

Onregelmatigheden in declaratiebestanden bij huisartsen, mondzorg, farmacie en GGZ

- Update -

17 februari 2014
Prof. dr. H.M.P. Kersten
Dr. W.J. Kowalczyk

Management samenvatting

De Nederlandse Zorgautoriteit (NZa) heeft aan het Fraude Detectie Expertise Centrum (afgekort FDEC) gevraagd om met behulp van datamining onregelmatigheden op te sporen in declaratiebestanden van huisartsen, mondzorg, farmacie en GGZ. Dit rapport bevat de voornaamste resultaten hiervan. Los van dit rapport zijn aan de NZa uitgebreide lijsten, overzichten en scripts ter beschikking gesteld.

Dit is de eerste keer dat een dergelijk grote hoeveelheid declaraties in de gezondheidszorg in Nederland bij elkaar zijn gebracht om “merkwaardige” transacties op te sporen. Door de toename in rekenkracht en RAM-geheugencapaciteit van computers vormen grote hoeveelheden records geen problemen meer. Voor de analyse zijn gebruikt 881 miljoen records van declaraties van huisartsen, 190 miljoen records van mondzorg, 619 miljoen records van farmacie en 4 miljoen records van GGZ. In totaal betreft het 1,7 miljard records.

In de maanden oktober en november is een uitgebreide data-analyse op deze data gedaan om declaraties op te sporen die “merkwaardig” zijn. Voor deze declaraties kunnen soms goede verklaringen worden gevonden, soms kunnen ze aangemerkt worden als niet conform de regels, soms als onrechtmatig en soms bestaat de indruk dat het om fraude zou kunnen gaan. Voor sommige cases is nader onderzoek noodzakelijk om uitspraak te kunnen doen in welke categorie ze zich bevinden.

Bij de data-analyse is uitgebreid gebruik gemaakt van domein- en expertkennis op het terrein van de zorg in Nederland. Deze kennis, gecombineerd met de vindingrijkheid van de onderzoekers en de rekenkracht van computers, leidt tot overzichten van zgn. *anomalieën*. Een anomalie (Engels: *anomaly*) is een opmerkelijke case in een groot databestand die sterk afwijkt van hetgeen van die variabele(n) (dienstverlener, product, dienst, klant/patiënt/verzekerde, tijd, tijdstip, etc.) verwacht kan worden.

Deze overzichten van anomalieën worden vervolgens door experts beoordeeld. Tabel 1 geeft een overzicht van de omvang van de bestanden per onderzoek segment. De kolom AGB-beh. geeft het aantal AGB-behandelaars aan terwijl bij farmacie dit het aantal AGB-afleveraars betreft. Het aantal records geeft het *totale aantal records* aan, dat wil zeggen inclusief correctierecords en records met de waarde nul als vergoed bedrag.

Jaar	Huisartsen		Mondzorg		Farmacie		GGZ	
	Aantal records	Aantal AGB-beh.	Aantal records	Aantal AGB-beh.	Aantal records	Aantal AGB-afl.	Aantal records	Aantal AGB-beh.
2009							1.165.353	2.758
2010	262.584.340	14.950	69.297.896	15.190	191.744.461	6.790	1.218.992	3.506
2011	304.654.670	17.052	68.449.999	15.550	208.515.505	6.909	1.251.854	4.119
2012	313.926.643	24.532	51.934.447	15.486	219.200.187	6.779		
Totaal	881.165.653		189.682.342		619.460.153		3.636.199	

Tabel 1. Overzicht van de gegevens voor de data-analyse

Opgemerkt moet worden dat er een groot verschil is tussen de gedetecteerde cases: het kan gaan om slordigheid, gemakzucht, onwetendheid en onoplettendheid van alle partijen in de keten van het verlenen van zorg tot het vergoeden ervan, maar ook kan er sprake zijn van fraude.

In de vier segmenten die door ons zijn onderzocht, hebben we overtredingen van declaratieregels en anomalieën gevonden. Anomalieën zijn declaraties van zorgverleners die een sterk afwijkend gedrag vertonen. Om vast te stellen of merkwaardige cases daadwerkelijk frauduleus zijn, is nader onderzoek naar deze cases noodzakelijk. Voor een klein aantal cases is dit (via desk research) gebeurd.

Bij overtredingen is het daadwerkelijk vergoede bedrag vergeleken met het bedrag dat het volgens de declaratieregels zou moeten zijn. Dat verschil wordt in dit rapport het “te veel vergoede bedrag” genoemd. Het is mogelijk dat er wellicht andere redenen zijn waarom een vergoed bedrag afwijkt van het bedrag volgens de regels maar dat is niet het doel van dit onderzoek. Bij anomalieën wordt het bedrag ervan vergeleken met het bedrag dat hoort bij het gehanteerde referentiepunt. Het verschil wordt beschouwd als het “te veel vergoede bedrag”.

Om te bepalen of iets een anomalie is, hanteren we drie referentiepunten: het 95%-punt, het 99%-punt en het 99,9%-punt. Links van het 95%-punt bevindt zich 95% van het totaal vergoede bedrag, rechts ervan 5%, de duurste en hoogste vergoede bedragen. Voor de andere referentiepunten gelden corresponderende percentages. De drie referentiepunten bieden houvast bij het beoordelen van de uitschieters in de verschillende verdelingen. Voor het bepalen van het totale bedrag dat gemoeid is met overtredingen van declaratieregels en de anomalieën, hanteren we conservatief het 99%-referentiepunt. De andere referentiepunten kunnen goed gebruikt worden om scenario's mee door te rekenen (stel dat het aantal merkwaardigheden in een zorgsegment groter is, of juist kleiner, etc., etc.)

De databestanden die voor de analyses zijn gebruikt, zijn niet voor dat doel opgebouwd. Ze zijn verkregen via Vektis die deze bestanden gebruikt voor o.a. (macro-)economisch onderzoek. Wellicht mede daardoor ontbreekt soms belangrijke informatie op declaratieniveau.

Wanneer de zorgverzekeraars aangeboden declaraties weigeren, zijn deze niet zichtbaar in de bestanden: de bestanden bevatten immers uitgekeerde betalingen. Het verschijnsel van controle en correctie achteraf is goed zichtbaar in de databestanden: er zijn debetrecords (toegekende bedragen) en creditrecords (tegenboekingen bij een bijbehorende debetboeking).

De creditboeking is te herkennen aan een negatief bedrag; we noemen dit in dit rapport ook wel een correctieboeking. De redenen waarom een creditboeking plaatsvindt zijn divers: dat kan uiteenlopen van een correctie wanneer bij materiele controle blijkt dat de declaratie ten onrechte is uitbetaald tot fouten in computersystemen van de zorgverzekeraar, van verrekening van niet onder de polis vallende betalingen tot fouten in het systeem van de zorgindieners of zorgbehandelaar, van fouten in de aanlevering van data aan Vektis tot versiebeheer van de data bestanden bij de zorgverzekeraar.

Deze creditrecords zijn in alle vier segmenten zichtbaar; het treedt het meest op bij mondzorg en GGZ.

De creditrecords (correctierecords) worden niet altijd op dezelfde manier behandeld. Zo heeft Vektis voor het segment Farmacie de creditrecords zo veel als mogelijk gekoppeld aan de oorspronkelijke debetrecord(s) en heeft de oorspronkelijke debetrecord(s) vervolgens gecorrigeerd. Daardoor is in dat segment het percentage correctieboekingen zeer laag. Voor de andere segmenten heeft Vektis dit (nog) niet op deze wijze gedaan en zien we dit terug in de data bestanden.

Ook moet beseft worden dat de “wereld van de data” niet noodzakelijkerwijs altijd overeenkomt met de “wereld van de verleende zorg”. Hiermee wordt aangegeven dat hoewel uit de data kan blijken dat zorg geleverd is, dit in de praktijk niet het geval hoeft te zijn. Slechts materiële controle en audits kunnen dit zichtbaar maken.

We hebben op veel plaatsen gezien dat declaratieregels werden overtreden. Nagegaan moet worden of dit oprechte vergissingen zijn, of er opzet in het spel is, of dat de regels wellicht niet helder (genoeg) zijn. Ook de kwaliteit van de brongegevens kan een oorzaak zijn voor het signaleren van merkwaardigheden; hier is voor zover mogelijk rekening mee gehouden.

Wij adviseren dat domeinexperts zich vastbijten in de gevonden onregelmatigheden en merkwaardigheden in de declaratiebestanden. Dit biedt als voordeel dat duidelijk wordt hoe dit heeft kunnen gebeuren, hoe dit kan worden voorkomen, en hoe dit snel kan worden gesignaleerd door de zorgverzekeraar. Ook wordt dan inzichtelijk hoe groot het aantal *false positives* (Nederlands: fout positieven, vals alarm, loos alarm) is in de gedetecteerde cases.

In dit project is duidelijk naar voren gekomen dat de omvang van de bestanden geen rol speelt wat betreft het opsporen van anomalieën. Declaratiemonitoring (het volgen van een grote datastroom) is dus beschikbaar voor de toezichthouder als aanvullend toetsingsinstrument. Toezicht 1.0 (toezien op ordentelijke administratieve organisatie, procedures, integriteit, ethiek, deskundigheid, etc.) wordt dan Toezicht 2.0 waarbij het toezicht zich ook uitstrekt naar de digitale wereld waar het declaratieverkeer steeds meer plaatsvindt. Het biedt tevens de mogelijkheid om – wanneer een zorgverzekeraar of zorgverlener onder verscherpt toezicht wordt geplaatst – met de vinger aan de pols naar de declaratiestromen te kijken.

Op basis van het detecteren van anomalieën en de analyses van de gevonden cases wordt een schatting gegeven van de omvang (frequentie en bedrag) van de onregelmatigheden. Dit is gebaseerd op de data-analyse en de onwaarschijnlijkheidsscore van situaties.

De resultaten in de tabellen 2-5 hebben drie aandachtspunten:

1. Het kan een ondergrens zijn omdat – als er verder wordt gezocht, meer regels worden opgesteld, de staarten van meer meta-variabelen worden beoordeeld – er meer zal worden gevonden. En het is mogelijk dat er onregelmatigheden zijn die met datamining niet kunnen worden gedetecteerd. Bovendien geldt dat er ook links van het referentiepunt niet-correcte declaraties kunnen zijn die (nog) niet zijn vastgesteld. Dit zijn zogenaamde *false negatives* (het is mis maar het alarm gaat niet af). Deze categorie vraagt om nader onderzoek. Hoe ver de ondergrens van het werkelijke, onbekende percentage (en bijbehorend bedrag) ligt, is nog onbekend.
2. Het kan te hoog zijn. Immers, hoewel de overtredingen in declaratieregels hard zijn - in zowel aantal als bedrag - en de merkwaardigheden erg opvallend, kunnen er *false positives* (fout positieven, vals alarm) tussen zitten. Die worden pas gevonden als de betreffende dossiers nauwkeurig worden onderzocht.
3. In het data bestand van huisartsen over 2010 ontbrak circa 10% van de declaraties. De oorzaak was dat één zorgverzekeraar deze gegevens niet heeft aangeleverd aan Vektis. Daarom is voor die ontbrekende 10% in dat jaar bijgeschat.

Jaartal	Totaal bedrag vergoed in miljoenen euro A	Te veel vergoed op basis van overtreding declaratieregels B	Te veel vergoed op basis van anomalieën (99%) C	Totaal te veel vergoed (99%) D = B+C	Totaal te veel vergoed als percentage van totaal vergoed E = D/A
2010	2.521	11,1	5,9	17,0	0,7
2011	2.619	15,4	6,2	21,6	0,8
2012	2.595	11,6	5,4	17,0	0,7

Tabel 2. Resultaten voor huisartsen, 2010 - 2012

Jaartal	Totaal bedrag vergoed in miljoenen euro A	Te veel vergoed op basis van overtreding declaratieregels B	Te veel vergoed op basis van anomalieën (99%) C	Totaal te veel vergoed (99%) D = B+C	Totaal te veel vergoed als percentage van totaal vergoed E = D/A
2010	2.179	0,8	0,9	1,7	0,1
2011	2.180	0,7	1,0	1,7	0,1
2012	2.179	2,5	1,2	3,7	0,2

Tabel 3. Resultaten voor mondzorg, 2010 - 2012

Jaartal	Totaal bedrag vergoed in miljoenen euro A	Te veel vergoed op basis van overtreding declaratieregels B	Te veel vergoed op basis van anomalieën (99%) C	Totaal te veel vergoed (99%) D = B+C	Totaal te veel vergoed als percentage van totaal vergoed E = D/A
2010	5.531	12,4	0,3	12,7	0,2
2011	5.280	10,5	0,9	11,4	0,2
2012	4.669	11,1	2,0	13,1	0,3

Tabel 4. Resultaten voor farmacie, 2010 - 2012

Jaartal	Totaal bedrag vergoed in miljoenen euro A	Te veel vergoed op basis van overtreding declaratieregels in miljoen euro B	Totaal te veel vergoed als percentage van totaal vergoed in miljoen euro B / A
2009	3.990	4,2	0,1
2010	3.980	3,6	0,1
2011	3.980	4,2	0,1

Tabel 5. Resultaten voor GGZ, 2009 – 2011

Deze tabellen geven een goed overzicht van de te veel vergoede bedragen doordat declaratieregels zijn overtreden of dat zorgverleners als anomalie zijn aangemerkt. Voor 2010 en 2011 zijn de cijfers over alle vier zorgsegmenten beschikbaar in de tabellen 2 tot 4.

Ervaring bij fraudeonderzoeken (creditcards) leert dat daar een fraudepercentage van 0,1% als realistisch wordt aangemerkt, *wanneer ook de administratieve en controleprocessen hiervoor goed op orde zijn*. Een dergelijke benchmark ontbreekt bij deze zorgsegmenten. In de situatie waarin de administratieve en controleprocessen goed op orde zijn is het referentiepunt van 99,9 uitermate geschikt om achter merkwaardige cases aan te jagen. Anders ligt het wanneer men algemene uitspraken wil doen over het optreden van merkwaardigheden, onregelmatigheden en afwijkende cases. Daarom menen wij - mede op basis van onze ervaringen met andere onderzoeken naar merkwaardigheden/anomalieën - dat het uitgangspunt om het 99%-referentiepunt hier het beste is. Ook het referentiepunt van 95% biedt houvast voor het inschatten van de omvang omdat er ook links van het referentiepunt declaraties zullen zijn die niet door de thans aanwezige detectieregels worden gevonden.

Voor 2010 kunnen we concluderen dat het overtreden van declaratieregels leidt tot een te veel vergoed bedrag van 27,9 miljoen. Hier komt nog een bijdrage van de gevonden anomalieën bij van 7,1 miljoen hetgeen het totaal op een bedrag brengt van 35 miljoen euro in 2010.

Voor 2011 is het totale bedrag dat te veel vergoed is 38,9 miljoen euro (aan overtreding van de declaratieregels 30,8 miljoen plus 8,1 miljoen door de bijdrage van anomalieën).

Wanneer men niet uitgaat van het 99%-punt maar van 99,9% (wat betekent dat men zeer conservatief rekent), dan zijn de schattingen voor 2010 28,5 miljoen euro en voor 2011 32,5 miljoen euro. Anderzijds, hanteert men het referentiepunt van 95% waardoor men iets minder conservatief rekent, dan zijn de te veel vergoede bedragen voor 2010 en 2011 respectievelijk 62,6 miljoen euro en 67,1 miljoen euro.

Inhoud

Management samenvatting	2
Voorwoord	10
1. Inleiding	11
1.1. Introductie	11
1.2. Wat is fraude?	11
1.3. Probleemstelling	12
1.4. Schatting van het bedrag gemoeid met het overtreden van regels en anomalieën	15
2. Data en data-analyse	19
2.1. Data verkregen via Vektis	19
2.2. Gebruikte analysetechnieken en zoekmethoden	23
3. Resultaten van de data-analyses	25
3.1 Inleiding	25
3.2 Resultaten voor huisartsen	26
3.2.1 Overzicht van variabelen	26
3.2.2 Overzicht van frequentie van declaraties van huisartsen	28
3.2.3 Overzicht van het aantal prestaties per declaratie van huisartsen	29
3.2.4 Overzicht van declaratiebedragen	31
3.2.5 Overschrijdingen van het inschrijftarief	31
3.2.6 Een combinatie van M&I verrichting en een consult op dezelfde dag	33
3.2.7 Meerdere consulten van een verzekerde op één dag	34
3.2.8 Merkwaardige individuele cases	34
3.2.9 Schatting van het bedrag gemoeid met het overtreden van regels en anomalieën	37
3.3 Resultaten voor mondzorg	40
3.3.1 Overzicht van de variabelen	40
3.3.2 Overzicht van frequentie van declaraties bij mondzorg	42
3.3.3 Overzicht van het aantal prestaties per declaratie van tandartsen	43
3.3.4 Overzicht van declaratiebedragen	45
3.3.5 Ongeldige combinatie van declaraties	45
3.3.6 Dubbele declaraties	45
3.3.7 Overzicht van enkele anomalieën	46
3.3.8 Schatting van het bedrag gemoeid met het overtreden van regels en anomalieën	56
3.3.9 Overzicht mondzorg	59
3.3.10 Specifieke analyse voor de basisverzekering	60

3.3.11	Specifieke analyse voor de jeugd in de basisverzekering.....	63
3.4	Resultaten voor farmacie	67
3.4.1	De beschikbare data	67
3.4.2	Overzicht van frequentie van declaraties bij farmacie.....	70
3.4.3	Overzicht van declaratiebedragen	72
3.4.4	Extreem hoge gemiddeld bedrag per declaratie.....	72
3.4.5	Verzekerden die langs veel apotheekpraktijken gaan	75
3.4.6	Verslavende medicijnen	76
3.4.7	Overtarifiering door toeslag eerste ter handstelling.....	77
3.4.8	Overtarifiering weekterhandstelling	78
3.4.9	Dubbele declaraties	78
3.4.10	Overschrijdingen WMG tariefcode.....	79
3.4.11	Overzicht van enkele anomalieën	80
3.4.12	Schatting van het bedrag gemoeid met het overtreden van regels en anomalieën	81
3.4.13	Overzicht farmacie	84
3.4	Resultaten voor GGZ	85
3.5.1	Overzicht van de variabelen	85
3.5.2	Overzicht van frequentie van declaraties in de GGZ.....	87
3.5.3	Dubbele facturen.....	89
3.5.4	Meerdere DBC's tegelijkertijd geopend voor een patiënt	90
3.5.5	Perioden met overlappende verblijf-DBC's.....	91
3.5.6	Schatting van het bedrag gemoeid met het overtreden van regels en anomalieën	93
3.5.7	Overzicht GGZ.....	93
4.	Overzicht van de resultaten	94
4.1	Inleiding	94
4.2	Schatting van het bedrag gemoeid met het overtreden van regels en anomalieën	95
5.	Conclusies, observaties en aanbevelingen	96
5.1	Conclusies.....	96
5.2	Observaties.....	98
5.3	Aanbevelingen.....	99

Appendix 1. Toelichting op methoden	100
Appendix 2. Toetsen voor huisartsen.....	111
Appendix 3. Toetsen voor mondzorg	113
Appendix 4. Toetsen voor farmacie	122
Appendix 5. Toetsen voor GGZ.....	124
Appendix 6. Toelichting op de data.....	126
6.1 Algemeen.....	126
6.2 Huisartsen.....	127
6.3 Mondzorg	129
6.4 Farmacie	132
6.5 GGZ	134
Appendix 7 Resultaten op basis van Cortana, huisartsen 2010	135

Voorwoord

Het opsporen van merkwaaardigheden in de miljoenen declaraties van huisartsen, mondzorg, farmacie en GGZ is een geweldige uitdaging. Wat is een opmerkelijke case, wat is een verklaarbare uitschieter en wanneer is er sprake van toch echt wel een heel vreemde declaratie?

