/beperking, lichamelijke handicap, verstandelijke handicap of zintuiglijke handicap;
7. Gevraagde en geïndiceerde zorgfuncties. Dit is een (combinatie) van de volgende functies: persoonlijke verzorging (PV), verpleging (VP), begeleiding individueel (BGI), begeleiding groep (BGG), behandeling individueel (BHI), behandeling groep (BHG), kortdurend verblijf (KVB), intramurale zorg;
8. Gevraagde en geïndiceerde zorgomvang: gevraagde en geïndiceerde klassen (extramurale zorg) of gevraagd en geïndiceerd zorgzwaartepakket (intramurale zorg). Voor geïndiceerde zorgomvang observeren we eveneens de volgtijdigheid, dat wil zeggen de periode waarvoor een bepaalde klasse of een bepaald zorgzwaartepakket is geïndiceerd;
9. Geldende toetsingspraktijk (toetsingspercentage en moment van toetsing);
10. Voor HiT en SIP aanvragen: een indicator die aangeeft of de aanvraag al dan niet getoetst is.

Voor getoetste aanvragen observeren we daarnaast nog het volgende:
11. Toetsingsresultaat: conform of niet conform
12. Reden voor toetsingsresultaat. De volgende redenen voor een bepaald toetsingsre-

sultaat zijn mogelijk voor HiT aanvragen:

- (a) Geen afwijkende verbijzonderingen of beperkingen
- (b) Afwijkende verbijzonderingen, niet leidend tot een ander besluit
- (c) Afwijkende beperkingen, niet leidend tot een ander besluit
- (d) Afwijkende grondslag
- (e) Afwijkende leveringsvoorwaarden
- (f) Afwijkende functies / ZZP / dagbesteding bij ZZP
- (g) Afwijkende klassen: Klasse is te hoog aangevraagd
- (h) Afwijkende klassen: Klasse is te laag aangevraagd
- (i) Afwijkende verbijzonderingen
- (j) Afwijkende beperkingen
- (k) Niet integraal

Voor SIP aanvragen worden de volgende redenen onderscheiden:

- (a) Beschrijving SIP n.v.t. op cliënt
- (b) Kernvragen / rekenhulp niet juist beantwoord
- (c) Beschrijving n.v.t. EN kernvragen / rekenhulp niet juist

13. Detail reden voor niet conform toetsing. Hierbij kunnen een of meerdere van de volgende detail redenen worden geselecteerd voor HiT aanvragen:

- (a) Relatie ziekte/aandoening/stoornis/beperking
- (b) Woon- en leefsituatie
- (c) Behandeling/revalidatie/aanpassing van de omgeving
- (d) Gebruikelijke zorg
- (e) Wettelijk voorliggende voorzieningen
- (f) Algemeen gebruikelijke voorzieningen
- (g) Mantelzorg
- (h) Toegang tot begeleiding
- (i) Niet van toepassing
- (j) Overige

Voor SIP aanvragen kan naast bovengenoemde redenen (a) tot en met (g) (ook) een (of meerdere) van de volgende opties worden opgegeven:

- (a) Verbijzonderingen en beperkingen
- (b) Grondslag
- (c) Aanleren van de zorg
- (d) Leeftijd cliënt
- (e) Frequentie en/of omvang van de zorg te laag
- (f) Frequentie en/of omvang van de zorg te hoog
- (g) Prognose
- (h) Indicatiecriteria functie of ZZP
- (i) Vragen specifiek gerelateerd aan de SIP
- (j) Huidige geldige indicatie
- (k) Overige

E Toepassing van selectiecriteria op de dataset

We passen een aantal selectiecriteria toe op de uiteindelijke dataset. Dit leidt tot verliezen in aantallen aanvragen zoals weergegeven in Tabel E.1. De eerste selectie is het gevolg van jaarlijkse (inhoudelijke) wijzigingen in de SIP protocollen, bijvoorbeeld door verschuivingen van bepaalde soorten zorg naar de Zorgverzekeringswet (Zvw) of de Wet maatschappelijke ondersteuning (Wmo). Ten opzichte van 2012 is in 2013 één indicatieprotocol komen te vervallen, te weten SIP 34. Om de gegevens over 2012 en 2013 vergelijkbaar te houden, worden SIP 34 aanvragen uit de data verwijderd. In de periode voor de start van het experiment was 14,3% van alle SIP aanvragen ingediend door een zorgaanbieder in een van de SIP groepen, een SIP 34 aanvraag.