Ons onderzoek in deze hooiberg van 1,7 miljard declaraties is onder tijdsdruk uitgevoerd en heeft geleid tot een eerste rapportage in december 2013. Op dit eerste rapport is veel positieve feedback gekomen en dit heeft tot verbeteringen geleid.

Ook is in de feedback erop gewezen dat niet alle correctieboekingen in het onderzoek naar declaratiebestanden waren verwerkt. Het betreft hier creditrecords: dit zijn records met tegenboekingen. Met deze boekingen worden debetboekingen verrekend. Door het niet verwerken van alle correctieboekingen zijn de oude berekeningen in het te veel vergoede bedrag, met name in mondzorg en deels in GGZ, te hoog. In dit rapport is met deze creditboekingen rekening gehouden.

Dit rapport is tot stand gekomen door de ondersteuning van de teamleden van het Fraude Detectie Expertise Centrum (FDEC) en het Leiden Institute of Advanced Computer Science (LIACS). Het gaat hier om Benjamin van der Burgh, Floris Kleyn, Marvin Meeng en Arno Knobbe.

Verder zijn wij blij met de samenwerking met Vektis voor de aanlevering van de data en de beantwoording van de vragen over zowel de data als de opmerkelijke cases die aan hen zijn voorgelegd. Ook het opbouwende (en scherpe en alerte) commentaar op de verschillende conceptversies wordt erg gewaardeerd.

We kijken terug op een goede en plezierige samenwerking met leidinggevenden en medewerkers/sters van de Nederlandse Zorgautoriteit. Hun inhoudelijke kennis heeft veel bijgedragen aan onze bevindingen, daarvoor dank. Echter, de verantwoordelijkheid over de inhoud van dit rapport berust geheel bij het Fraude Detectie Expertise Centrum.

1. Inleiding

1.1. Introductie

Door de Nederlandse Zorgautoriteit (afgekort NZa) is aan het Fraude Detectie Expertise Centrum (afgekort FDEC) gevraagd om een analyse uit te voeren op een groot data bestand betreffende vier zorgsegmenten: huisartsen, mondzorg, farmacie en GGZ, voor de periode 2010 tot en met 2012. FDEC is gespecialiseerd in de analyse van zeer grote data bestanden oplopend tot miljarden transacties. Doelstelling van deze analyses is het opsporen van afwijkingen in gebruik, bedragen, frequentie, etc., etc. Daartoe worden (combinaties van) wiskundige en statistische technieken ingezet.

Uit de opdrachtschrijving blijkt duidelijk de kaders en het doel van dit project naar onregelmatigheden in de declaratiebestanden:

“Voor het verzoek om onderzoek te doen naar de omvang van fraude wil de NZa allereerst d.m.v. datamining een analyse uitvoeren op de declaratiebestanden teneinde onregelmatigheden daarin boven tafel te krijgen. Via deze technieken kunnen opvallende afwijkingen in declaratiebestanden die mogelijk op niet-correcte declaraties wijzen, gedetecteerd worden. Deze opvallende onregelmatigheden zijn niet op voorhand aan te merken als niet-correct en zeker ook niet als fraude. In voorkomende gevallen kan er ook een goede verklaring zijn, reden waarom verdiepend onderzoek noodzakelijk zal zijn.”

Onregelmatigheden en afwijkingen van wat – op grond van feiten en achtergrondvariabelen - kan worden verwacht, bieden aanknopingspunten voor de beantwoording van de vraag “gaat het hier om fraude?”. In deze rapportage gaan we **niet** in op de schatting van de omvang van de fraude in de zorg in Nederland. Wél geven we een overzicht van overtredingen, onregelmatigheden en opmerkelijke afwijkingen die zijn gevonden en de bedragen die daarmee zijn gemoeid.

1.2. Wat is fraude?

Onze ervaring met fraudedetectie in veel verschillende domeinen is dat steeds weer de cruciale vraag optreedt:

Gaat het bij deze gevonden afwijking om toeval, om een vergissing, om een verkeerd begrip of interpretatie, om opzettelijke fraude of om systematische fraude en/of samenspanning?

Daarom is het doen van uitspraken over de totale fraude in (segmenten van) de zorg gecompliceerd. Onze ervaring is dat zelfs binnen één organisatie de opvatting over wat als fraude moet worden aangemerkt sterk kan verschillen. De grondhouding van medewerkers en leiding verschilt immers: sommigen gaan uit van het goede van de mens, anderen denken dat mensen geneigd zijn tot het kwade. Onze ervaring bij eerdere fraudedetectie bij een groot aantal organisaties leert dat – mede wellicht daarom – niet alle gevonden afwijkingen en merkwaardigheden leiden tot aangifte en een rol voor het Openbaar Ministerie. Soms ook is het lastig om volledig waterdicht te bewijzen dat het om fraude gaat.

Omdat op voorhand niet gelijk duidelijk is waarmee we te maken hebben, hanteren we het begrip anomalie:

Een anomalie is een opmerkelijke case in een zeer groot data bestand die sterk afwijkt van hetgeen van die variabele(n) (dienstverlener, product, dienst, klant/patiënt/verzekerde, tijd, tijdstip, etc.) verwacht kan worden.

Wanneer een anomalie is gedetecteerd, wordt deze case nader onderzocht. Soms vindt men een verklaring voor de anomalie maar vaak ook niet. Wanneer blijkt dat het echt “mis” is, vindt er door de uitkerende instantie een gesprek plaats met de anomalie-veroorzaker om een verklaring te vinden. Mocht die verklaring er niet zijn, of niet aannemelijk zijn, dan vindt terugvordering plaats.

Om het aantal *false positives* (gedetecteerd maar er is toch niets aan de hand) zo klein mogelijk te houden is *feedback* op de gevonden anomalieën essentieel: door domeinexperts worden de gevonden anomalieën beoordeeld. Hun *feedback* leidt tot scherpere en nog doeltreffender detectie regels.

1.3. Probleemstelling

De vraag is om zinvolle anomalieën in de declaratiebestanden van vier zorgsegmenten te vinden die inzicht bieden in onregelmatigheden in de zorg in Nederland. Het signaleren van een anomalie betekent niet als vanzelf dat het hier gaat om fraude. Grofweg zijn er drie mogelijkheden, in volgorde van toenemende ernst:

1. Een onregelmatigheid
2. Een onrechtmatigheid
3. Een geval van fraude

Hoewel het zoeken van een speld in een hooiberg nu, met krachtige computers en slimme methoden goed mogelijk is, wordt een succesvolle *anomalie detectie* bepaald door een goede samenwerking tussen domein/fraude-deskundigen en de data analisten. Het werk leunt zwaar op de kennis en ervaring van de fraude-experts én op de gedegen expertise van de domeindeskundigen: van profielen van fraudeurs en hun werkwijze tot kennis omtrent de regels, de voorwaarden en de uitzonderingen. De experts leveren zowel input voor het zoeken in de databases als *feedback* op de gevonden resultaten. Op die manier leidt dit iteratieve proces tot een effectieve en efficiënte manier van detectie.

Of een anomalie ook daadwerkelijk fraude is, is lastig vast te stellen. Uiteindelijk gaat de rechter hierover. Wél is het goed mogelijk om vast te stellen of er sprake is van onrechtmatigheid, regelmaat, vergissing of verrijking. Anomalie detectie *sec* is niet voldoende maar wel een eerste stap. De volgende stap is het beoordelen van de gevonden anomalieën. Deze vallen in een aantal klassen uiteen:

- a. Onrechtmatigheden door dat vergoedingsregels zijn overtreden
- b. Opvallende gelijkenissen met bekende fraudeurs in het segment
- c. Afwijkende patronen ten aanzien van vergoedingen
- d. Afwijkende patronen van/bij de zorgaanbieder, rekening houdend met zijn/haar specifieke kenmerken omtrent patiënten, mede in relatie tot zijn/haar vakgenoten
- e. Afwijkingen in hun groep patiënten

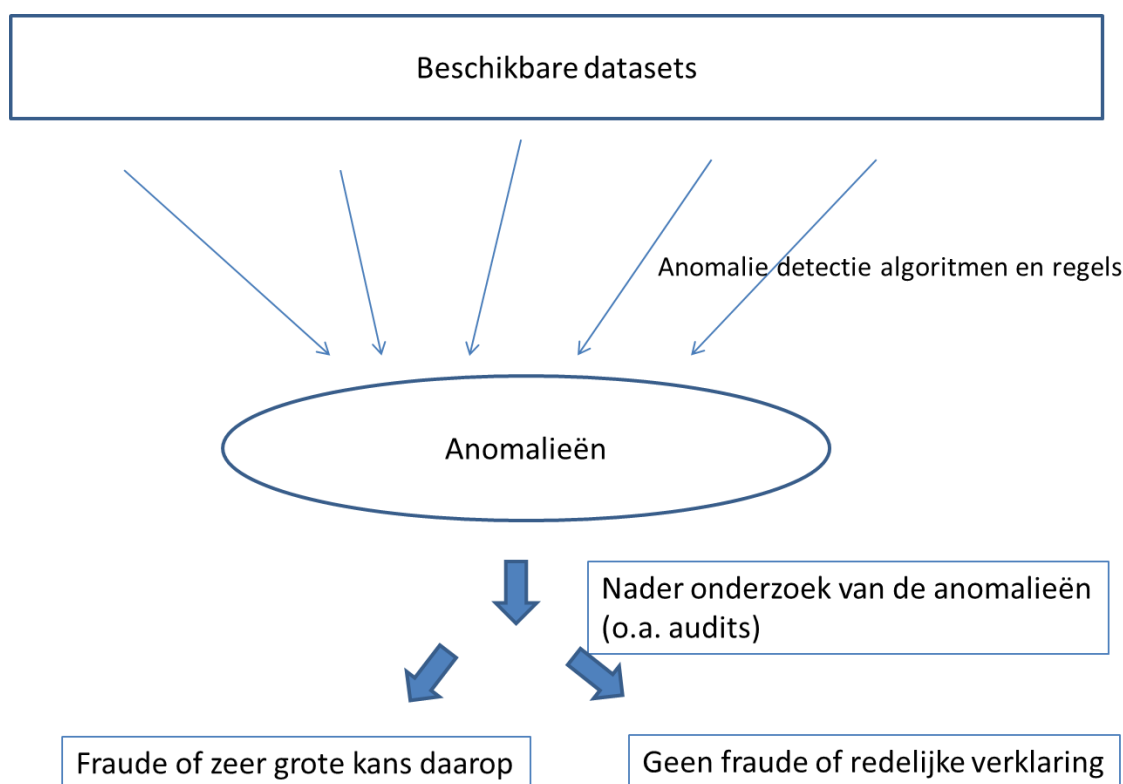
De cases van categorie a. zijn relatief eenvoudig: aanwezige regels zijn overtreden en het bedrag dat hiermee gemoeid is, kan worden berekend. We kunnen niet vaststellen of zorgaanbieders *in de geest van de regels* hebben gedeclareerd en evenmin of hun overtredingen zich ook uitstrekken naar andere declaraties.

De cases in de groepen b – e moeten nader worden beoordeeld om te bepalen wat hiermee aan de hand is. Het onderscheid tussen een toevallige (of verklaarbare) afwijking en bewezen opzet is groot. De wiskundige waarschijnlijkheid dat iemand op een fraudeur lijkt, heeft geen rechtsgrond en kan ook toeval zijn. Maar toch: het kan een begin zijn om eens nader op deze cases in te gaan.

Hier komen de fraude-experts en domeindeskundigen weer in beeld. Zij bestuderen de gevonden anomalieën en de toelichting waarom/ de wijze waarop de anomalieën door de data analisten zijn gevonden. Zij kunnen dan beoordelen of de case een false positive (vals alarm: gedetecteerd maar er is niets aan de hand) is. *Feedback* aan de data analisten vindt plaats en de detectie regels worden bijgesteld.

De deskundigen kunnen ook tot het oordeel komen dat de cases toch wel erg merkwaardig zijn en besluiten tot een audit. Dat is in dit onderzoek niet gebeurd: een nader onderzoek naar de patiënt en zorgverlener heeft slechts plaatsgevonden op basis van de beschikbare declaratiebestanden. De audit moet uitsluitsel geven op de vraag of hier sprake is van fraude¹. Pas ná de resultaten van de audit kan vastgesteld worden hoeveel van de gegenereerde cases fraude betreft, en via dat cijfer kan een schatting voor Nederland worden bepaald. Natuurlijk is het mogelijk om het auditprogramma efficiënt op te zetten door steekproeftechnieken te hanteren.

Figuur 1 geeft het stappenplan van dit onderzoeksproces schematisch weer.



Figuur 1. Anomalie detectie en de stappen erna

¹ Hier wordt niet ingegaan op de opzet van de audit.

In dit tekstblok wordt een voorbeeld van de *feedback* van een expert gegeven voor één *outlier* (uitschieter of uitbijter) op het aantal verstrekkingen beschreven. Het betreft hier een verzekerde waarvan we bij farmacie in 2010 in totaal 3.890 declaraties aantreffen en in 2011 in totaal 3.182 declaraties. Dat zijn uitzonderlijk hoge aantallen die al bij de exploratieve analyse werden gevonden (zie verder paragraaf 3.4.2.). Deze *feedback* van de expert (die bijvoorbeeld kijkt om welke medicijnen het gaat, de dosering, de frequentie, etc., etc.) leidt zowel tot verdere bijstelling van de detectie regels als tot opvolging van de case.

“Opmerkelijke casus en voor een apotheek ook zeer discutabel; er is mogelijk sprake van:
- *misbruik van iemands BSN door op deze manier verslaafden van een portie te voorzien;*
- *declaratiefraude door op 1 persoon veel te veel geneesmiddelen te boeken;*
- *een handel van de betreffende patiënt en apotheker in de geneesmiddelen*
zoals in de bijlage vernoemd.

We hanteren dus een tweesporen aanpak: het vaststellen of declaratieregels overtreden zijn én het opsporen van merkwaardigheden/anomalieën waarbij dat laatste in een iteratief proces gebeurt:

- a. Het signaleren van een anomalie
- b. Het voorleggen van deze anomalie aan een expert op het betreffende zorgdomein
- c. De uitslag van de expert: is hier inderdaad sprake van een onregelmatigheid, een onrechtmatigheid of wellicht de verdenking van fraude.²

Het is mogelijk om experts te vragen om een a priori kansschatting te geven of een anomalie fraude betreft of niet. Daarvan is in dit project afgezien omdat de voornaamste focus is het opsporen van onregelmatigheden in de grote data bestanden. De inschatting van deze kans door experts kan de efficiëntie van de opvolging van de onregelmatigheden sterk verhogen.

Wanneer blijkt dat declaratieregels zijn overtreden, wordt voor die cases gekeken naar het verschil tussen hetgeen daadwerkelijk vergoed is en hetgeen volgens de declaratieregels vergoed had mogen worden. Dit verschil, waarvoor overigens ook andere moverende redenen mogelijk zouden kunnen zijn, wordt het *te veel vergoede bedrag* genoemd. Inzicht te verkrijgen in de mogelijke redenen die het verschil zouden kunnen verklaren, is niet het doel van dit onderzoek. Wij beperken ons tot een analyse op basis van de declaratiebestanden waarin wordt weergegeven wat er concreet in de praktijk is vergoed voor zorg. Het totaal te veel vergoede bedrag is de som over alle gevonden cases.

² We vragen **niet** aan de expert om een schatting te geven voor de kans dat de case frauduleus is. Een dergelijke aanpak is mogelijk en kan leiden tot een schatting van het fraudepercentage in Nederland. Zie bijvoorbeeld Ger van Gils, Peter G.M. der Heijden en Olav Laudy (2003), *Fraude in particuliere inboedelverzekeringen*, Utrecht: BOA.

1.4. Schatting van het bedrag gemoeid met het overtreden van regels en anomalieën

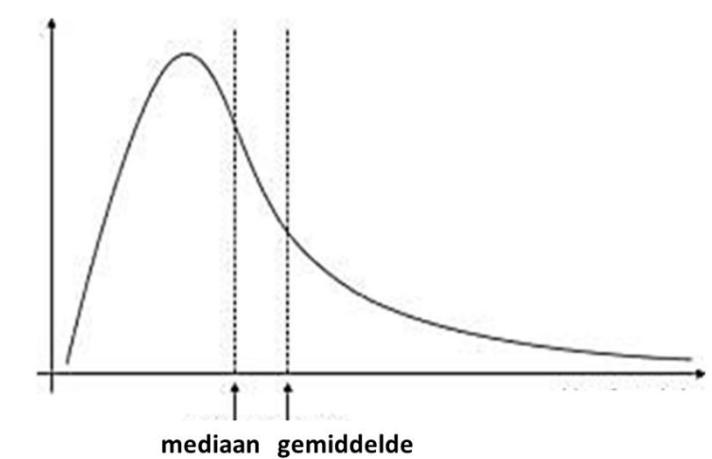
In dit project zijn de declaraties van de zorg geleverd door huisartsen, mondzorg, farmacie en GGZ onderzocht. Daarbij is zowel gekeken naar het overtreden van regels als naar merkwaardigheden en afwijkende patronen. Zo zijn bijvoorbeeld de zorgverleners vergeleken met hun collega's, en is er gekeken naar de verdeling van de declaraties en verrichtingen.

Op basis daarvan kan een schatting worden gemaakt van de omvang (frequentie en bedrag) van het overtreden van declaratieregels en de omvang (frequentie en bedrag) dat gemoeid is met de meest opvallende merkwaardigheden/afwijkingen. Of een merkwaardigheid echt fraude is, wordt uiteindelijk bepaald door de rechter of door de deskundigen van de zorgverzekeraars in kwestie.