Ten tweede verwijderen we aanvragen voor SIP 35, omdat dit SIP protocolnummer, zoals eerder aangegeven, vanaf 1 januari 2013 niet langer onderdeel uitmaakt van het experiment. In de periode voor de start van het experiment was slechts 1,1% van alle SIP aanvragen ingediend door een zorgaanbieder in één van de SIP groepen, een SIP 35 aanvraag.

Er is een tweetal zorgaanbieders in de SIP groepen met alleen SIP 34 of SIP 35 aanvragen. Deze twee zorgaanbieders hebben wel HiT aanvragen en indicatiemeldingen ingediend, maar vanwege de focus op deze twee soorten SIP aanvragen, is het niet wenselijk om deze zorgaanbieders mee te wegen in de meting van effecten van de variatie in toetsingspercentages in de SIP groepen. Dat betekent dat we één zorgaanbieder in de laag toetsingspercentage groep en één zorgaanbieder in de hoog toetsingspercentage groep verliezen. In totaal zijn door deze zorgaanbieders, afgezien van SIP 34 en SIP 35 aanvragen, 319 aanvragen ingediend gedurende de periode waarover we data hebben.

Ten derde verwijderen we ERAI zorgaanbieders uit de data, omdat deze zorgaanbieders na de start van ERAI niet langer HiT of SIP aanvragen indienen.³⁰ De ERAI zorgaanbieders in de HiT (SIP) groepen zijn, vóór de start van het experiment, verantwoordelijk voor 10,3% (15,0%) van de ingediende HiT (SIP) aanvragen (exclusief SIP 34 en SIP 35).³¹

Ten vierde heeft een aantal zorgaanbieders sinds een bepaald moment gedurende het ex-

³⁰Er is één zorgaanbieder, onderdeel van zowel de HiT prestatie groep als de SIP prestatie groep, die pas sinds 1 april 2013 is gestart met deelname aan ERAI. Omdat we door de constructie van tweeweeks periodes deze laatste week aan data in de analyses niet benutten, nemen we deze zorgaanbieder in principe wel mee in de analyses. Deze zorgaanbieder heeft 27 ERAI aanvragen ingediend.

³¹Het aandeel van ERAI zorgaanbieders in het totaal aantal HiT (SIP) aanvragen verschilt aanzienlijk over HiT (SIP) groepen. In de SIP laag toetsingspercentage groep is het aandeel van ERAI zorgaanbieders 26,2%. Dit is voornamelijk te wijten aan Buurtzorg (AF-code 10342), een grote zorgaanbieder in termen van aantal ingediende SIPs voor de start van ERAI. Daarentegen is het aandeel van ERAI zorgaanbieders in het aantal SIP aanvragen slechts 3,7% in de SIP hoog toetsingspercentage groep. Voor HiT groepen loopt het aandeel van ERAI zorgaanbieders uiteen van 8,9% in de HiT achteraf groep tot 13,1% in de HiT prestatie groep.