De werkwijze is in grote lijnen als volgt:

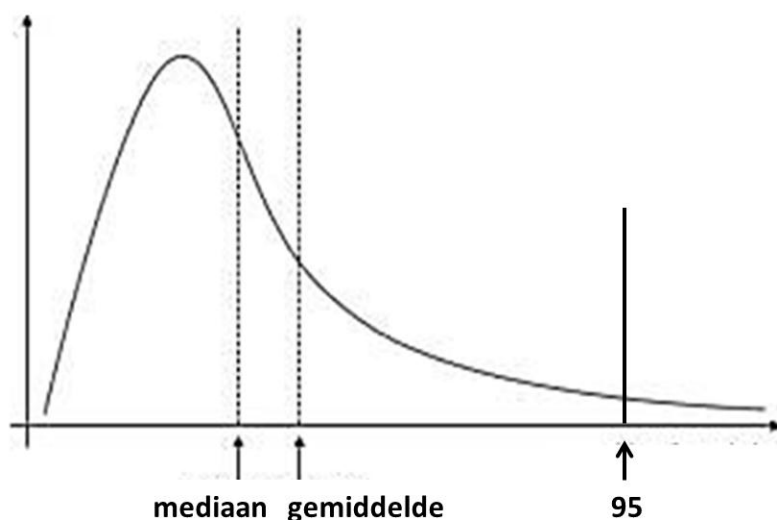
1. Als er sprake is van een overtreding van een regel dan berekenen we het verschil tussen wat gedeclareerd is en wat volgens de regel gedeclareerd had mogen worden.
2. Als er sprake is van een anomalie dan is de mate van afwijkendheid leidend in de berekening van het te veel vergoede bedrag. We kijken naar de verdeling van de gedeclareerde bedragen per patiënt/verrichting, etc. Dit zijn doorgaans scheve verdelingen. De aandacht wordt respectievelijk gericht op de top 5%, top 1% en top 0,1% van de cases die bijdragen aan het totale bedrag dat vergoed is. Deze drie groepen van cases worden beschouwd als afwijkend. Voor deze groepen wordt het te veel vergoede bedrag berekend door hun vergoeding te verminderen met het bedrag dat bij het 95-, 99- en 99,9- percentiel behoort.
3. De totalen van 1. en 2. worden bij elkaar opgeteld.

We illustreren onderdeel 2 met een aantal figuren en een rekenvoorbeeld. In figuur 2 is een voorbeeld gegeven van een verdeling waarbij ook het gemiddelde en de mediaan zijn weergegeven. De mediaan geeft de middelste waarneming aan wanneer de waarnemingen op grootte zijn gerangschikt. De verdeling is – zoals in veel situaties bij anomalie detectie – scheef en we zijn in het bijzonder geïnteresseerd in de cases die zich in de rechterstaart van de verdeling bevinden.

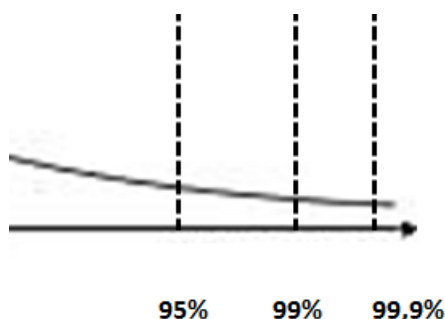


Figuur 2. Voorbeeld van een verdeling met daarin de mediaan en het rekenkundig gemiddelde

In figuur 3 is ook het 95%-punt aangegeven. 95% van het totaal vergoede bedrag bevindt zich links van dit punt, 5% bevindt zich rechts van dit punt³. Leidend is het vergoede bedrag, niet het aantal personen/zorgaanbieders, etc. We kunnen de aandacht verder richten op de cases die zich nog verder in de rechterstaart van de verdeling bevinden. We kijken dan bijvoorbeeld naar de top 1% of, nog verder naar rechts, de top 0,1%.



Figuur 3. Voorbeeld van een verdeling met het 95-percentiel



Figuur 4. Voorbeeld van een verdeling met het 95, 99- en 99,9-percentiel

De referentiepunten 99,9%, 99% en 95% zijn enigszins arbitrair. Doorgaans wordt de grens bepaald door wat de opsporingsafdeling/domeinexperts aan kunnen (qua aantal) én door het aantal *false positives* (fout positieven, vals alarm). Naarmate dit referentiepunt lager is (bijvoorbeeld van 99% naar 95%), neemt de kans om tussen de gedetecteerde cases onschuldigen aan te treffen toe. Anderzijds blijven er bij een hoger referentiepunt mogelijk niet-correcte declaraties onterecht buiten beeld. In dit project worden deze drie grenzen gehanteerd om inzicht te krijgen in de gevoeligheid van de uitkomsten bij verschillende afkapgrenzen.

³ Omdat dit doorgaans scheve verdelingen zijn, is dit niet noodzakelijkerwijs de twee sigma-grens.

Een rekenvoorbeeld kan dit verder illustreren.

Stel dat er gekeken wordt naar het vergoede bedrag voor een specifieke categorie van verslavende middelen, relatief ten opzichte van het gemiddelde over alle patiënten (dit is toets 22, zie Appendix 4). We zien in de farmaciedata over 2012 dat het totale bedrag dat voor deze categorie is vergoed € 18.417.222 is. Het gemiddelde bedrag is € 48,53 berekend over alle verzekerden. De mediaan is € 19,56.

Het bedrag dat overeenkomt met het 95%-punt is € 1.335. Links van dit punt bevinden zich 379.126 patiënten, goed voor een bedrag van € 17.495.798. Rechts van dit 95%-punt bevinden zich 362 patiënten die samen goed zijn voor € 921.424.

Het bedrag dat overeenkomt met het 99%-punt is € 5.979. Links van dit punt bevinden zich 379.471 patiënten, goed voor een bedrag van € 18.228.345. Rechts van dit 99%-punt bevinden zich 17 patiënten die samen goed zijn voor € 188.877.

Tenslotte, het bedrag dat overeenkomt met het 99,9%-punt is € 23.282. Links van dit punt bevinden zich 379.487 patiënten, goed voor een bedrag van € 18.390.213. Rechts van dit 99,9%-punt bevindt zich 1 patiënt, goed voor € 27.009.

Het “te veel vergoede bedrag” wordt als volgt berekend: totaal vergoede bedrag rechts van het referentiepunt minus het bedrag dat hoort bij het referentiepunt maal het aantal patiënten rechts van het referentiepunt.

Voor het 95%-punt betekent dit: € 921.424 minus (362 maal € 1.335) = € 438.154.

Voor het 99%-punt is dit € 188.877 minus (17 maal € 5.979) = € 87.234.

Voor het 99,9%-punt is dit € 27.009 minus (1 maal € 23.282) = € 3.727.

Dit leidt tot de volgende cijfers, zie tabel 1.

Farmacie	Het percentage onregelmatigheden (als percentage van totaal aantal cases)				Schatting van het hiermee gemoeide bedrag, in miljoen euro			
	Overtreding regels	95% niveau	99% niveau	99,9% niveau	Overtreding regels	95% niveau	99% niveau	99,9% niveau
Anomalieën opiaten		0,1 (=362/379.488)	0,0	0,0		0,438	0,087	0,004

Tabel 1. Voorbeeld van de anomalie-berekeningen voor opiaten, 95%, 99% en 99,9%-referenties

Duidelijk is te zien dat wanneer men de grens verlaagt, het bedrag dat gemoeid is met cases in de rechterstaart oploopt. Tegelijkertijd neemt het aantal onschuldigen (*false positives*, fout positieven, vals alarm) toe. Het lastige is dat het percentage fraude onbekend is en we dus geen houvast hebben bij de keuze van het referentiepunt. Ervaring bij fraudeonderzoeken (creditcards) leert dat daar een fraudepercentage van 0,1% als realistisch wordt aangemerkt (*wanneer ook de administratieve en controleprocessen hiervoor op orde zijn*). Een dergelijke benchmark ontbreekt bij deze zorgsegmenten.

Deze methode kan ook worden toegepast om bijvoorbeeld anomalieën te vinden voor de kosten van vullingen, extracties, röntgenfoto's, etc. Zo geldt voor vullingen (in 2010 en 2011) dat er vijf prestatiecodes zijn, elk met een bepaalde prijs. Voor elke AGB berekenen we dan de gemiddelde kosten van een vulling voor alle patiënten die een vulling hebben gekregen. Vervolgens wordt de rechterstaart van deze verdeling onderzocht: hierin bevinden zich de relatief erg dure tandartsen.

In Kersten en Kowalczyk (zie Jaarboek Compliance 2009) is beschreven hoe de nadruk op snelle verwerking van declaraties zijn weerslag heeft op het aspect rechtmatigheid. Het snel opsporen van onregelmatigheden (“dicht op de bal spelen”) is ook vanuit efficiëntie overwegingen belangrijk: daarmee voorkomt men de groei van de onregelmatigheden, de te veel vergoede bedragen die worden uitgekeerd en het later terugvorderen van veel geld⁴. Elders in dit rapport beschrijven we een verzekerde die bij veel verschillende apotheken medicijnen krijgt (20 apotheken). Het jaar daarna is het aantal apotheken waar hij geneesmiddelen heeft gekregen gestegen tot 87. Voorlopig kunnen we in dit onderzoek slechts spreken over bedragen die gemoeid zijn met anomalieën in verdelingen en databestanden.

Bij het sommeren van het totaal van 1 (het bedrag dat te veel vergoed is doordat declaratieregels zijn overtreden) en het totaal van 2 (anomalie detectie) is geen rekening gehouden met of de veroorzaker van de anomalie gesignaleerd is in beide groepen. Dat is niet waarschijnlijk omdat de benaderingen echt verschillend zijn. Mocht dit wel (in theorie is het mogelijk) het geval zijn, dan is correctie hiervoor pas zinvol wanneer de lijst met merkwaardigheden (anomalieën) geanalyseerd is door domeinexperts en bekend is of er sprake is van “er is echt iets goed mis” dan wel *false positives* (“afwijkend maar er is niets aan de hand”).

⁴ Overigens is het *near real time* monitoren met de huidige IT-hulpmiddelen goed mogelijk. Er is een voorbeeld van een analyse-afdeling van een zorgverzekeraar die – om de grootschalige management informatiesystemen en besognes van de IT-afdeling niet te belemmeren – elke dag een kopie van het declaratiebestand krijgt en dat onderzoekt om anomalieën. Het is een mooi voorbeeld van “dicht op de bal spelen”.

2. Data en data-analyse

2.1. Data verkregen via Vektis

Door de NZa zijn gegevens opgevraagd bij Vektis over de jaren 2010 t/m 2012, uitgezonderd GGZ. Voor GGZ zijn de gegevens 2009 t/m 2011 opgevraagd omdat er onvoldoende gegevens over 2012 beschikbaar waren. Het ging hierbij om zeer grote aantallen records. In Appendix 6 wordt uitgebreid ingegaan op de beschikbare variabelen en de kwaliteit ervan. Hier kan worden vermeld dat de data geschikt is om anomalie detectie op uit te voeren maar dat er zorgen zijn over bepaalde variabelen én de aard en omvang van records met negatieve bedragen. Uit de analyses komen ook aanbevelingen om de kwaliteit van de data voor detectiedoeleinden in de toekomst te verbeteren⁵. De samenwerking met Vektis wordt als opbouwend en nuttig ervaren.

De door Vektis aangeleverde gegevens dienen/dienden vooral om statistische informatie en algemene trends op te baseren. Omdat het doel van de dataverzameling dus anders is dan wanneer dit voor detectiedoeleinden zou zijn, staan er in het databestand gegevens die minder geschikt zijn om merkwaardige cases op te sporen; anderzijds ontbreken er gegevens die juist heel effectief kunnen zijn in de detectie.

De gegevens die als basis dienen voor de analyse hebben in het proces van dataverzameling en verwerking een aantal bewerkingen ondergaan. Zo zijn er data verwerkt die wel aangepast die bijvoorbeeld incompleet, foutief of niet-consistent waren. Inzicht in het effect hiervan op de uiteindelijke resultaten van onze analyses ontbreekt. De databestanden van Vektis bevatten goedgekeurde en uitbetaalde declaraties. Dit betekent dat de gegevens niet noodzakelijkerwijs hetzelfde zijn als bij zorgverzekeraars zelf. Of deze verschillen groot zijn, zal blijken bij de ultieme toets bij de zorgverzekeraar of een merkwaardige case, een anomalie is dan wel een fraudegeval.

De privacy van verzekerden is gewaarborgd doordat encryptie van het Burger Service Nummer (BSN) heeft plaatsgevonden; daardoor is er een nieuw uniek nummer ontstaan waarbij we niet de persoon kunnen traceren maar wél de verrichtingen en vergoedingen aan de verzekerde kunnen relateren. Ook de geboortedatum van de verzekerde is verhuuld doordat we alleen de leeftijd op 1 juli van dat jaar kennen. In het data bestand farmacie beschikken we over de postcode (op 4-cijferig niveau), bij de andere zorgsegmenten (huisartsen, mondzorg en GGZ) is de postcode niet aanwezig.

De gegevens van de behandelaar en de indiener zijn bekend. In het project zijn afdoende maatregelen genomen om die privacy te waarborgen door strenge toegangsbeveiliging op de data, *policies* wat betreft de data en dataverkeer, geanonimiseerde eindrapportage, beveiligde opslag van documenten en een zorgvuldige selectie van de leden van het team. Alle gegevens, zowel brongegevens als afgeleide data bestanden worden aan het eind van de opdracht aan NZa overgedragen. Bij FDEC vindt daarna vernietiging van de gegevens plaats.

Voor het vaststellen van “merkwaardige” cases⁶ onderscheiden we drie benaderingen met bijbehorende kernvariabele. De variabelen waarvan dan uitgegaan wordt, zijn:

- Verzekerde
- Aanbieder = indiener
- Behandelaar

⁵ Zie Appendix 6.

⁶ We hanteren hier doorgaans de term “anomalie” voor.

Daarnaast zijn er variabelen die veel informatie bieden zoals datum van de prestatie, bedrag, leeftijd, etc., evenals afgeleide variabelen (meta-variabelen) zoals het aantal consulten per verzekerde, het aantal verrichtingen per behandelaar en de verhouding tussen verschillende verrichtingen per behandelaar. Een overzicht van de variabelen wordt in Appendix 6 gegeven.

Bij de anomalie detectie worden de analyses zowel voor de behandelaar als voor de indiener uitgevoerd. Het is immers mogelijk dat een zorgverlener via zijn/haar eigen AGB-code declareert maar het is ook mogelijk dat het indienen op praktijkniveau of via een factoringmaatschappij plaatsvindt. Merkwaardigheden kunnen op verschillende plaatsen in het declaratieproces ontstaan. Het is ook mogelijk dat dit bij/via verzekerden plaatsvindt; dit wordt apart onderzocht. Doorgaans gaat het dan om “papieren” declaraties (die natuurlijk ook door de verzekerde digitaal kunnen worden ingediend bij de verzekeraar); de code van de indiener is dan 0.

Een bijzonder vervelend probleem is het optreden van ontbrekende waarnemingen zoals lege velden in de bestanden en het veelvuldig voorkomen van codes die aangeven dat iets onbekend is of ontbreekt. Het gaat daarbij vaak om de code “0”, “1234567” en “9999999” (of vergelijkbare codes). Vooral als dat bij *key*-variabelen optreedt zoals prestatiecode en AGB-behandelaar, maar ook bij BSN, verstoort dit de detectie. Immers, het is mogelijk dat juist *omdat* de variabele ontbreekt er een grotere kans op fraude is maar bovendien kan dan niet beoordeeld worden of de vergoeding juist is (prestatiecode onbekend) en aan wie die wordt uitbetaald.

Daarnaast speelt het verschijnsel van creditrecords of tegenboekingen (correctierecords) een grote rol: dit zijn records met negatieve bedragen. In dit rapport noemen we deze creditrecords ook wel correctieboekingen. Soms gaat het om miljoenen creditrecords en de verdeling over de vier zorgsegmenten én jaren verschilt sterk. Voor 2012 bestaat de indruk dat nog niet alle correcties zijn verwerkt door de zorgverzekeraar/Vektis.

Bij de creditrecords is het opvallend dat bij huisartsen vooral in het eerste kwartaal van een jaar er veel van worden aangetroffen; bij mondzorg is dat patroon anders. Een mogelijke verklaring hiervoor kan zijn (wijzigende) regelgeving die per 1 januari van dat jaar in gaat.

In de bestanden van farmacie zijn door Vektis de correcties zo veel als mogelijk verwerkt in de originele declaratie indien het mogelijk was deze te koppelen. Daardoor is het aantal correctierecords bij farmacie laag. Deze werkwijze volgt Vektis (nog) niet bij de bestanden voor de huisartsen, mondzorg en GGZ waardoor we daar soms hoge aantallen correctierecords in de bestanden zien.

De oorsprong en reden van de correctierecords zijn niet eenduidig vast te stellen. Eén van de redenen kan zijn dat de zorgverzekeraar deze records heeft toegevoegd en daarmee de indiener “corrigeert” voor declaraties die een te hoog bedrag betreffen. Ook is het mogelijk dat een uitgekeerde vergoeding niet onder de polisvoorwaarden viel en de verzekeraar dit verrekent met de verzekerde. En natuurlijk kan een zorgindienaar of zorgbehandelaar onderkennen dat er een fout in zijn/haar computersysteem is gemaakt die hersteld moet worden. Er zijn daarnaast nog meer redenen waarom we een correctierecord aan kunnen treffen.

Dit verschijnsel is in alle vier segmenten zichtbaar; het treedt het meest op bij mondzorg en het effect is daar het grootst. De creditrecords komen vooral bij prestaties die gebruikt zijn in onze toetsen.

Voor een juiste benadering⁷ van de correctierecords introduceren we **het netto aantal records**: het aantal records met een positief bedrag minus het aantal records met een negatief bedrag. Er zijn ook records in de data bestanden met een bedrag gelijk aan 0; daarom is het mogelijk dat het terugrekenen naar het bruto en/of aantal positieve records ogenschijnlijk niet klopt.

Als voorbeeld geven we hier een case van een verzekerde waarvoor in 2010 in totaal 123 records in het bestand van mondzorg werden aangetroffen. Nader onderzoek leerde dat het hier om 63 records ging met een positief bedrag (uitbetalingen) en 60 records met een negatief bedrag (creditrecords = correctierecords). Het netto aantal records voor deze verzekerde bedraagt dus $63 - 60 = 3$.

De tabellen 1 tot en met 4 geven een overzicht van het verschijnsel van de correctieboekingen voor de vier zorgsegmenten.

Jaartal	Huisartsen				
	Aantal correctie records	Netto aantal records	Totale bedrag correcties in miljoen euro	Percentage van de correcties in totaal aantal records	Percentage van het bedrag van de correcties in het totaal gedeclareerd bedrag
2010	1.321.459	259.914.793	9,2	0,5	0,4
2011	2.630.912	299.373.280	16,0	0,9	0,6
2012	1.248.005	311.413.569	6,9	0,4	0,3

Tabel 1. Correctieboekingen in het segment huisartsen 2010-2012

Jaartal	Mondzorg				
	Aantal correctie records	Netto aantal records	Totale bedrag correcties in miljoen euro	Percentage van de correcties in totaal aantal records	Percentage van het bedrag van de correcties in het totaal gedeclareerd bedrag
2010	765.810	67.763.707	37,5	1,1	1,7
2011	841.923	66.702.404	37,9	1,3	1,7
2012	302.079	49.801.964	22,3	0,6	1,0

Tabel 2. Correctieboekingen in het segment mondzorg 2010-2012

⁷ Er zijn ook andere manieren om deze creditrecords te behandelen.

Jaartal	Farmacie				
	Aantal correctie records	Netto aantal records	Totale bedrag correcties in miljoen euro	Percentage van de correcties in totaal aantal records	Percentage van het bedrag van de correcties in het totaal gedeclareerd bedrag
2010	23.727	191.636.481	1,1	0,0	0,0
2011	14.427	208.431.051	0,8	0,0	0,0
2012	16.634	219.113.995	0,7	0,0	0,0

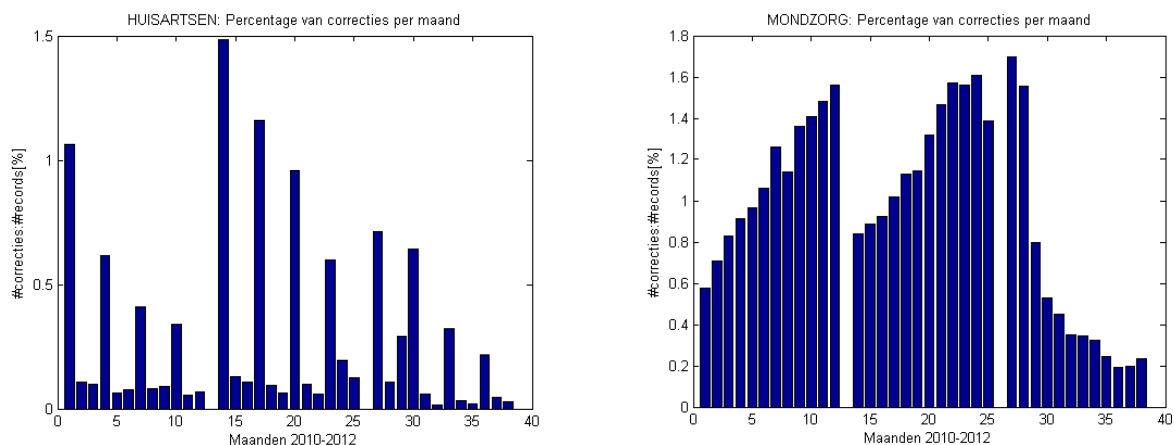
Tabel 3. Correctieboekingen in het segment farmacie 2010-2012

Jaartal	GGZ				
	Aantal correctie records	Netto aantal records	Totale bedrag correcties in miljoen euro	Percentage van de correcties in totaal aantal records	Percentage van het bedrag van de correcties in het totaal gedeclareerd bedrag
2009	18.679	1.120.268	104,5	1,7	2,6
2010	17.919	1.176.607	93,7	1,5	2,4
2011	19.211	1.206.324	80,7	1,6	2,0

Tabel 4. Correctieboekingen in het segment GGZ, inclusief verrekenpercentages, 2009 - 2011

Hoewel het aantal correctierecords vergeleken met het totaal aantal records relatief klein is, kan de impact groot zijn wanneer we specifieke zorgsegmenten en declaraties onderzoeken. Bij Mondzorg en GGZ zijn deze percentages hoog: tenminste 1% van de records was een creditrecord.

Wanneer we verdeling van de correctierecords nader onderzoeken (in figuur 1 worden de maandoverzichten van huisartsen en mondzorg gegeven) dan kan worden geconstateerd dat het aantal correctierecords in 2012 enigszins achterblijft met de omvang van dit verschijnsel in 2010 en 2011. Naar verwachting moeten over 2012 nog correctierecords door zorgverzekeraars aan Vektis worden aangeboden.



Figuur 1. Overzicht van percentage creditrecords per maand, huisartsen en mondzorg, 2010- 2012

Door wie een declaratie wordt ingediend, is deels te herleiden in het bestand. Wie en de wijze waarop de declaratie is ingediend, kan een sterke variabele zijn om onregelmatigheden in declaraties op te sporen; sommige verzekeraars werken daarom al met (scanning-)software die extra alert is op “beschadigingen en wijzigingen” op de declaratie die door een verzekerde wordt ingediend. Wanneer de declaratie rechtstreeks door de zorgaanbieder wordt ingediend, ontbreekt het “controlerend oog” van de verzekerde zelf. Tenslotte, als een factoringmaatschappij voor de verzameling en aanbidding van de declaraties aan de zorgverzekeraar zorgt, is dit weer anders. De wijze van aanbidding zal ongetwijfeld bij vervolgonderzoek meer aandacht krijgen.

Datzelfde geldt voor de zorgverzekeraar (UZOV): in dit onderzoek is niet getracht verschillen tussen verzekeraars op te sporen. Hun prudentie en oplettendheid (van de IT-systemen en hun medewerkers) zullen ongetwijfeld verschillen. Doordat hier de gegevens van alle declaraties beschikbaar zijn, kan worden nagegaan of er verschillen in aangetroffen declaraties van dezelfde zorgaanbieder zijn bij verschillende zorgverzekeraars.

2.2. Gebruikte analysetechnieken en zoekmethoden

Het zoeken naar anomalieën is geenszins een “*in the blind*”- doorzoeken van de data. Er wordt gericht gezocht op basis van wat al bekend is op het terrein van fraude in de zorg, op basis van de kennis van domeinexperts, op basis van de kennis van fraude experts en op basis van andere succesvolle fraudedetectie projecten. Los daarvan kijken een aantal projectmedewerkers met een frisse en onbevangen blik naar de data om tunnelvisie en *framing* te voorkomen in dit project.

Om er voor te zorgen dat de analyse niet gehinderd wordt door kleine aantallen wordt voor de meeste modelmatige technieken (zie appendix 1) een minimum aantal verrichtingen⁸ vereist. Hierdoor worden kleine afwijkende groepen (bijvoorbeeld mondhygiënist en tandartsen in ruste) buiten de analyse gehouden. Deze kunnen overigens afzonderlijk worden bekeken.

⁸ Dit minimum aantal verrichtingen verschilt van analyse tot analyse.

In Appendix 2 t/m 5 worden overzichten gegeven van logische toetsen die door domeinexperts zijn opgesteld voor de vier zorgsegmenten. Deze toetsen bleken zeer nuttig te zijn in de detectie. Bij sommige van die toetsen bleken weinig overtredingen te worden gevonden (in aantal of omvang van het te veel vergoede bedrag), andere toetsen lieten forse aantallen en bedragen zien. De opbrengst van deze laatste toetsen wordt gepresenteerd.

Los daarvan zijn succesvolle detectie resultaten geboekt met (afgeleide) variabelen (mede in relatie tot een bijbehorende referentiegroep) zoals (niet uitputtend overzicht):

- Aantal niet-standaard prestatiecodes per behandelaar
- Percentage huisbezoeken per huisarts
- Aantal bezoeken per dag van een behandelaar
- Aantal consulten per dag van een behandelaar
- Aantal consulten van één patiënt op een dag
- Gemiddelde tijd per patiënt van een behandelaar
- Combinatie van consulten op één dag
- Prestaties al dan niet in relatie met ketenzorg
- Aantal dagen waarop een prestatie heeft plaatsgevonden
- Verhouding enkele/dubbele zorg, verhouding licht/zwaar; verhouding kort/lang consult, etc.
- Verhouding gewone gebitsreiniging/uitgebreide, verhouding eenvoudige extractie/moeilijke extractie, verhouding 1-, 2- en 3-vlaksvulling, etc.
- Analyse per leeftijdsgroep
- Eigen bijdrage
- Overschrijden van maximum bedragen
- Verhouding panoramische foto/gewone foto
- Verhouding verrichtingen binnen een praktijk tussen medicus en ondersteuner
- Verschil tussen berekende vergoeding en toegekende vergoeding
- Dubbele facturen, ook dubbel voor zowel materiaal als honorarium
- Meer dan één zorgverzekering van een verzekerde
- Declareren op praktijkcode én eigen AGB-code
- Het met een te hoge frequentie declareren van het inschrijftarief
- Samenhangende verrichtingen: prestatiecodes die bij elkaar horen zoals verdoven, daarna extractie, etc.
- Verhouding tussen het aantal inschrijvingen en het aantal daadwerkelijke consulten

Daarnaast vormen deze gegevens een goede bron voor het afleiden van andere inzichten zoals bijvoorbeeld het totale bedrag uitgekeerd per verzekerde per zorgverzekeraar in deze segmenten.

3. Resultaten van de data-analyses

3.1 Inleiding

Voor elk van de vier zorgsegmenten die *in scope* zijn, worden in dit hoofdstuk overzichten en enkele opvallende resultaten gegeven. Voor een methodologische toelichting op de wijze waarop merkwaardige cases zichtbaar zijn geworden, wordt verwezen naar Appendix 1.

De data bestanden bevatten de declaraties die daadwerkelijk door de zorgverzekeraars zijn vergoed. Het betreft de jaren 2010, 2011 en 2012 voor wat betreft huisartsen, mondzorg en farmacie. Voor GGZ beschikken we over de jaren 2009, 2010 en 2011. Van de meest recente jaren (met name 2012) kunnen in de toekomst nog correcties (creditrecords) plaatsvinden, voor eerdere jaren ligt dit niet voor de hand.

De structuur van de paragrafen geeft de werkwijze goed weer: na een algemene beschrijving van de beschikbare variabelen en gegevens volgen als eerste stap de resultaten van een exploratieve analyse van de gegevens. Daar komen aantal patiënten⁹, omzet, aantal declaraties, gemiddeld bedrag per declaratie, gemiddelde per patiënt, zorgaanbieder, etc. aan de orde. Al uit de exploratieve analyse komen veel merkwaardigheden naar voren, een aantal *highlights* wordt belicht.

Daarna wordt kort gerapporteerd over de uitslagen van de toetsen die gericht zijn op het opsporen van overtredingen van declaratieregels. Enkele markante resultaten worden gegeven. Niet alle toetsen “leveren” evenveel op: enkele toetsen zijn zeer effectief, andere toetsen leiden tot kleinere aantallen overtredingen en/of kleinere bedragen die overigens ook de juiste aandacht verdienen. De toetsen die relatief niet veel geld op leveren, komen in dit rapport niet aan de orde hoewel ze gezamenlijk voor een substantieel bedrag kunnen zorgen.

Daarna wordt stilgestaan bij anomalieën die met meer geavanceerde technieken zijn opgespoord. Elk zorgsegment wordt afgesloten met een overzicht van te veel vergoede bedragen.

⁹ De aanduiding patiënt en verzekerde zijn synoniemen van elkaar in dit rapport.

3.2 Resultaten voor huisartsen

3.2.1 Overzicht van variabelen

Voor een uitgebreide toelichting op de data waarop de detectie is gebaseerd, zie Appendix 6.

Het gaat om goedgekeurde declaratierecords dat wil zeggen dat deze declaraties vergoed zijn. In een klein aantal gevallen worden negatieve bedragen aangetroffen, dit zijn creditrecords die we ook wel correctieboekingen noemen. Door te aggregeren kan inzicht per verzekerde worden verkregen, net als inzicht per huisarts, per praktijk, per prestatiecode, per kwartaal, per jaar, etc., etc.

Het aantal records voor dit onderzoek is groot, in totaal 881 miljoen. Het betreft vergoedingen voor een totaal van 7,5 miljard euro in de jaren 2010 – 2012.

De verdeling van het aantal records over deze jaren wordt in tabel 1 weergegeven. De indruk bestaat dat in 2010 een deel van de gegevens ontbreekt. De oorzaak hiervan is dat niet alle zorgverzekeraars declaraties bij Vektis hebben aangeleverd. Na een algemene controle op het data bestand is geconcludeerd dat het gemis (circa 10% van de records) geen belemmering vormt voor het opsporen van anomalieën.

Anderzijds is het opvallend dat het aantal AGB-behandelaars in 2012 zo groot is. Dit is onderzocht en daaruit bleek dat één zorgverzekeraar bij deze variabele in de fout is gegaan en 6.770 ongeldige AGB-codes heeft aangeleverd. Het bedrag dat met deze ongeldige AGB-codes is gemoeid is 6 miljoen euro en is circa 2% van het totale bedrag dat door deze zorgverzekeraar in 2012 is vergoed voor huisartsen. We gaan uit van een administratieve fout en nemen deze extra 10% niet mee in de schattingen en berekeningen. Het vormt geen obstakel in deze fase van de anomalie detectie. Wanneer bij vervolgstappen nagegaan moet worden wat er met één of meer van deze AGB-behandelaars aan de hand is, is dat wel een hindernis.

Jaartal	Aantal records	Totaal bedrag vergoed in miljoenen euro	Aantal AGB-behandelaars	Aantal AGB-indieners
2010	262.584.340	2.292	14.950	12.976
2011	304.654.670	2.619	17.052	13.022
2012	313.926.643	2.595	24.532	18.738
Totaal	881.165.653	7.506		

Tabel 1. Aantal records betreffende vergoedingen aan huisartsen in de jaren 2010 – 2012

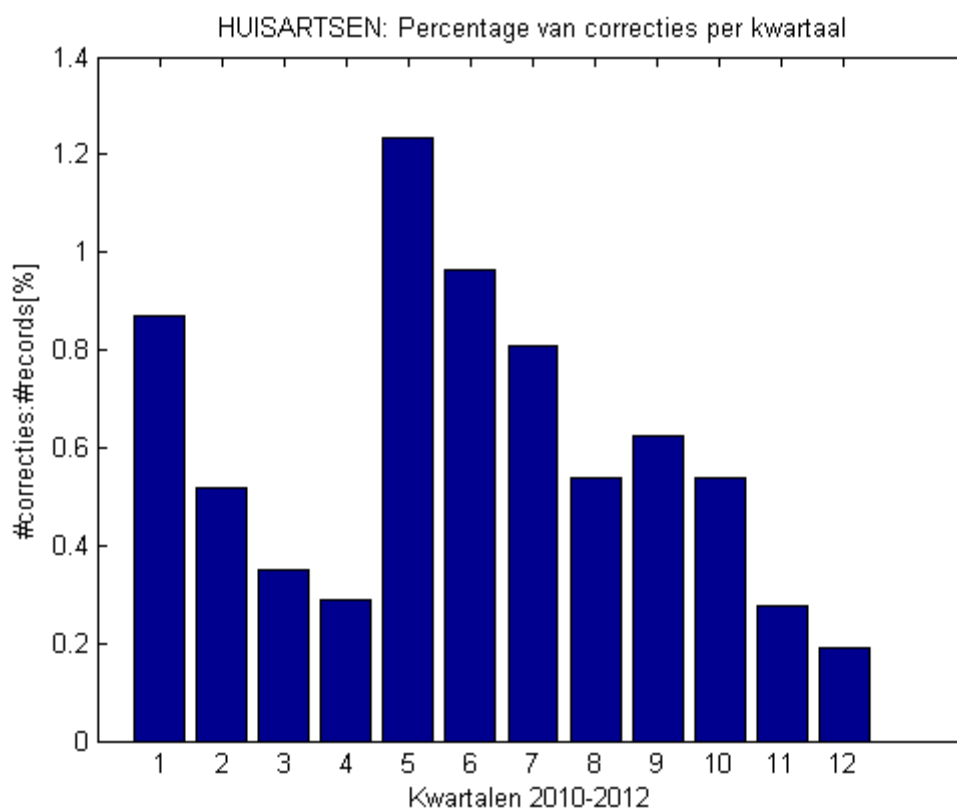
Los van alle scripts en overzichten die aan de NZa ter beschikking zijn gesteld, worden hier enkele opmerkelijke observaties gepresenteerd. Ze zijn niet uitputtend en proberen een beeld te geven van de breedte van de analyses.

Zoals beschreven in hoofdstuk 2 bevindt zich in het data bestand een aanzienlijk aantal creditrecords (correctierecords). De creditrecords zorgen ervoor dat het totale bedrag dat aan huisartsen in 2010 is toegekend met circa 9 miljoen daalt, in 2011 met 16 miljoen en in 2012 met 7 miljoen euro.

Jaartal	Huisartsen				
	Aantal correctie records	Netto aantal records	Totale bedrag correcties in miljoen euro	Percentage van de correcties in totaal aantal records	Percentage van het bedrag van de correcties in het totaal gedeclareerd bedrag
2010	1.321.459	259.914.793	9,2	0,5	0,4
2011	2.630.912	299.373.280	16,0	0,9	0,6
2012	1.248.005	311.413.569	6,9	0,4	0,3

Tabel 2. Correctieboekingen in het segment huisartsen 2010-2012

Figuur 1 geeft een overzicht van de verdeling van de correctieboekingen over de tijd. In het eerste kwartaal is het percentage correctieboekingen het hoogst en in de loop van het jaar neemt het af. Het hoge niveau in 2011 en de eerste twee kwartalen van 2012 is opvallend.



Figuur 1. Overzicht van het percentage correctieboekingen in de loop van de tijd, 2010 – 2012

3.2.2 Overzicht van frequentie van declaraties van huisartsen

Een illustratief voorbeeld is de frequentie waarmee verzekerden voorkomen in het bestand van de huisartsen. Het gaat om de top 10 per jaar; elk jaar bevinden zich andere verzekerden in deze top tien. Zie tabel 3.

2010	2011	2012
457	424	788
445	417	721
413	407	476
350	360	346
306	328	314
304	326	312
303	319	307
298	318	301
296	305	299
294	305	286

Tabel 3. Top 10 voor 2010, 2011 en 2012 van de frequentie van verzekerden, netto records.

Zo is te zien dat er één verzekerde is die in 2012 voor maar liefst 788 records bij de huisarts zorgde. Wanneer dieper op deze case wordt ingegaan, ontstaat het volgende beeld.

Volgens de gegevens in het data bestand gaat het hier om een vrouw van 58 jaar die (onder andere) haar huisarts 2 – 4 keer per dag belt (in totaal 648 telefonische consulten), alsmede 36 korte consulten en 32 lange consulten. Het totaal vergoede bedrag is € 6.289. Er zijn twee behandelaars die voor haar hebben gedeclareerd: één diende 758 declaraties in, de andere 30 declaraties. Er zijn geen correctierecords aangetroffen.

3.2.3 Overzicht van het aantal prestaties per declaratie van huisartsen

Een declaratie heeft altijd betrekking op één verzekerde. Het is dan ook te verwachten dat er niet veel meer dan één prestatie per declaratie voorkomt. Dit blijkt echter in het data bestand niet zo te zijn, zie hiervoor tabel 4 waarin een (beperkt) overzicht is gegeven.

In een declaratierecord kan - dus - een aantal prestaties staan (al dan niet onterecht) dat groter is dan 1. Het bedrag in de kolom declaratiebedrag is het totaal bedrag voor alle prestaties samen binnen dat declaratierecord, en dit bedrag is vergoed. Relatief grote bedragen in een declaratie is een verschijnsel dat ook in ander onderzoek is gevonden: verschrijvingen in het aantal zijn hiervan soms de oorzaak. In plaats van “1” toetst men 10, 100 of 1000. Maar ook verschrijvingen in de dosis zijn in dat andere onderzoek gesignaleerd. Het bedrag dat met dit soort “verschrijvingen” is gemoeid, schiet snel omhoog. En bovendien is een dergelijke verschrijving eenvoudig te proberen. Het is natuurlijk opmerkelijk dat deze declaraties wel vergoed zijn. Een logische verklaring is niet gevonden; duidelijk is dat het hier om grote bedragen kan gaan. Bij de opvolging van de cases met het hoogste aantal prestaties per declaratie moet worden nagegaan of dit reële cases zijn of niet¹⁰. Dit verschijnsel wordt nu niet meegenomen bij de berekening van het te veel vergoede bedrag in die jaren.

Navraag bij Vektis leerde dat Vektis dit aantal prestaties soms ook (her)berekent. Het is dan ook niet onomstotelijk aan te geven in hoeverre deze aantallen prestatiecodes worden ingebracht door de indiener, door de zorgverzekeraar of door Vektis. Deze variabele verdient bij vervolgonderzoek de aandacht.