periment tot het einde van het experiment geen aanvraag meer ingediend, bijvoorbeeld als gevolg van beëindiging of samenvoeging van AF (Aanmeldfunctionaliteit)-contracten. We hebben niet direct informatie over welke zorgaanbieders dit betreft en bovendien weten we niet exact per wanneer deze zorgaanbieders inactief zijn geworden. Wanneer we de gegevens over het aantal ingediende HiT of SIP aanvragen per week bekijken, zien we dat er ook tussenliggende periodes zijn waarin zorgaanbieders geen aanvragen indienen. Bij het definiëren van uitgevallen zorgaanbieders houden we rekening met het historische patroon in het wekelijks aantal ingediende aanvragen. Meer specifiek gebruiken we de volgende definitie: een zorgaanbieder wordt aangemerkt als zijnde ‘uitgevallen’ wanneer de duur van de aaneengesloten periode zonder aanvragen tot het einde van de periode waarover we data hebben (dus tot en met 7 april) minstens twee weken langer is dan de maximale lengte van een tussenliggende periode zonder aanvragen. Voor tussenliggende periodes zonder aanvragen, wordt het aantal aanvragen (dat later gebruikt wordt in de statistische analyses) op 0 gezet. Daarentegen wordt een missende waarde toegekend voor weken zonder aanvragen aan het eind van de periode waarover we gegevens hebben. Als robuustheidsanalyse zullen we de regressies ook uitvoeren op gegevens waaruit de zorgaanbieders die als zijnde uitgevallen worden gedefinieerd, volledig zijn verwijderd, zelfs in de periodes dat ze wel aanvragen hebben ingediend. Tevens zullen we als robuustheidstest een striktere definitie van een inactieve zorgaanbieder hanteren. Meer details hierover worden besproken in sectie 6.3.

De hierboven gegeven definitie van een uitgevallen zorgaanbieder impliceert dat 18 zorgaanbieders in de HiT groepen en 12 zorgaanbieders in de SIP groepen als uitgevallen worden aangemerkt. Om precies te zijn gaat het om 11 zorgaanbieders in de HiT controle groep, drie zorgaanbieders in de HiT achteraf toetsing groep en vier zorgaanbieders in de HiT prestatie groep. Voor de SIP groepen gaat het om drie zorgaanbieders in de controlegroep, één (vier) zorgaanbieder(s) in de laag (hoog) toetsingspercentage groep en vier zorgaanbieders in de SIP prestatie groep. Het verwijderen van alle aanvragen van uitgevallen zorgaanbieders uit de data, zoals in de robuustheidsanalyses in sectie gedaan zal worden, leidt tot een verdere reductie van het aantal aanvragen met 3,9%, zoals is gerapporteerd in Tabel E.1.

Tabel E.1: Selectiecriteria en verlies in aantal aanvragen.

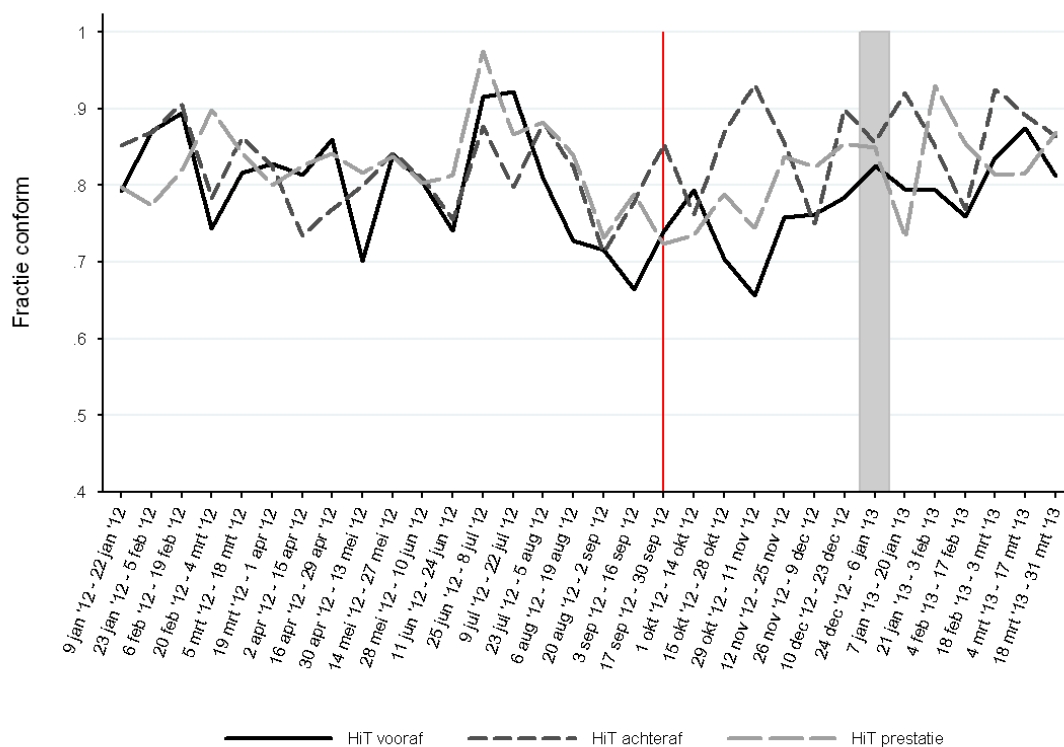
Selectie criterium	Aantal aanvragen resterend	Aantal (%) verlies	
<i>Initieel aantal aanvragen</i>	1.200.825		
zonder SIP 34	1.156.291	44.534	(3,7%)
zonder SIP 35	1.152.019	4.272	(0,4%)
zonder ERAI zorgaanbieders	950.707	201.312	(17,5%)
zonder SIP 34 en SIP 35 zorgaanbieders	950.388	319	(0,03%)
Uitgevallen zorgaanbieders	913.497	36.891	(3,9%)