¹⁰ Vektis stelt dat de variabele “aantal prestaties per declaratie” een zwakke variabele is. Toch zien we dat er een relatie is tussen het bedrag dat is vergoed en het aantal prestaties dat op de declaratie staat. Ook het verschijnsel van “verschrijvingen” zorgt ervoor dat deze analyse betekenisvol is. Nader onderzoek is noodzakelijk om bijvoorbeeld na te gaan of dit data veld wel met de juiste variabele is gevuld.

Aantal prestaties per declaratie ¹¹	2010	2011	2012
1	259.871.408	299.159.436	310.827.512
2	57.806	74.556	95.738
3	3.358	3.546	3.807
4	771	701	517
5	216	195	163
6	124	125	52
7	70	138	50
8	44	86	22
9	50	45	12
10	311	42	43
11	64	17	16
12	12	34	23
13	4	7	2
14	13	52	10
15	22	15	7
16	4	4	4
17	1	6	3
18	0	0	3
19	0	1	3
20	20	11	4
25	3	205	1
28	4	316	1
30	20	46	15
40	2	0	1
50	3	0	5
60	4	0	8
70	0	1	0
80	1	1	0
90	1	0	1
100	6	0	3
200	2	0	0
240	0	32	0
360	2	0	1
500	2	0	0
1000	4	3	1
2930	0	1	0
4095	0	0	1

Tabel 4. Aantal prestaties per declaratie, 2010 – 2012 (na 20 is het detailniveau verkleind, en dit detailniveau is nog verder verkleind boven 100 prestaties per declaratie)

¹¹ Er zijn ook records waarbij het aantal prestaties gelijk is aan 0. Deze zijn niet in deze tabel opgenomen.

3.2.4 Overzicht van declaratiebedragen

Op basis van de databestanden kan een goed beeld worden geschetst van de verschillen tussen de jaren. Zo is uit tabel 5 te lezen dat het totale gedeclareerde bedrag voor huisartsen in deze drie jaren nagenoeg constant is geweest, namelijk zo'n 2,5 miljard. Het iets lagere bedrag in 2010 is te wijten aan het feit dat niet alle zorgverzekeraars declaraties hebben aangeleverd, zie ook paragraaf 3.2.1.

We zien verder dat het gemiddelde bedrag per declaratie in 2012 zo'n 5% lager is ten opzichte van 2011. Dit kan mogelijk worden verklaard door de lagere inschrijftarieven die in 2012 gelden.

Jaartal	Totaal bedrag gedeclareerd in miljoenen euro	Gemiddeld bedrag van een declaratie in euro	Mediaan van het bedrag van een declaratie in euro
2010	2.292	8,82	4,50
2011	2.619	8,75	3,71
2012	2.595	8,33	4,13

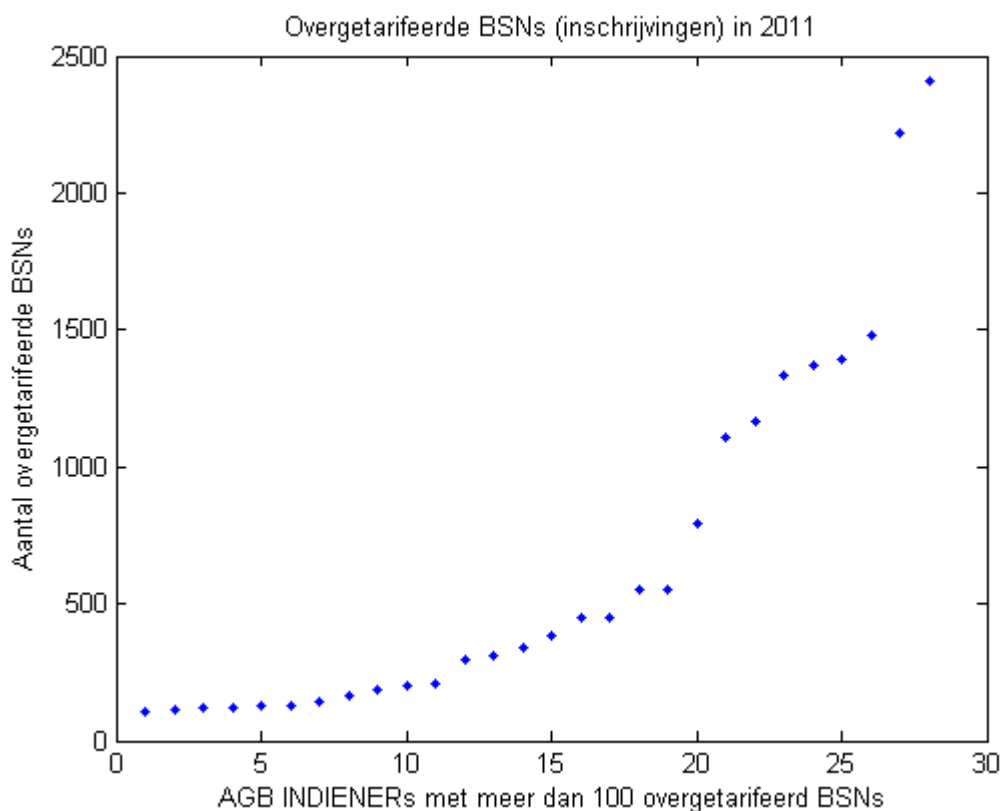
Tabel 5. Overzicht van de declaraties, 2010 – 2012

3.2.5 Overschrijdingen van het inschrijftarief

Huisartsen mogen per kwartaal voor elke op de eerste van het kwartaal ingeschreven verzekerde een inschrijftarief (codes 11000, 11100, 11101, 11102, 11103, 11104) declareren. Dit betekent dat per verzekerde maximaal vier declaraties per BSN per jaar mogen optreden.

We noemen een BSN **overgetarifeerd** wanneer voor deze persoon 5 of meer inschrijfdeclaraties (na correcties) hebben plaatsgevonden. Figuur 2 geeft een overzicht van de resultaten van dit verschijnsel voor de huisartsen met meer dan 100 overgetarifeerde verzekerden, voor het jaar 2011¹², voor de codes 11000, 11100, 11101, 11102, 11103 en 11104. Tabel 6 geeft enkele resultaten. In 2010 zijn 48.665 verzekerden overgetarifeerd, in 2011 zijn dit er 31.355 en in 2012 gaat het om 43.180 verzekerden.

¹² Vergelijkbare figuren zijn er voor 2010 en 2012



Figuur 2. Overzicht van AGB-indieners¹³ met overgetarifeerde verzekerden, 2011

De schatting van het te veel in rekening gebrachte bedrag bedraagt voor 2010: € 1,8 miljoen, voor 2011¹⁴: € 0,7 miljoen en voor 2012: € 1,5 miljoen. Dit teveel betaalde bedrag is berekend als zijnde het verschil tussen het uitgekeerd bedrag en het bedrag waar de indiener recht op had.¹⁵

Jaartal	Bedrag overgetarifeerd in miljoen euro
2010	1,8
2011	0,7
2012	1,5

Tabel 6. Overzicht van het bedrag overgetarifeerd op inschrijftarieven, 2010 – 2012

¹³ Vergelijkbare figuren zijn er per huisarts en per jaar.

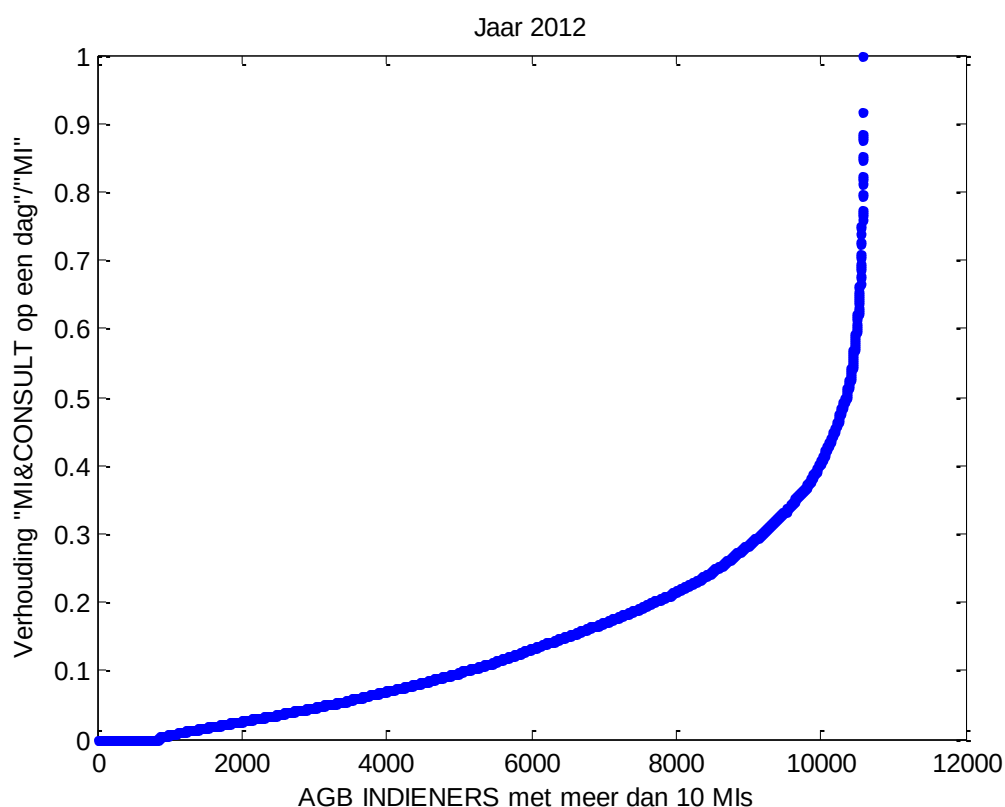
¹⁴ In 3.2.8.1 wordt een bijzondere case beschreven van een huisarts die in 2011 voor meer dan 2 miljoen euro declareerde met de inschrijftarieven. Deze case is niet in deze tabel opgenomen om de vergelijking over de jaren 2010-2012 niet te verstoren.

¹⁵ Een aantal inschrijftarieven kon niet worden gecontroleerd bijvoorbeeld het tarief dat betrekking heeft op achterstandswijken. We hebben niet de beschikking over de postcode. Deze schattingen zijn dus aan de lage kant.

3.2.6 Een combinatie van M&I verrichting en een consult op dezelfde dag

Toets 19 (zie ook Appendix 2) is een erg bruikbare toets. We hebben getoetst hoe vaak er op dezelfde dag zowel een M&I verrichting als een consult in rekening worden gebracht voor één patiënt. Dit is toegestaan wanneer een patiënt daadwerkelijk twee keer komt. Het kan echter ook zo zijn dat een M&I verrichting wordt gedaan maar dat er twee dingen in rekening worden gebracht: namelijk een M&I verrichting én een consult. Wanneer we er vanuit gaan dat wanneer er zowel een M&I verrichting als een consult in rekening is gebracht voor de zorgvraag van één patiënt op één dag, en dat het consult dus onterecht in rekening is gebracht, dan gaat het om circa 500.000 overtredingen per jaar. Elk jaar worden er circa 3 miljoen M&I verrichtingen gedaan en daarvan worden er 500.000 (16%) gecombineerd met de declaratie van een kort consult.

We construeren een variabele die aangeeft in hoeverre deze merkwaardige combinatie heeft plaatsgevonden en vergelijken dit met het aantal M&I verrichtingen. Hieruit blijkt dat er ook behandelaars zijn die voor bijna elke M&I verrichting die zij in rekening brengen ook een consult declareren. De grafiek wordt in figuur 3 gegeven.



Figuur 3. Verhouding tussen het aantal ongeldige combinaties (M&I en Consult) en het aantal M&I verrichtingen, huisartsen, 2012

Het te veel vergoede bedrag is berekend door er van uit te gaan dat het hoogste bedrag dat vergoed is, het correcte bedrag is. Al de andere vergoedingen voor die patiënt op die dag voor M&I en consult worden als te veel vergoed aangemerkt.

Dit betekent dat er – per jaar - voor circa 6 miljoen teveel vergoed is, zie hiervoor tabel 7. In 2010 ging het om 473.984 gevallen, in 2011 om 489.241 cases en in 2012 om 493.857 situaties.

Jaartal	Bedrag overgetarifeerd in miljoen euro
2010	5,7
2011	6,4
2012	6,2

Tabel 7. Te veel vergoede bedragen per jaar door niet-toegestane combinatie M&I verrichting en consult

3.2.7 Meerdere consulten van een verzekerde op één dag

Voor elke BSN is gekeken of het vaak voorkwam dat er meerdere consulten werden gedeclareerd op één dag, bijvoorbeeld twee korte consulten of een kort én een lang consult. Indien een patiënt voor meerdere zorgvragen bij een huisarts komt, en deze zorgvragen in één consult worden behandeld, dan mag er ook maar één consult in rekening worden gebracht. Met de toets op meerdere consulten kunnen we hier een inschatting van maken. Indien een patiënt daadwerkelijk twee keer op een dag bij de huisarts komt dan mogen er wel meerdere consulten in rekening worden gebracht. De uitkomst van deze toets bevat hierdoor ook *false positives*.

Er bleken meerdere consulten op één dag voor te komen. Het te veel gedeclareerde bedrag bedraagt voor 2010 2,6 miljoen, voor 2011 6,3 miljoen en voor 2012 3,9 miljoen.

Het te veel toegekend bedrag per case is berekend door er vanuit te gaan dat het hoogste bedrag de juiste toekenning is, en de som van alle andere lagere bedragen het teveel vergoede bedrag is.

3.2.8 Merkwaardige individuele cases

3.2.8.1 Case 1: huisarts niet in het AGB-bestand en uitzonderlijk hoge declaraties

Uit de analyses is een opmerkelijke case naar voren gekomen, laten we deze huisarts X noemen. Huisarts X declareerde elk kwartaal voor 1.400 patiënten maar op 1 januari 2011 werden 1.166 declaraties met hoge declaratiebedragen (elk € 1.783 euro, een factor 100 te hoog) ingediend, in totaal € 2.078.978. Later werd voor een bedrag van € 35.560 gecorrigeerd. Hij komt in 2012 (als AGB-code indiener) niet meer voor in de AGB-registratie en op de AGB-code van de behandelaar zijn in 2012 geen declaraties meer ingediend.

3.2.8.2 Case 2: huisarts die sterk afwijkt met zijn “meest nabije buren”

Dit is de huisarts die ook besproken is als voorbeeld in Appendix 1, paragraaf 3.2. Op basis van één van de detectietechnieken is hij/zij naar voren gekomen als afwijkend in 2010.

Deze huisarts had:
2.594 lange consulten
4.053 korte consulten
119 lange bezoeken
1.209 korte bezoeken

Gemiddeld is dit percentage lange consulten 15,0%. Onder de huisartsen die erg veel op hem lijken, is dit percentage 26,3%. Het percentage lange consulten van deze huisarts is 39,0% (namelijk $2594/(2594 + 4053)$).

Ook voert deze huisarts relatief veel huisbezoeken uit ten opzichte van het aantal consulten dat hij verricht. Doordat hij relatief veel korte bezoeken heeft in plaats van lange bezoeken is hij er duidelijk niet op uit om daar het geld te verdienen.

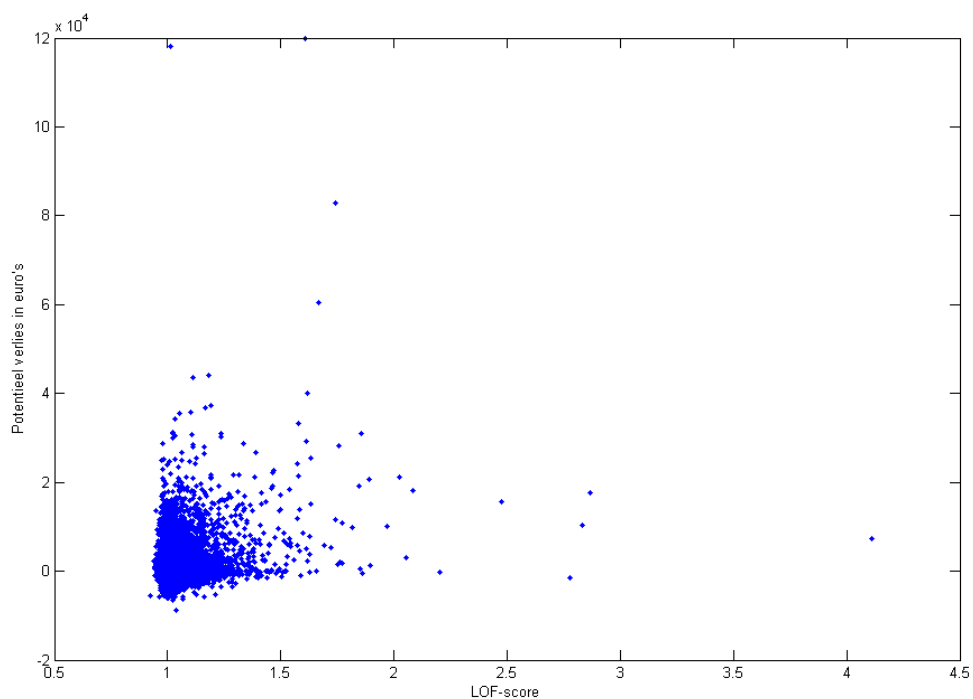
Het lijkt een arts te zijn die echt tijd neemt voor de patiënt (als die patiënt naar de praktijk komt) en hij legt veel huisbezoeken af. Maar die huisbezoeken zijn allemaal wel weer kort (1.209 van de 1.328) wat in tegenspraak lijkt met tijd nemen voor de patiënt.

3.2.8.3 Andere opmerkelijke “buren”

Figuur 4 geeft een overzicht aan van de relatie tussen de LOF-score¹⁶ en het mogelijk ten onrechte gedeclareerde bedrag bij huisartsen. Scatterplots zoals deze zijn een krachtig hulpmiddel om merkwaardige/opvallende huisartsen op te sporen. Immers, hoe hoger de LOF-score (uitgezet op de x-as), hoe afwijkender de huisartsen ten opzichte van hun collega's die het meest op hen lijken. De vier stippen die uiterst rechts liggen (LOF-score groter dan 2,5) zorgen samen voor een schatting van het te veel vergoede bedrag van circa 15.000 euro. Nader onderzoek van deze huisartsen is noodzakelijk.

Verder vallen de scores met een hoog mogelijk ten onrechte vergoede bedrag op. Ook deze cases (in het bijzonder de case van bijna € 120.000) verdienen bijzondere aandacht, net als de vergelijkbare cases in 2011 en 2012.

¹⁶ De Local Outlier Score (LOF) geeft aan elk datapunt een score. Een datapunt kan zijn een huisarts, een verzekerde, een declaratie, etc. Hoe lager de dichtheid van een punt ten opzichte van zijn buren is, hoe hoger de score en dus hoe meer een datapunt afwijkt van zijn dichtstbijzijnde buren. Zie verder Appendix 1.



Figuur 4. Scatterplot van LOF-score tegen het potentieel verlies in euro's, huisartsen, 2010.

3.2.9 Schatting van het bedrag gemoeid met het overtreden van regels en anomalieën

Zoals bekend is er een onderscheid gemaakt tussen het overtreden van declaratieregels enerzijds en het vaststellen van opvallende cases anderzijds. Bij de eerste categorie gaat het om duidelijke/harde regels, bij de tweede categorie moeten experts zich buigen over de duiding van de merkwaardige cases.

3.2.9.1 Overtreding van declaratieregels

De berekening van het te veel vergoede bedrag bij het overtreden van de declaratieregels is relatief eenvoudig. Dit betekent dat de bedragen in de onderstaande tabel relatief hard zijn.

Overtreding declaratieregels	Te veel vergoede bedrag 2010 in miljoen	Te veel vergoede bedrag 2011 in miljoen	Te veel vergoede bedrag 2012 in miljoen
Inschrijftarieven	1,8	0,7	1,5
Combinatie M&I én consult op 1 dag (19)	5,7	6,4	6,2
Meerdere consulten op één dag	2,6	6,3	3,9
Case 1 (3.2.8.1)	-	2,0	-
Totaal te veel vergoed	11,1¹⁷	15,4	11,6
Totaal vergoed	2.521 ¹⁸	2.619	2.595
Als percentage van het totale bedrag in dat jaar vergoed aan huisartsen	0,4	0,6	0,4

Tabel 8. Overzicht van de te veel vergoede bedragen in 2010, 2011 en 2012

¹⁷ Bijgeschat voor de 10% records die in het data bestand van 2010 ontbreekt.

¹⁸ Bijgeschat voor de 10% records die in het data bestand van 2010 ontbreekt.

3.2.9.2 Anomalieën

In case 2 (zie paragraaf 3.2.8.2) hebben we een aanbieder laten zien die een afwijkende verhouding had tussen lange en korte consulten. Op basis van het gekozen referentiepunt (95, 99 of 99,9) kan worden berekend welk bedrag er te veel in rekening is gebracht. Een voorbeeld hiervan is beschreven in paragraaf 1.4. Voor huisartsen hebben we ons gericht op vier verhoudingsgetallen, zie tabel 9.

Huisartsen 2010	Schatting van het te veel vergoede bedrag, in miljoen euro		
	95% niveau	99% niveau	99,9% niveau
Verhouding korte/lange consulten per dag	21,2	4,2	0,4
Verhouding korte/lange huisbezoeken	0,6	0,2	0,0
Verhouding telefonische consult/totaal consulten	2,1	0,5	0,0
Verhouding vaccinatie/totaal consulten	1,6	0,5	0,0
Totaal bedrag¹⁹	27,7	5,9	0,4

Huisartsen 2011	Schatting van het te veel vergoede bedrag, in miljoen euro		
	95% niveau	99% niveau	99,9% niveau
Verhouding korte/lange consulten per dag	23,4	4,7	0,5
Verhouding korte/lange huisbezoeken	0,5	0,1	0,0
Verhouding telefonische consult/totaal consulten	2,6	0,8	0,3
Verhouding vaccinatie/totaal consulten	2,3	0,6	0,0
Totaal bedrag	28,8	6,2	0,8

Huisartsen 2012	Schatting van het te veel vergoede bedrag, in miljoen euro		
	95% niveau	99% niveau	99,9% niveau
Verhouding korte/lange consulten per dag	22,5	4,5	0,5
Verhouding korte/lange huisbezoeken	0,4	0,1	0,0
Verhouding telefonische consult/totaal consulten	2,2	0,4	0,0
Verhouding vaccinatie/totaal consulten	1,8	0,4	0,0
Totaal bedrag	26,9	5,4	0,5

Tabel 9. Overzicht van de te veel vergoede bedragen bij anomalieën, 2010 – 2012

¹⁹ Bijgeschat voor de 10% records die in het data bestand van 2010 ontbreekt.

3.2.9.3 Overzicht huisartsen

Deze slotparagraaf brengt alle deelresultaten samen voor het huisartsensegment. De resultaten spreken voor zich. Wij achten het referentiepunt van 99% het meest reëel, zie hiervoor tabel 11.

Jaartal	Totaal bedrag vergoed in miljoenen euro A	Te veel vergoed op basis van overtreding declaratieregels B	Te veel vergoed op basis van anomalieën (95%) C	Totaal te veel vergoed (95%) D = B+C	Totaal te veel vergoed als percentage van totaal vergoed E = D/A
2010	2.521 ²⁰	11,1	27,7	38,8	1,5
2011	2.619	15,4	28,8	44,2	1,7
2012	2.595	11,6	26,9	38,5	1,5

Tabel 10. Overzicht van de totalen te veel vergoede bedragen, 2010 - 2012, 95% referentiepunt

Jaartal	Totaal bedrag vergoed in miljoenen euro A	Te veel vergoed op basis van overtreding declaratieregels B	Te veel vergoed op basis van anomalieën (99%) C	Totaal te veel vergoed (99%) D = B+C	Totaal te veel vergoed als percentage van totaal vergoed E = D/A
2010	2.521 ¹⁵	11,1	5,9	17,0	0,7
2011	2.619	15,4	6,2	21,6	0,8
2012	2.595	11,6	5,4	17,0	0,7

Tabel 11. Overzicht van de totalen te veel vergoede bedragen, 2010 - 2012, 99% referentiepunt

Jaartal	Totaal bedrag vergoed in miljoenen euro A	Te veel vergoed op basis van overtreding declaratieregels B	Te veel vergoed op basis van anomalieën (99,9%) C	Totaal te veel vergoed (99,9%) D = B+C	Totaal te veel vergoed als percentage van totaal vergoed E = D/A
2010	2.521 ¹⁵	11,1	0,4	11,5	0,5
2011	2.619	15,4	0,8	16,2	0,6
2012	2.595	11,6	0,5	12,1	0,5

Tabel 12. Overzicht van de totalen te veel vergoede bedragen, 2010 - 2012, 99,9% referentiepunt

²⁰ Bijgeschat voor ontbrekende 10% in het data bestand in 2010.

3.3 Resultaten voor mondzorg

3.3.1 Overzicht van de variabelen

Deze paragraaf 3.3 heeft betrekking op de mondzorg als totaal, dat wil zeggen op zowel de basisverzekering als de aanvullende verzekering. In paragraaf 3.3.10 beschouwen we de basisverzekering afzonderlijk; over de jeugd in de basisverzekering wordt verslag gedaan in paragraaf 3.3.11. Het onderwerp mondzorg wordt afgesloten met een slotparagraaf waar we een vergelijking maken tussen de percentage te veel vergoed voor deze drie domeinen.

Een overzicht van de variabelen wordt in Appendix 6 gegeven. Het gaat om declaratierecords dat wil zeggen dat deze declaraties vergoed zijn. Door te aggregeren kan inzicht per verzekerde worden verkregen, net als inzicht per tandarts, per tandartspraktijk, per prestatiecode, per orthodontist, per kwartaal, per jaar, etc., etc.

Er is een stapsgewijze aanpak: eerst een exploratieve analyse, vervolgens de toetsen op overtreding van declaratieregels en daarna anomalie detectie. De paragraaf wordt afgesloten met een overzicht van te veel vergoede bedragen.

Het aantal records voor dit onderzoek is groot, in totaal 190 miljoen. Het betreft vergoedingen voor een totaal van 6,5 miljard euro in de jaren 2010 – 2012. Bij de analyse is rekening gehouden met het feit dat de prestatiecodes over de periode 2010/2011 verschillen van die van 2012.

De verdeling van het aantal records over deze jaren wordt in tabel 1 weergegeven.

Jaartal	Aantal records	Totaal bedrag vergoed in miljoenen euro	Aantal AGB-behandelaars	Aantal AGB-indieners
2010	69.297.896	2.179	15.190	8.439
2011	68.449.999	2.180	15.550	8.325
2012	51.934.447	2.179	15.486	8.289
Totaal	189.682.342	6.538		

Tabel 1. Aantal records betreffende vergoedingen aan mondzorg in de jaren 2010 – 2012

Los van alle scripts en overzichten die aan de NZa ter beschikking zijn gesteld, worden hier enkele opmerkelijke observaties gepresenteerd. Ze zijn niet uitputtend en proberen een beeld te geven van de breedte van de analyses.

Het data bestand van de mondzorg bevat veel creditrecords (correctieboekingen)²¹. Tabel 2 geeft een overzicht van de correctieboekingen voor de mondzorg. We constateren bij sommige tandartsen een grote diversiteit: het aantal correctie boekingen is dan enorm hoog vergeleken bij de normale declaraties. Het is mogelijk dat dit niet aan de tandarts ligt maar aan andere partijen in het proces dat heeft geleid tot de data bestanden waarover we nu beschikken. Het gaat in totaal om 765.810 correcties (2010), 841.923 (2011) en 302.079 (2012)²², waarmee respectievelijk een totaal bedrag overeenkomt van € 37,5 miljoen, € 37,9 miljoen en € 22,3 miljoen. Procentueel gaat het respectievelijk om 1,7%, 1,7 % en 1,0% van het totaal gedeclareerde bedrag in het mondzorgsegment in die jaren.

²¹ Een deel van de correctieboekingen wordt mogelijk verklaard door de polisvoorwaarden van de aanvullende verzekering, waarin declaraties slechts deels vergoed worden.

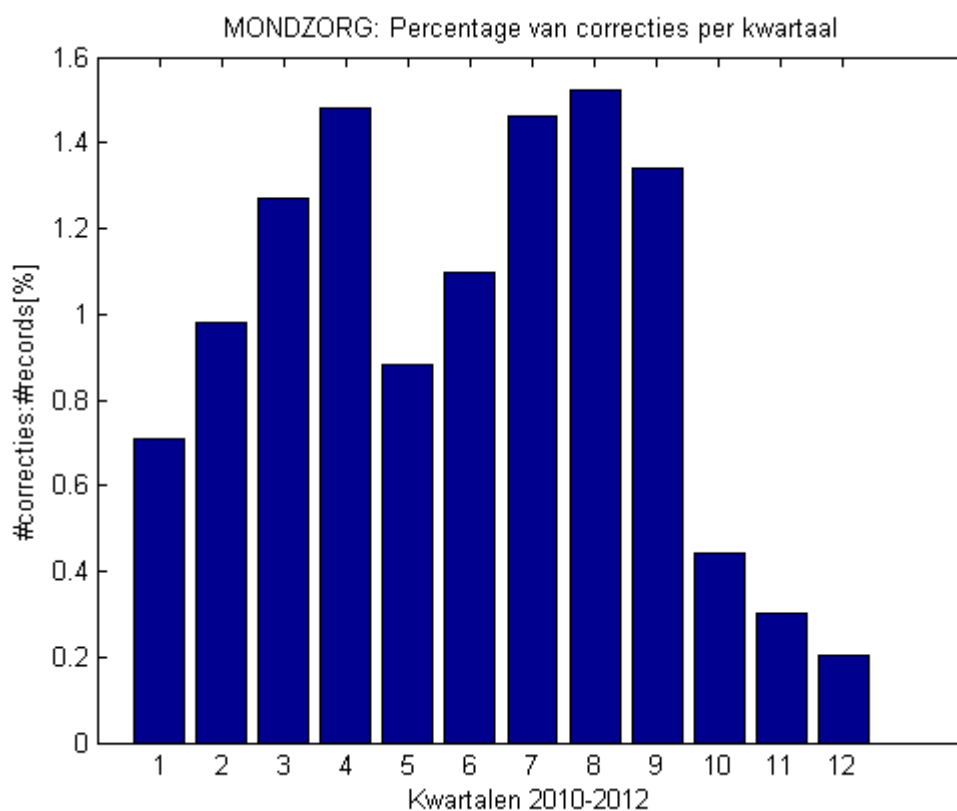
²² Meest correcties bij de frequentie 1.

Het lagere aantal correctierecords over 2012 is een indicatie dat een aantal correcties wellicht nog aan Vektis aangeboden moet worden en nog niet in het thans beschikbare bestand over 2012 zijn opgenomen.

Jaartal	Mondzorg				
	Aantal correctie records	Netto aantal records	Totaal bedrag correcties in miljoen euro	Percentage van de correcties in totaal aantal records	Percentage van het bedrag van de correcties in het totaal gedeclareerd bedrag
2010	765.810	67.763.707	37,5	1,1	1,7
2011	841.923	66.702.404	37,9	1,3	1,7
2012	302.079	49.801.964	22,3	0,6	1,0

Tabel 2. Correctieboekingen in het segment mondzorg 2010-2012

Figuur 1 laat zien dat het percentage correcties in de laatste drie kwartalen van 2012 vrij laag is: er moeten waarschijnlijk nog correcties over 2012 door zorgverzekeraars aan Vektis worden aangeleverd. Ook is te zien dat – in tegenstelling met huisartsen – het percentage correctieboekingen – in de loop van het jaar toeneemt.



Figuur 1. Overzicht van het percentage correctieboekingen in de loop van de tijd, 2010 – 2012

3.3.2 Overzicht van frequentie van declaraties bij mondzorg

Een illustratief voorbeeld is de frequentie waarmee verzekerden voorkomen in het bestand van de tandartsen. We geven de top 10 van deze verzekerden met het aantal records (gesorteerd op het aantal netto records) waarmee we ze tegen kwamen in het onderzoek. Zie tabel 3.

Een rekenvoorbeeld om te verduidelijken: in 2011 heeft een verzekerde 441 records. Daarvan hebben er 436 een positief bedrag en 5 een negatief bedrag. Dit brengt het netto aantal records voor deze verzekerde op $436 - 5 = 431$.

2010 Totaal aantal records	2010 Correctierecords bij deze verzekerde	2010 Netto aantal records	2011 Totaal aantal records	2011 Correctierecords bij deze verzekerde	2011 Netto aantal records	2012 Totaal aantal records	2012 Correctierecords bij deze verzekerde	2012 Netto aantal records
2.167	0	2.167	441	5	431	239	0	239
311	0	311	300	0	300	156	0	156
246	0	246	287	0	287	135	0	135
244	0	244	174	0	174	129	4	125
227	0	227	174	0	174	124	0	124
203	0	203	167	0	167	121	0	121
199	0	199	165	0	165	120	0	120
198	0	198	163	0	163	118	1	117
195	0	195	193	17	159	116	0	116
190	0	190	159	0	159	114	0	114

Tabel 3. Top 10 van de frequentie van verzekerden, inclusief en exclusief correctierecords, mondzorg, jaren 2010 – 2012

Zo is te zien dat er één verzekerde is die in 2010 voor maar liefst 2.167 declaratierecords bij de tandarts zorgde. Dit is een bijzonder merkwaardige case.

Het gaat om een vrouw van 49-jarige leeftijd die overwegend door één behandelaar is behandeld. Onder die 2.167 declaraties troffen we geen enkele creditrecord aan. Het gaat om een totaal bedrag van € 146.710 en nagenoeg alle prestaties die mogelijk zijn, zijn verricht. Namelijk 42 verschillende codes, slechts 1 röntgenfoto, 138 keer de code "Vorbereiding praktijkruimte ten behoeve van chirurgische verrichtingen", nagenoeg altijd op een doordeweekse dag.

Haar tandarts valt niet bijzonder op, hij/zij is niet verbonden aan een instelling, is in 1993 geregistreerd en werkt in een klein dorp.

Nader onderzoek leert dat het in deze tabel met dertig patiënten (10 patiënten voor 2010, 2011 en 2012) om 29 *verschillende* patiënten gaat: er is er één die voor 311 records zorgde in 2010 en voor 239 in 2012. Wellicht scoorde hij/zij ook in 2011 hoog maar bleef hij/zij net buiten de top 10. Dit zijn ongebruikelijk veel declaraties voor één persoon en hier moet iets bijzonders aan de hand zijn.

3.3.3 Overzicht van het aantal prestaties per declaratie van tandartsen

In paragraaf 3.2.3 is beschreven dat er bij huisartsen in sommige declaraties wel erg veel prestaties werden gedeclareerd. Ditzelfde verschijnsel nemen we ook waar bij mondzorg. Tabel 4 geeft het overzicht van het aantal prestaties. Daarin is te zien dat in het algemeen het gaat om één prestatie per declaratie.

Het hoogste aantal prestaties per declaratie bedroeg 8.008 (2010), 9.667 (2011) en 6.174 (2012).²³

De prestatiecode kan worden aangemerkt als een *key*-variabele in deze analyse omdat het rechtstreeks verwijst naar de verrichte prestatie. In een klein percentage van de records ontbrak deze en dit is nader onderzocht (het veld 'prestatiecode' is helemaal leeg). Zie hiervoor ook Appendix 6.3, tabel 2.

Het betekent dat in 2010 er 9,3 miljoen euro is uitbetaald terwijl niet in de declaratie staat waarvoor; in 2011 is dit 12,7 miljoen en in 2012 gaat het om 1,4 miljoen euro.

Bij vervolgonderzoek vraagt de kwaliteit van deze variabele beslist om meer aandacht. Bij de opvolging van de cases met het hoogste aantal prestaties per declaratie moet worden nagegaan of dit reële cases zijn of niet²⁴. Dit verschijnsel wordt nu niet meegenomen bij de berekening van het te veel vergoede bedrag in die jaren. Zie ook de opmerkingen over deze variabele bij het hoofdstuk over huisartsen.

²³ Wellicht is hier sprake van een soort bundeling van facturen of een afwijkende correctie-werkwijze van (een) zorgverzekeraar(s).

²⁴ Vektis stelt dat de variabele "aantal prestaties per declaratie" een zwakke variabele is. Toch zien we dat er een relatie is tussen het bedrag dat is vergoed en het aantal prestaties dat op de declaratie staat. Ook het verschijnsel van "gemakkelijke verschrijvingen" zoals geconstateerd in andere FDEC-onderzoeken, zorgt ervoor dat deze tabel betekenisvol is. Nader onderzoek is noodzakelijk om bijvoorbeeld na te gaan of dit data veld wel met de juiste variabele is gevuld.

Aantal prestaties per declaratie ²⁵	2010	2011	2012
1	64.908.983	64.300.870	49.370.496
2	2.426.874	2.328.545	1.640.298
3	371.560	358.761	170.846
4	226.162	218.699	124.335
5	96.520	95.822	35.866
6	249.723	260.898	83.723
7	97.211	98.411	12.153
8	76.220	78.956	10.147
9	97.982	103.444	26.884
10	43.198	47.348	5.618
11	13.100	13.722	2.177
12	41.243	45.665	16.184
13	5.390	5.599	1.037
14	5.688	5.906	1.712
15	6.839	7.477	3.332
16	5.113	5.106	1.414
17	1.038	1.169	389
18	4.511	5.520	2.489
19	1.319	1.427	229
20	11.269	12.977	863
21	1.012	1.174	660
22	973	1.758	188
23	432	674	158
24	4.696	6.488	1.157
25	275	287	144
26	379	771	199
27	546	498	357
28	924	2557	187
29	35	41	29
30 ²⁶	997	902	481
33	219	191	179
36	595	539	433
39	116	114	65
42	168	129	99
45	178	197	124
48	195	209	258
60	161	108	71
100	44	8	2
101-199	113	140	73
200-299	16	10	11
300 - 999	21	7	11
Meer dan 1000	328	9	16

Tabel 4. Aantal prestaties per declaratie, mondzorg, 2010 – 2012

²⁵ Er zijn ook records waarbij het aantal prestaties gelijk is aan 0. Deze zijn niet in deze tabel opgenomen.

²⁶ Vanaf 30 zijn alleen de frequenties opgenomen wanneer ze meer dan 100 zijn, en hoge frequenties vanaf 100.