F Fractie conform per HiT type

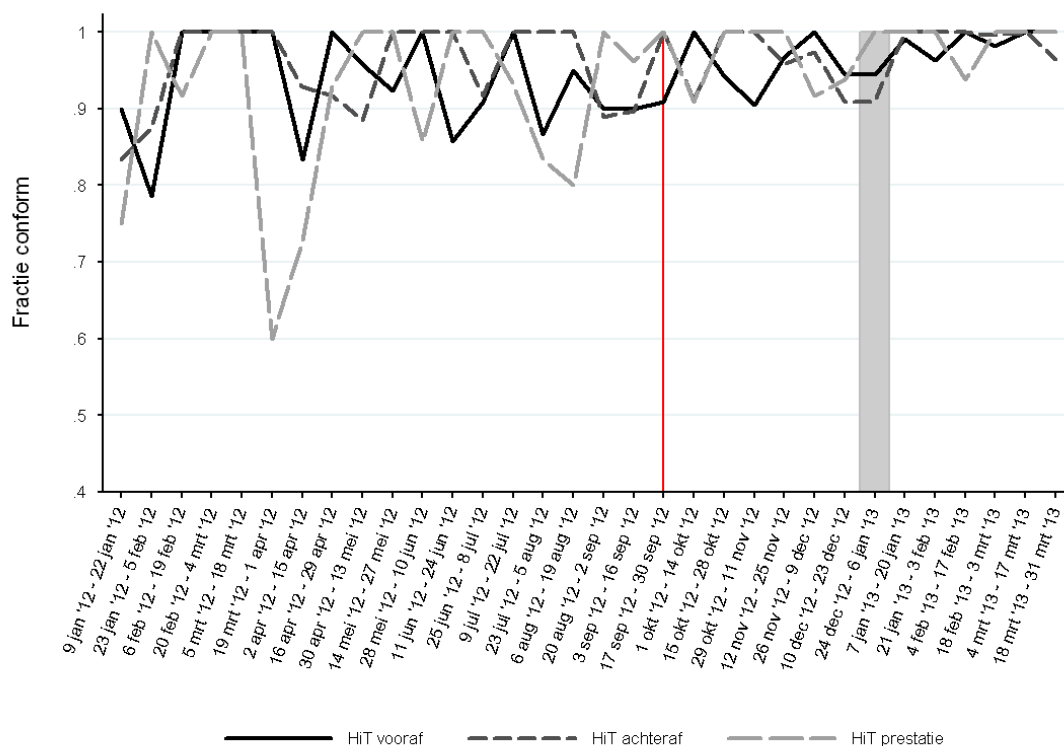
De figuren in deze appendix geven voor iedere type HiT aanvraag de gemiddelde fractie conform getoetste aanvragen over zorgaanbieders per periode van 2 weken. Door deze fractie conform per type HiT weer te geven kan worden bekeken of er sterke verschillen zijn in conformpercentage voor diverse HiT types en of dit een verklaring kan vormen voor de sterke fluctuatie in de conformpercentages die is gevonden in Figuur 5.5 in sectie 5.2.3. Belangrijk om op te merken is dat het aantal toetsingen waarop de diverse figuren zijn gebaseerd sterk verschilt: van de in totaal 16.876 getoetste HiT aanvragen in één van de HiT groepen is 73% een HiT type 1 aanvraag, 12% een HiT type 2 aanvraag, 12% een HiT type 4 aanvraag en 4% een HiT type 3 of 6 aanvraag.

Wanneer we de figuren voor de verschillende typen HiT aanvragen bekijken, valt op dat het patroon in Figuur 5.5 in sectie 5.2.3 voornamelijk bepaald wordt door HiT type 1, aangezien dat het meest voorkomende type aanvraag is, terwijl de ontwikkeling in de fractie conform getoetste aanvragen voor de overige types daar behoorlijk van afwijkt. Zo lijkt bijvoorbeeld in Figuur F.2 wel degelijk een iets stijgende trend in de fractie conform getoetste aanvragen te ontstaan sinds de start van het experiment, alhoewel deze trend in alle groepen, inclusief de controlegroep, in gelijke mate lijkt op te treden. Bovendien ligt het conformpercentage voor dit type aanvraag in iedere periode wat hoger dan dat voor alle HiT aanvragen bij elkaar.

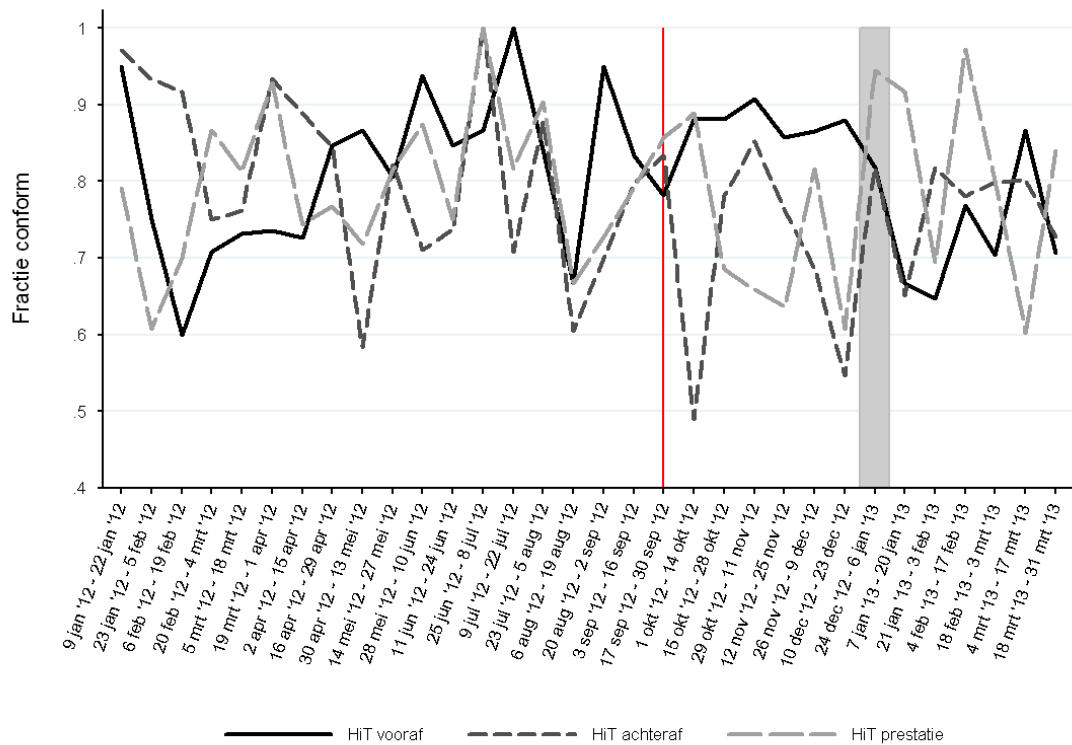
Figuur F.1: Fractie conform getoetste HiT type 1 aanvragen, per HiT groep.