3.3.4 Overzicht van declaratiebedragen

Op basis van de databestanden kan een goed beeld worden geschetst van de verschillen tussen de jaren. Zo is uit tabel 5 te lezen dat het totale bedrag aan mondzorg in deze drie jaren nagenoeg constant is geweest, namelijk ongeveer 2,2 miljard. We zien verder dat het gemiddelde bedrag per declaratie in 2012 ongeveer 32% is gegroeid ten opzichte van 2011; een verklaring hiervoor is de wijziging in prestaties in 2012. Ook is te zien dat de verdeling van de declaraties schever is geworden: het verschil tussen het rekenkundig gemiddelde en de mediaan is in 2012 sterk gegroeid.

Jaartal	Totaal bedrag vergoed in miljoenen euro	Gemiddeld bedrag van een declaratie in euro	Mediaan van het bedrag van een declaratie in euro
2010	2.179	31,45	19,47
2011	2.180	31,85	19,63
2012	2.179	41,95	22,00

Tabel 5. Overzicht van de totale vergoeding, gemiddelde en mediaan, mondzorg, 2010 – 2012

3.3.5 Ongeldige combinatie van declaraties

Een aantal prestaties kan niet in combinatie met elkaar in rekening worden gebracht. Zie hiervoor ook Appendix 3 waarbij het voor de jaren 2010 en 2011 gaat om de toetsen 56, 57, 59-62; voor 2012 gaat het om vergelijkbare toetsen (2012: 46 en 47).

Per saldo is het te veel vergoede bedrag voor de ongeldige combinaties (het gaat om verschillende niet toegestane combinaties waarbij C11, C12, C13, C22, C28, C29 M50, M55 en M59 betrokken zijn) voor 2010 61.000 euro en voor 2011 is dat 97.000 euro.

Voor 2012 gelden andere prestatiecodes en niet-toegestane combinaties. Voor niet-toegestane combinaties (het gaat om verschillende niet toegestane combinaties waarbij A111, A211, A231 en A261 betrokken zijn) is het hiermee gemoeide bedrag 1,1 miljoen euro (17.389 gevallen). Het bedrag in 2012 voor ongeldige combinaties waarbij C112 en C114 betrokken zijn, bedraagt 0,8 miljoen euro (8.807 gevallen). Dat brengt het totaalbedrag voor 2012 op 1,9 miljoen euro.

3.3.6 Dubbele declaraties

Verder is er gekeken in hoeverre er dubbele declaraties voorkwamen, dat wil zeggen declaraties met het zelfde BSN, prestatie en datum maar *door verschillende indieners*. Met andere woorden: meerdere personen in de mondzorg hebben een factuur voor deze unieke combinatie van BSN, prestatie en datum ingediend. In dit onderzoek zijn er in 2010 13.239 van zulke declaraties gevonden, in 2011 14.336 en in 2012 9.420.

Het te veel toegekende bedrag per case dat hiermee is gemoeid is als volgt berekend: het hoogste bedrag dat toegekend is aan een indiener is aangemerkt als het juiste bedrag. De som van alle andere toegekende bedragen voor deze case wordt beschouwd als te veel vergoed²⁷.

In totaal werd er te veel vergoed aan dubbele declaraties in 2010 € 635.016, in voor 2011 € 639.749 en in 2012 € 604.848.

²⁷ Dit is een conservatieve schattingsmethode

3.3.7 Overzicht van enkele anomalieën

We hebben gekeken naar het aantal identieke prestaties dat een patiënt heeft gekregen op één en dezelfde dag en dat is vergoed. Deze analyse stoelt natuurlijk op de veronderstelling dat de datum waarop de prestatie geleverd is een harde variabele is. Wanneer de prestatie in het buitenland heeft plaatsgevonden, hoeft dit gegeven niet altijd beschikbaar te zijn.

De resultaten zijn uiterst merkwaardig. Het gaat hier om echte patiënten getuige hun encrypted BSN en al deze declaraties zijn vergoed. In alle gevallen is de zorgverlener bekend en hebben de vergoedingen plaatsgevonden.

Doordat debetrecords en creditrecords hier veel optreden, kan men – zoals bijvoorbeeld bij vullingen – niet te snel concluderen dat er honderden vullingen hebben plaatsgevonden op één en dezelfde dag bij één en dezelfde patiënt. We zullen zien (zie bijvoorbeeld tabel 6) dat er heel veel records zijn die vullingen betreffen (in 2011: 180 records voor één patiënt). Onder die 180 records bevinden zich zowel debet- als creditrecords. Als we daar goed rekening mee houden, ontstaat een overzicht zoals in tabel 7 wordt gegeven. Alle records zijn gesaldeerd en we zien dat in 2010 een patiënt bovenaan staat die – met alle correcties – nog steeds 36 records over vullingen op die één en dezelfde dag heeft. Tegelijkertijd valt een patiënt op in 2012: één record met een bedrag van € 3.762. Blijkbaar gaat het hier niet om één vulling (doorgaans geldt: één prestatie per record) maar zijn hier meer prestaties gedeclareerd die op deze één en dezelfde dag hebben plaatsgevonden.

Soortgelijke verschijnselen zijn zichtbaar bij wortelkanaalbehandelingen, extracties, röntgenfoto's en sealings. Met dit soort analyses worden de grootste uitschieters zichtbaar, zeker wanneer men zich niet tot de top 10 beperkt en de top 50 patiënten onderzoekt.

Daarnaast hebben we gekeken naar afwijkende declaraties bij jonge mensen. Het betreft in al deze analyses daadwerkelijk vergoede bedragen.

3.3.7.1 Aantal vullingen op één dag

We hebben onderzocht hoeveel records er optreden bij verzekerden die op één en dezelfde dag betrekking hebben. In deze paragraaf richten we ons op vullingen. Er worden bij sommige patiënten zeer veel records aangetroffen, zie tabel 6.

2010	2011	2012
150	180	44
80	72	36
75	63	36
66	62	36
58	54	36
54	54	36
52	49	35
52	48	35
52	48	35
49	48	33

Tabel 6. Aantal records over vullingen bij één en dezelfde patiënt op één en dezelfde dag, 2010 -2012

Nader deskresearch naar de patiënt met 150, respectievelijk 80 records betreffende vullingen op één dag in 2010 leerde het volgende.

Het betreft vullingen V11 en V13. Volgens de data was het vergoede bedrag van deze patiënt voor de tandarts op deze dag € 392,06 en heeft de patiënt 18 vullingen gekregen.

Bij de verzekerde met in 2010 80 records over vullingen op één dag blijkt het te gaan om “maar” 16 vullingen op deze dag waarvoor in totaal € 983,84 is vergoed.

Uit dit tekstblok kan men afleiden dat correctieboekingen een grote rol spelen. Wanneer we daarom uitgaan van het netto aantal records en de patiënten worden gesorteerd op het bedrag dat aan de tandarts is toegekend voor deze vullingen op één en dezelfde dag, ontstaat het volgende beeld, zie tabel 7. In deze tabel is te zien dat er in 2010 patiënten zijn die meer dan 25 declaraties van vullingen op één en dezelfde dag hebben gekregen. Het hoogste aantal in 2011 is 40 netto records op één dag terwijl het maximum in 2012 44 records is betreffende vullingen op één en dezelfde dag.

Daarnaast valt in 2012 het veel voorkomen van “ronde” bedragen in deze top 10 op. Verder vallen de grote verschillen op in vergoede bedragen: het hoogste eindbedrag dat in deze specifieke analyse naar aantal vullingen op één dag gevonden is, bedraagt € 3.762,00.

2010		2011		2012	
Netto aantal records	Eindbedrag €	Netto aantal records	Eindbedrag €	Netto aantal records	Eindbedrag €
36	2.070,20	40	2.293,60	1 ²⁸	3.762,00
2	1.978,02	2	2.200,65	44	3.028,00
31	1.729,42	32	1.983,68	32	2.774,00
2	1.725,94	2	1.782,22	28	2.632,00
28	1.721,72	1	1.735,72	32	2.431,48
4	1.611,63	28	1.735,72	36	2.160,00
25	1.537,25	40	1.735,72	2	2.088,00
1	1.524,60	29	1.689,23	32	2.080,00
5	1.491,16	3	1.689,23	26	2.045,10
2	1.486,10	26	1.611,74	28	2.016,00

Tabel 7. Top 10 (gesorteerd op eindbedrag) van het netto aantal records over vullingen en het eindbedrag voor vullingen op één en dezelfde dag, 2010-2012

Het hoogste aantal declaraties voor vullingen dat een patiënt op één en dezelfde dag heeft gekregen in 2012 bedraagt 44. De kosten per vulling voor die dag zijn voor deze patiënt € 68,82 (uitgaande van één vulling per record). Het bedrag dat per vulling wordt gedeclareerd verschilt nogal in deze groep patiënten waarvoor allemaal op één en dezelfde dag zoveel declaraties voor vullingen werden vergoed.

Over de medische aspecten kunnen we geen uitspraken doen: nadere analyse en beoordeling is nodig.

²⁸ In het veld “aantal prestaties” staat het getal 19. Dus mogelijk gaat het om 19 vullingen. Dit is wellicht ook mogelijk bij de andere kleine aantallen in de kolommen over 2010, 2011 en 2012. Dit is verder niet onderzocht.

3.3.7.2 Aantal wortelkanaalbehandelingen op één dag

We hebben onderzocht hoeveel records er in totaal zijn bij verzekerden die op één en dezelfde dag betrekking hebben op wortelkanaalbehandelingen. Er worden bij sommige patiënten zeer veel records aangetroffen, zie tabel 8. Dit zijn debet- en creditrecords.

2010	2011	2012
18	16	12
16	16	12
16	14	10
14	14	10
14	12	10
14	12	10
12	12	9
11	10	9
10	10	8
10	10	8

Tabel 8. Aantal records over wortelkanaalbehandelingen bij één en dezelfde patiënt op één en dezelfde dag, 2010-2012

Nader deskresearch van de patiënt met in totaal 18 records met wortelkanaalbehandelingen in 2010 leerde dat:

Op 15 september 2010 werden in totaal 33 records opgevoerd waaronder 18 R13 (wortelkanaalbehandeling). Voor deze patiënt kreeg de tandarts voor zijn werk voor deze patiënt op 15 september € 815,78.

In totaal waren er voor deze patiënt over heel 2010 153 records. Veel van de records zijn correctierecords maar na de correcties bleek dat voor deze patiënt in 2010 35 wortelkanaalbehandelingen zijn vergoed, voor een totaal bedrag van € 1.850.

Deze case laat zien hoe groot soms het effect van creditrecords (correctieboekingen) kan zijn in de mondzorg. Wanneer we dan ook uitgaan van het netto aantal records en de patiënten worden gesorteerd op het bedrag dat aan de tandarts is toegekend voor deze wortelkanaalbehandelingen op één en dezelfde dag, ontstaat het volgende beeld, zie tabel 9. In deze tabel is te zien dat in 2012 acht netto records voor wortelkanaalbehandelingen op één en dezelfde dag geen uitzondering zijn. Ook vallen in dat jaar de viervouden in het aantal records op.

Verder vallen de verschillen op in eindbedragen: betaalde een patiënt in 2010 circa 800 euro, in 2012 is dat circa 1.700 euro.

2010		2011		2012	
Netto aantal records	Eindbedrag €	Netto aantal records	Eindbedrag €	Netto aantal records	Eindbedrag €
2	819,85	3	1.126,22	8	1.893,52
8	819,84	1	1.053,90	8	1.722,48
5	789,12	4	848,72	8	1.722,48
5	789,10	5	836,92	8	1.696,12
5	707,13	4	795,59	8	1.500,00
1	696,88	5	754,26	4	1.400,00
4	696,88	4	702,60	8	1.369,92
1	696,00	4	702,60	4	1.357,20
1	645,63	4	686,80	4	1.357,20
1	645,63	1	650,94	4	1.357,20

Tabel 9. Top 10 (gesorteerd op eindbedrag) van het netto aantal records over wortelkanaalbehandelingen en het eindbedrag voor wortelkanaalbehandelingen op één en dezelfde dag, 2010-2012

Ook hier weer zijn de grote verschillen te zien in de kosten die worden vergoed voor deze patiënten waarbij op een en dezelfde dag zoveel declaraties voor wortelkanaalbehandelingen plaatsvinden.

Over de medische aspecten kunnen we geen uitspraken doen. Medische experts moeten zich buigen over de (voor ons) merkwaardige viervouden in 2012 en over wat een patiënt op één en dezelfde dag aan kan wat betreft wortelkanaalbehandelingen.

3.3.7.3 Aantal extracties op één dag

Cases kunnen opmerkelijk zijn vanuit het patiëntperspectief, vanuit tandartsperspectief en vanuit declaratieperspectief. In tabel 10 wordt een overzicht gegeven van het aantal records betreffende extracties voor één patiënt op één en dezelfde dag. Zo is af te lezen dat er in 2013 een patiënt is waarvoor er in totaal 123 records waren betreffende extracties bij die persoon op die dag. Die 123 records bestaan echter uit debet- en creditrecords, daarom moet saldering plaatsvinden.

2010	2011	2012
123	104	96
102	81	79
85	66	75
63	63	72
63	62	69
63	60	63
57	58	61
57	57	60
56	57	60
55	56	60

Tabel 10. Aantal records over extracties bij één en dezelfde patiënt op één dag en dezelfde dag, 2010-2012

Wanneer we uitgaan van het netto aantal records en de patiënten worden gesorteerd op het bedrag dat aan de tandarts is toegekend voor deze extracties op één en dezelfde dag, ontstaat het volgende beeld, zie tabel 11.

Verder valt in 2010 en 2011 het grote verschil dat vergoed wordt per netto extractie-record. Die verschillen lopen uiteen van circa 50 euro tot 1.600 euro.

2010		2011		2012	
Netto aantal records	Eindbedrag €	Netto aantal records	Eindbedrag €	Netto aantal records	Eindbedrag €
24	1.291,28	29	1.472,20	48	2.712,28
6	1.229,76	28	1.446,48	36	2.091,48
30	1.158,02	1	1.425,77	31	1.727,00
32	1.152,90	26	1.246,13	24	1.704,00
23	1.132,40	24	1.239,84	1 ²⁹	1.625,00
22	1.131,37	27	1.167,50	4	1.599,96
2	1.127,28	27	1.152,02	23	1.577,80
22	1.127,28	24	1.137,76	1	1.560,00
22	1.127,28	29	1.121,00	28	1.532,00
24	1.127,28	11	1.084,86	30	1.518,00

Tabel 11. Top 10 (gesorteerd op eindbedrag) van het netto aantal records over extracties en het eindbedrag voor extracties op één en dezelfde dag, 2010-2012

Interessant is te zien dat er bij een patiënt in 2011 één record betreffende extractie is waarvoor € 1.425,77 is vergoed. Vergelijkbare cases treffen we aan in 2012 waar er één extractie record is waarvoor € 1.625,00 werd vergoed en één waarvoor € 1.560,00 werd uitgekeerd. Wellicht dat er op dit declaratierecord meerdere prestaties waren vermeld.

²⁹ Wellicht dat er meer dan één prestatie is gedeclareerd. Dit is verder niet onderzocht.

3.3.7.4 Aantal röntgenfoto's op één dag

Net als bij vullingen, wortelkanaalbehandelingen en extracties kunnen we ook hier concreet zoeken naar echt opmerkelijke cases. Die cases kunnen opmerkelijk zijn vanuit het patiëntperspectief, vanuit tandartsperspectief en vanuit declaratieperspectief. In tabel 12 wordt een overzicht gegeven van het aantal records betreffende röntgenfoto's die voor een patiënt op één en dezelfde dag zijn aangetroffen. Zo is af te lezen dat er in 2011 een patiënt is waarvoor er maar liefst 54 records waren die aangeven dat er bij die persoon op die dag röntgenfoto's zijn genomen.

2010	2011	2012
40	54	40
36	54	40
32	54	40
31	54	40
31	48	40
30	45	40
30	38	40
30	38	38
28	35	33
24	33	32

Tabel 12. Aantal records over röntgenfoto's bij één en dezelfde patiënt op één en dezelfde dag, 2010-2012

Nader deskresearch van de patiënt die bovenaan staat in 2010 leerde dat:

De patiënt met 40 records van röntgenfoto's op 6 december 2010 kreeg 191 prestaties gedeclareerd voor een totaal bedrag van € 12.562,73 waaronder 40 X10 (röntgenfoto's) en ook veel andere prestaties uit de serie Rxx, Vxx, Txx, Gxx, etc. Dit was het derde en laatste bezoek van deze patiënt in 2010 aan deze tandarts. De eerste twee consulten vielen niet echt op: elk consult had drie records waarvan de bedragen optelden tot € 383,29.

Overigens vormt het declaratiegedrag van deze tandarts voor deze patiënt een merkwaardig landschap. Er zijn erg veel correctieboekingen, hetgeen één en ander niet overzichtelijker maakt.

In hoofdstuk 2 en in de paragrafen hiervoor is al de invloed van correctieboekingen beschreven. Wanneer we dan ook uitgaan van het netto aantal records en de patiënten worden gesorteerd op het eindbedrag dat aan de tandarts is toegekend voor deze röntgenfoto's op één en dezelfde dag, ontstaat het volgende beeld, zie tabel 13. In deze tabel is te zien dat er in 2012 verzekerden zijn waarbij wel heel veel records over röntgenfoto's op één dag zijn vergoed.

Verder vallen de grote verschillen op in vergoede bedragen. Deze lopen uiteen van circa 60 euro tot 940 euro.

2010		2011		2012	
Netto aantal records	Eindbedrag €	Netto aantal records	Eindbedrag €	Netto aantal records	Eindbedrag €
3	1.291,50	4	861,97	4	2.008,01
1	861,00	1	680,09	2	1.805,74
1	861,00	2	673,39	12	1.074,56
1	861,00	12	607,52	3	983,00
3	633,70	1	557,91	1	940,38
1	614,90	3	557,41	1	937,88
4	582,11	3	535,20	4	911,40
1	516,60	1	500,00	15	900,00
1	516,60	1	459,20	2	895,00
1	516,60	6	455,64	8	815,72

Tabel 13. Top 10 (gesorteerd op eindbedrag) van het netto aantal records over röntgenfoto's en het eindbedrag voor röntgenfoto's op één en dezelfde dag, 2010-2012

Ook hier weer zijn de grote verschillen te zien in de kosten die worden vergoed voor deze patiënten waarbij op een en dezelfde dag zoveel records voor röntgenfoto's worden aangetroffen.

Over de medische aspecten kunnen we geen uitspraken doen: zo is de duiding van de persoon die op één dag 12 foto's kreeg gedeclareerd en een ander die 15 foto's kreeg gedeclareerd moeilijk (zie 2012).

3.3.7.5 Aantal sealings op één dag

Dezelfde werkwijze zoals beschreven in de vorige paragrafen kan worden gehanteerd voor sealings/sealen. Er is gekeken hoe vaak het voorkwam dat bij een en dezelfde persoon op een en dezelfde dag declaratierecords met betrekking tot sealen voorkwamen. We geven eerst een overzicht van het totale aantal records, daarna geven we een overzicht van de gesaldeerde records.

2010	2011	2012
90	54	120
60	54	80
57	54	65
55	51	64
53	48	64
48	48	64
48	48	64
48	48	64
48	48	64
45	48	64

Tabel 14. Aantal records over sealings bij één en dezelfde patiënt op één en dezelfde dag, 2010-2012

Opvallend in deze tabel is niet alleen het grote aantal records dat bij sommige patiënten wordt aangetroffen (op één en dezelfde dag) maar ook het opmerkelijk hoge aantal achtvouden.