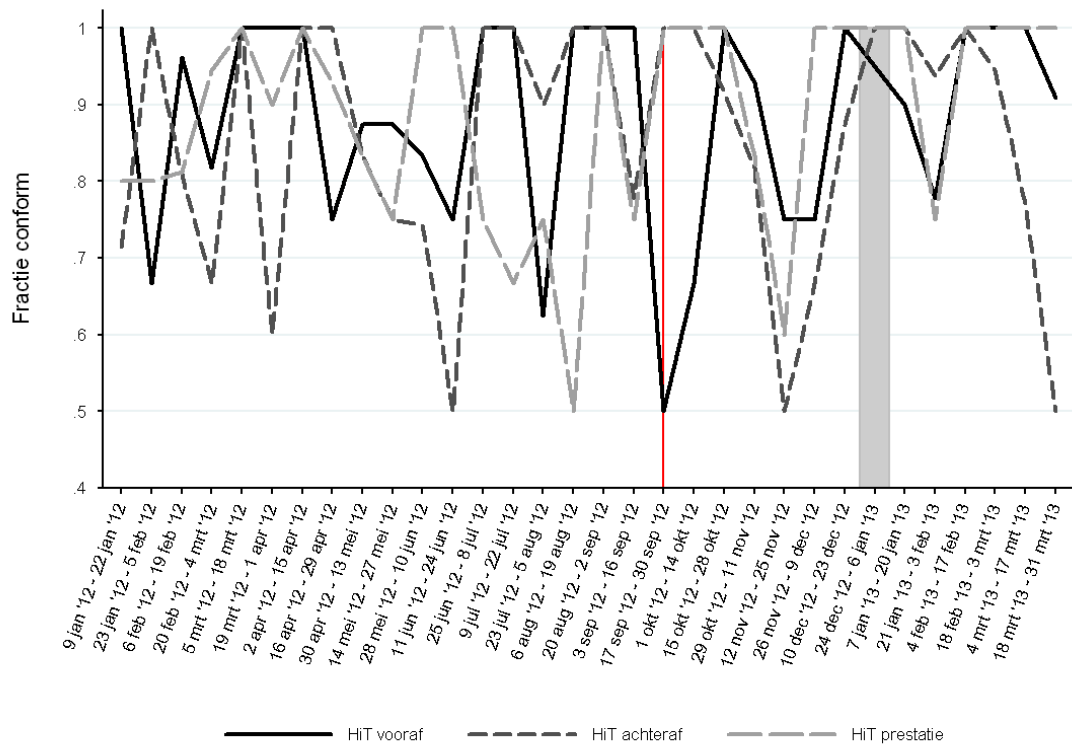
Figuur F.2: Fractie conform getoetste HiT type 2 aanvragen, per HiT groep.



Figuur F.3: Fractie conform getoetste HiT type 4 aanvragen, per HiT groep.



Figuur F.4: Fractie conform getoetste overige HiT aanvragen (type 3 en 6), per HiT groep.



G Aanvullende regressieresultaten

Tabel G.1: Robuustheidsanalyses: polynoom in tijd.

Panel A: Variatie in het toetsingsmoment.						
	# HiT aanvragen		# HiT toetsingen		fractie conform HiT	
HiT achteraf, 2012	0,254	(0,257)	0,167**	(0,069)	0,005	(0,021)
HiT achteraf, 2013	1,178**	(0,579)	0,414***	(0,152)	0,036*	(0,020)
HiT prestatie, 2012	-0,233*	(0,134)	0,042	(0,047)	-0,015	(0,021)
HiT prestatie, 2013	-0,398	(0,304)	0,023	(0,083)	0,005	(0,023)
gem. controlegr.	3,381	(0,092)	0,769	(0,025)	0,823	(0,009)
R^2	0,751		0,598		0,092	
aantal obs.	9.286		9.286		5.727	
aantal ZABs	299		299		299	
Panel B: Variatie in het toetsingspercentage.						
	# SIP aanvragen		# SIP toetsingen		fractie conform SIP	
SIP laag, 2012	-0,446	(0,747)	-0,771***	(0,114)	-0,002	(0,024)
SIP laag, 2013	-0,977	(1,143)	-0,925***	(0,134)	-0,003	(0,027)
SIP hoog, 2012	-0,883	(0,849)	0,776***	(0,145)	-0,003	(0,018)
SIP hoog, 2013	-0,827	(1,141)	0,881***	(0,201)	0,001	(0,027)
SIP prestatie, 2012	-0,113	(0,619)	0,335***	(0,110)	-0,014	(0,016)
SIP prestatie, 2013	0,168	(1,057)	1,417***	(0,332)	0,014	(0,024)
gem. controlegr.	18,681	(0,622)	1,394	(0,052)	0,868	(0,010)
R^2	0,920		0,676		0,142	
aantal obs.	7.067		7.067		4.857	
aantal ZABs	226		226		224	

Cluster-robuste standaardfouten tussen haakjes. Het gerapporteerde gemiddelde in de controlegroep is het gemiddelde per week, gemeten over de periode voor de start van het experiment. * significant op significantieniveau van 10%, ** significant op significantieniveau van 5%, *** significant op significantieniveau van 1%.

Om te corrigeren voor algemene tijdstrends is een 2^e-graads polynoom in tijd en een dummy voor de kerstperiode opgenomen. De bijbehorende coëfficiënten zijn niet in de tabel weergegeven.

Tabel G.2: Robuustheidsanalyses: kwartaaldummies.

Panel A: Variatie in het toetsingsmoment.						
	# HiT aanvragen		# HiT toetsingen		fractie conform HiT	
HiT achteraf, 2012	0,270	(0,247)	0,135**	(0,068)	0,020	(0,023)
HiT achteraf, 2013	1,044*	(0,609)	0,376**	(0,158)	0,030	(0,022)
HiT prestatie, 2012	-0,217*	(0,130)	0,010	(0,049)	-0,000	(0,024)
HiT prestatie, 2013	-0,531	(0,338)	-0,014	(0,093)	-0,001	(0,024)
gem. controlegr.	3,381	(0,092)	0,769	(0,025)	0,823	(0,009)
R^2	0,751		0,599		0,093	
aantal obs.	9.286		9.286		5.727	
aantal ZABs	299		299		299	

Panel B: Variatie in het toetsingspercentage.						
	# SIP aanvragen		# SIP toetsingen		fractie conform SIP	
SIP laag, 2012	-0,445	(0,805)	-0,738***	(0,119)	0,013	(0,026)
SIP laag, 2013	-1,558	(1,246)	-0,950***	(0,149)	-0,004	(0,030)
SIP hoog, 2012	-0,881	(0,894)	0,810***	(0,144)	0,011	(0,020)
SIP hoog, 2013	-1,414	(1,205)	0,856***	(0,210)	0,000	(0,029)
SIP prestatie, 2012	-0,111	(0,676)	0,369***	(0,107)	0,001	(0,019)
SIP prestatie, 2013	-0,416	(1,126)	1,393***	(0,349)	0,013	(0,027)
gem. controlegr.	18,681	(0,622)	1,394	(0,052)	0,868	(0,010)
R^2	0,920		0,676		0,143	
aantal obs.	7.067		7.067		4.857	
aantal ZABs	226		226		224	

Cluster-robuste standaardfouten tussen haakjes. Het gerapporteerde gemiddelde in de controlegroep is het gemiddelde per week, gemeten over de periode voor de start van het experiment. * significant op significantieniveau van 10%, ** significant op significantieniveau van 5%, *** significant op significantieniveau van 1%.

Om te corrigeren voor algemene tijdstrends zijn kwartaaldummies opgenomen. Bovendien is een aparte dummy voor de kerstperiode opgenomen. De geschatte coëfficiënten voor deze dummy variabelen zijn niet gerapporteerd in de tabel.