Onder deze records bevinden zich behalve debetrecords ook creditrecords. Wanneer we uitgaan van het netto aantal records (zie paragraaf 2.1) ontstaat het overzicht dat gegeven wordt in tabel 15.

2010		2011		2012	
Netto aantal records	Eindbedrag €	Netto aantal records	Eindbedrag €	Netto aantal records	Eindbedrag €
57	740,42	43	565,89	64	1152,00
16	548,27	33	436,69	64	1152,00
35	458,60	33	436,69	64	1088,00
32	450,92	32	434,10	60	1077,00
32	430,42	32	434,10	60	1050,00
32	430,42	16	410,85	64	1024,00
32	430,42	30	408,26	64	1024,00
2 ³⁰	417,61	28	403,08	56	1008,00
31	407,36	30	391,11	60	1005,00
31	407,36	3	385,01	64	1004,80

Tabel 15. Top 10 (gesorteerd op eindbedrag) van het netto aantal records over sealings en het eindbedrag voor sealings voor deze patiënt op één en dezelfde dag, 2010-2012

Los van de soms grote verschillen die te zien zijn in de kosten per record van sealings die er voor deze patiënten zijn, vallen de achtvouden op in 2012. Ook in 2010 en 2011 zijn die zichtbaar maar in mindere mate in deze top 10.

³⁰ Wellicht zijn er meer dan één prestatie in deze twee records gedeclareerd.

3.3.7.6 Behandelingen bij jonge mensen

Een onderwerp waar domeinexperts beslist naar moeten kijken, zijn enkele bijzondere verrichtingen bij jonge mensen. De resultaten van deze paragraaf (4 tot 5 miljoen euro vergoed per jaar) zijn niet meegenomen in de totaalschatting van te veel vergoede bedragen bij mondzorg omdat medische experts nog naar de cases en declaraties moeten kijken. Vooral de orale implantologie bij jongeren valt op.

Behandeling bij jongeren < 17 jaar	2010 Aantal	2010 Bedrag in miljoen euro	2011 Aantal	2011 Bedrag in miljoen euro	2012 Aantal	2012 Bedrag in miljoen euro
Wortelkanaal	15.467	2,2	14.619	2,1	14.065	3,3
Kroon	2.548	0,3	3.770	0,4	5.550	0,6
Prothese	632	0,1	641	0,1	17 ³¹	0,0
Orale implant	398.388	2,4	406.193	2,3	511 ³²	0,1
Totaal		5,0		4,9		4,0

Tabel 16. Bijzondere verrichtingen bij mensen jonger dan 17 jaar, 2011-2012

Wanneer we verder inzoomen op deze groep en ons richten op de kinderen jonger dan 10 jaar dan vinden we als totaal bedragen voor 2010: 1,5 miljoen euro, voor 2011: 1,5 miljoen euro en voor 2012 tenslotte: 0,9 miljoen euro.

³¹ Wellicht wordt hier een deel gemist door afwijkende codes in 2012.

³² Wellicht wordt hier een deel gemist door afwijkende codes in 2012.

3.3.8 Schatting van het bedrag gemoeid met het overtreden van regels en anomalieën

Zoals bekend is er een onderscheid gemaakt tussen het overtreden van declaratieregels enerzijds en het vaststellen van merkwaardige cases anderzijds. De eerste categorie zijn harde regels, bij de tweede categorie moeten experts zich buigen over de duiding van de merkwaardige cases.

3.3.8.1 Overtreding van declaratieregels

De berekening van het te veel vergoede bedrag bij het overtreden van de declaratieregels is relatief eenvoudig. Voor het vaststellen of er sprake is van een overtreding, is “soepel” beoordeeld, dat wil zeggen dat de criteria zijn verzacht om twijfel uit te sluiten.

Dit betekent dat de bedragen in de onderstaande tabel relatief hard zijn.

Toets	Te veel vergoede bedrag 2010 in miljoen	Te veel vergoede bedrag 2011 in miljoen	Te veel vergoede bedrag 2012 in miljoen
Ongeldige combinaties van prestatiecodes C11, C12, C12, C13, C22, C28, C29, M50, M55, M59	0,1	0,1	nvt
Ongeldige combinaties van prestatiecodes A111, A211, A231, A261, C112, C114	nvt	nvt	1,9
Dubbele facturen	0,6	0,6	0,6
Individuele case (3.3.2)	0,1	nvt	nvt
Totaal	0,8	0,7	2,5
Totaal vergoed	2.179	2.180	2.179
Als percentage van het totale bedrag vergoed	0,04	0,03	0,11

Tabel 17. Overzicht van de te veel vergoede bedragen in 2010, 2011 en 2012

3.3.8.2 Anomalieën

Sommige meta-variabelen zoals de verhouding tussen bepaalde verrichtingen en het totale aantal verrichtingen kunnen goed helpen om opmerkelijke situaties (lees: tandartsen die wel erg afwijkend zijn) op te sporen. Bij deze anomalie detectie hebben we ons vooral gericht op enkele specifieke onderdelen namelijk gebitsreiniging, vullingen, wortelkanaalbehandelingen, extracties, röntgenfoto's en sealings (zie ook 3.3.7 en Appendix 3).

De werkwijze is als volgt (zie ook paragraaf 1.4): Voor elke AGB berekenen we de gemiddelde kosten van een verrichting (bijvoorbeeld een vulling) voor alle patiënten die deze verrichting (vulling) bij hem/haar hebben gekregen. Dit levert een verdeling op van de kosten van deze verrichtingen van alle tandartsen. Daarin kunnen we relatief goedkope en "gemiddelde" tandartsen onderscheiden. Vervolgens wordt de rechterstaart van deze verdeling onderzocht: hierin bevinden zich de relatief erg dure tandartsen.

De resultaten worden in tabel 18 - 20 gegeven. Een uitleg over de berekeningen en de referentiepunten 95%, 99% en 99,9% is in paragraaf 1.4 aangereikt.

Mondzorg 2010	Schatting van het te veel vergoede bedrag, in miljoen euro		
Variabele	95% niveau	99% niveau	99,9% niveau
Gebitsreiniging	0,9	0,1	0,0
Vullingen	2,0	0,3	0,0
Wortelkanaal	0,2	0,1	0,0
Extracties	0,3	0,1	0,0
Röntgenfoto's	1,1	0,3	0,0
Sealings	0,3	0,0	0,0
Totaal bedrag	4,8	0,9	0,1

Tabel 18. Overzicht van de te veel vergoede bedragen bij anomalieën, 2010

Mondzorg 2011	Schatting van het te veel vergoede bedrag, in miljoen euro		
Variabele	95% niveau	99% niveau	99,9% niveau
Gebitsreiniging	0,9	0,2	0,0
Vullingen	2,2	0,4	0,0
Wortelkanaal	0,2	0,0	0,0
Extracties	0,3	0,1	0,0
Röntgenfoto's	1,0	0,3	0,0
Sealings	0,2	0,0	0,0
Totaal bedrag	4,8	1,0	0,1

Tabel 19. Overzicht van de te veel vergoede bedragen bij anomalieën, 2011

Mondzorg 2012	Schatting van het te veel vergoede bedrag, in miljoen euro		
Variabele	95% niveau	99% niveau	99,9% niveau
Gebitsreiniging	0,4	0,0	0,0
Vullingen	3,0	0,7	0,1
Wortelkanaal	0,5	0,1	0,0
Extracties	0,4	0,1	0,0
Röntgenfoto's	1,0	0,3	0,0
Sealings	0,3	0,0	0,0
Totaal bedrag	5,6	1,2	0,1

Tabel 20. Overzicht van de te veel vergoede bedragen bij anomalieën, 2012

3.3.9 Overzicht mondzorg

Deze paragraaf brengt alle deelresultaten samen voor het mondzorgsegment. De resultaten spreken voor zich. Wij achten het referentiepunt van 99% het meest reëel, zie hiervoor tabel 22.

Jaartal	Totaal bedrag vergoed in miljoenen euro A	Te veel vergoed op basis van overtreding declaratieregels B	Te veel vergoed op basis van anomalieën (95%) C	Totaal te veel vergoed (95%) D = B+C	Totaal te veel vergoed als percentage van totaal vergoed E = D/A
2010	2.179	0,8	4,8	5,6	0,3
2011	2.180	0,7	4,8	5,5	0,3
2012	2.179	2,5	5,6	8,1	0,4

Tabel 21. Overzicht van de totalen te veel vergoede bedragen, 2010 – 2012, 95% referentiepunt

Jaartal	Totaal bedrag vergoed in miljoenen euro A	Te veel vergoed op basis van overtreding declaratieregels B	Te veel vergoed op basis van anomalieën (99%) C	Totaal te veel vergoed (99%) D = B+C	Totaal te veel vergoed als percentage van totaal vergoed E = D/A
2010	2.179	0,8	0,9	1,7	0,1
2011	2.180	0,7	1,0	1,7	0,1
2012	2.179	2,5	1,2	3,7	0,2

Tabel 22. Overzicht van de totalen te veel vergoede bedragen, 2010 – 2012, 99% referentiepunt

Jaartal	Totaal bedrag vergoed in miljoenen euro A	Te veel vergoed op basis van overtreding declaratieregels B	Te veel vergoed op basis van anomalieën (99,9%) C	Totaal te veel vergoed (99,9%) D = B+C	Totaal te veel vergoed als percentage van totaal vergoed E = D/A
2010	2.179	0,8	0,1	0,9	0,0
2011	2.180	0,7	0,1	0,8	0,0
2012	2.179	2,5	0,1	2,6	0,1

Tabel 23. Overzicht van de totalen te veel vergoede bedragen, 2010 – 2012, 99,9% referentiepunt

3.3.10 Specifieke analyse voor de basisverzekering

De analyses voor de mondzorg zoals die hiervoor beschreven zijn, zijn eveneens uitgevoerd op twee deelverzamelingen: de basisverzekering en de jeugd binnen de basisverzekering. Het betreft eveneens de jaren 2010 – 2012. Ook bij deze deelanalyses is rekening gehouden met de correctieboekingen.

Voor de basisverzekering wordt in tabel 24 het overzicht over de overtreding van declaratieregels gegeven. Een vergelijking van tabel 24 met tabel 17 laat zien dat het percentage van het bedrag dat ontstaat door de overtreding van declaratieregels nagenoeg gelijk is.

Ook is duidelijk dat de overtreding van de declaratieregels wat betreft niet toegestane combinaties met name in de basisverzekering hebben plaatsgevonden, en zoals we later zullen zien: met name bij de jeugd in de basisverzekering.

Toets	Te veel vergoede bedrag 2010 in miljoen	Te veel vergoede bedrag 2011 in miljoen	Te veel vergoede bedrag 2012 in miljoen
Ongeldige combinaties van prestatiecodes C11, C12, C12, C13, C22, C28, C29, M50, M55, M59	0,1	0,1	nvt
Ongeldige combinaties van prestatiecodes A111, A211, A231, A261, C112, C114	nvt	nvt	0,4
Dubbele facturen	0,2	0,2	0,2
Individuele case (3.3.2)	0,1		
Totaal	0,4	0,3	0,6
Totaal vergoed	734	672	686
Als percentage van het totale bedrag vergoed	0,05	0,04	0,09

Tabel 24. Overzicht van de te veel vergoede bedragen door overtreding declaratieregels, basisverzekering, 2010 – 2012

Sommige meta-variabelen zoals de verhouding tussen bepaalde verrichtingen en het totale aantal verrichtingen kunnen goed helpen om opmerkelijke situaties (lees: tandartsen die wel erg afwijkend zijn) op te sporen. Bij deze anomalie detectie hebben we ons vooral gericht op enkele specifieke onderdelen namelijk gebitsreiniging, vullingen, wortelkanaalbehandelingen, extracties en röntgenfoto's (zie ook 3.3.7 en Appendix 3). De resultaten worden in tabel 25 – 27 gegeven. Een uitleg over de referentiepunten 95%, 99% en 99,9% is in paragraaf 1.4 aangereikt.

Mondzorg 2010	Schatting van het te veel vergoede bedrag, in miljoen euro		
Variabele	95% niveau	99% niveau	99,9% niveau
Gebitsreiniging	0,2	0,1	0,0
Vullingen	0,6	0,1	0,0
Wortelkanaal	0,1	0,0	0,0
Extracties	0,1	0,0	0,0
Röntgenfoto's	0,4	0,0	0,0
Sealings	0,3	0,0	0,0
Totaal bedrag	1,7	0,2	0,0

Tabel 25. Overzicht van de te veel vergoede bedragen bij anomalieën basisverzekering, 2010

Mondzorg 2011	Schatting van het te veel vergoede bedrag, in miljoen euro		
Variabele	95% niveau	99% niveau	99,9% niveau
Gebitsreiniging	0,2	0,0	0,0
Vullingen	0,4	0,1	0,0
Wortelkanaal	0,0	0,0	0,0
Extracties	0,1	0,0	0,0
Röntgenfoto's	0,2	0,0	0,0
Sealings	0,2	0,0	0,0
Totaal bedrag	1,1	0,1	0,0

Tabel 26. Overzicht van de te veel vergoede bedragen bij anomalieën basisverzekering, 2011

Mondzorg 2012	Schatting van het te veel vergoede bedrag, in miljoen euro		
Variabele	95% niveau	99% niveau	99,9% niveau
Gebitsreiniging	0,2	0,0	0,0
Vullingen	0,9	0,2	0,0
Wortelkanaal	0,0	0,0	0,0
Extracties	0,1	0,0	0,0
Röntgenfoto's	0,4	0,0	0,0
Sealings	0,3	0,0	0,0
Totaal bedrag	1,9	0,2	0,0

Tabel 27. Overzicht van de te veel vergoede bedragen bij anomalieën basisverzekering, 2012

De resultaten spreken voor zich. Wij gaan uit van het referentiepunt van 99%, zie hiervoor tabel 29.

Jaartal	Totaal bedrag vergoed in miljoenen euro A	Te veel vergoed op basis van overtreding declaratieregels B	Te veel vergoed op basis van anomalieën (95%) C	Totaal te veel vergoed (95%) D = B+C	Totaal te veel vergoed als percentage van totaal vergoed E = D/A
2010	734	0,4	1,7	2,1	0,3
2011	672	0,3	1,1	1,4	0,2
2012	686	0,6	1,9	2,5	0,4

Tabel 28. Overzicht van de totalen te veel vergoede bedragen, basisverzekering, 2010 – 2012, 95% referentiepunt

Jaartal	Totaal bedrag vergoed in miljoenen euro A	Te veel vergoed op basis van overtreding declaratieregels B	Te veel vergoed op basis van anomalieën (99%) C	Totaal te veel vergoed (99%) D = B+C	Totaal te veel vergoed als percentage van totaal vergoed E = D/A
2010	734	0,4	0,2	0,6	0,1
2011	672	0,3	0,1	0,4	0,1
2012	686	0,6	0,2	0,8	0,1

Tabel 29. Overzicht van de totalen te veel vergoede bedragen, basisverzekering, 2010 – 2012, 99% referentiepunt

Jaartal	Totaal bedrag vergoed in miljoenen euro A	Te veel vergoed op basis van overtreding declaratieregels B	Te veel vergoed op basis van anomalieën (99,9%) C	Totaal te veel vergoed (99,9%) D = B+C	Totaal te veel vergoed als percentage van totaal vergoed E = D/A
2010	734	0,4	0,0	0,4	0,1
2011	672	0,3	0,0	0,3	0,04
2012	686	0,6	0,0	0,6	0,1

Tabel 30. Overzicht van de totalen te veel vergoede bedragen, basisverzekering, 2010 – 2012, 99,9% referentiepunt

3.3.11 Specifieke analyse voor de jeugd in de basisverzekering

Voor de jeugd in basisverzekering (dat wil zeggen 17 jaar of jonger) wordt in tabel 31 het overzicht over de overtreding van declaratieregels gegeven terwijl tabel 34 het totaaloverzicht weergeeft (99%-referentiepunt).

Wanneer we dit vergelijken met de resultaten van de gehele mondzorg dan constateren we dat het in grote lijnen overeenkomt. Zoals eerder beschreven vragen de opmerkelijke verrichtingen bij jonge mensen (zie de betreffende paragraaf hierover) bijzondere aandacht.

Toets	Te veel vergoede bedrag 2010 in miljoen	Te veel vergoede bedrag 2011 in miljoen	Te veel vergoede bedrag 2012 in miljoen
Ongeldige combinaties van prestatiecodes C11, C12, C12, C13, C22, C28, C29, M50, M55, M59	0,1	0,1	nvt
Ongeldige combinaties van prestatiecodes A111, A211, A231, A261, C112, C114	nvt	nvt	0,4
Dubbele facturen	0,0	0,0	0,0
Totaal	0,1	0,1	0,4
Totaal vergoed	276	291	320
Als percentage van het totale bedrag vergoed	0,04	0,03	0,13

Tabel 31. Overzicht van de te veel vergoede bedragen in jeugd basisverzekering, 2010, 2011 en 2012

Sommige meta-variabelen zoals de verhouding tussen bepaalde verrichtingen en het totale aantal verrichtingen kunnen goed helpen om opmerkelijke situaties (lees: tandartsen die wel erg afwijkend zijn) op te sporen. Bij deze anomalie detectie hebben we ons vooral gericht op enkele specifieke onderdelen namelijk gebitsreiniging, vullingen, wortelkanaalbehandelingen, extracties en röntgenfoto's (zie ook 3.3.7). De resultaten worden in tabel 32 – 34 gegeven. Een uitleg over de referentiepunten 95%, 99% en 99,9% is in paragraaf 1.4 aangereikt.

Mondzorg 2010	Schatting van het te veel vergoede bedrag, in miljoen euro		
Variabele	95% niveau	99% niveau	99,9% niveau
Gebitsreiniging	0,1	0,0	0,0
Vullingen	0,4	0,1	0,0
Wortelkanaal	0,0	0,0	0,0
Extracties	0,1	0,0	0,0
Röntgenfoto's	0,1	0,0	0,0
Sealings	0,2	0,0	0,0
Totaal bedrag	0,9	0,1	0,0

Tabel 32. Overzicht van de te veel vergoede bedragen bij anomalieën in jeugd basisverzekering, 2010

Mondzorg 2011	Schatting van het te veel vergoede bedrag, in miljoen euro		
Variabele	95% niveau	99% niveau	99,9% niveau
Gebitsreiniging	0,2	0,0	0,0
Vullingen	0,4	0,1	0,0
Wortelkanaal	0,0	0,0	0,0
Extracties	0,0	0,0	0,0
Röntgenfoto's	0,1	0,0	0,0
Sealings	0,2	0,0	0,0
Totaal bedrag	0,9	0,1	0,0

Tabel 33. Overzicht van de te veel vergoede bedragen bij anomalieën in jeugd basisverzekering, 2011

Mondzorg 2012	Schatting van het te veel vergoede bedrag, in miljoen euro		
Variabele	95% niveau	99% niveau	99,9% niveau
Gebitsreiniging	0,2	0,0	0,0
Vullingen	0,7	0,2	0,0
Wortelkanaal	0,0	0,0	0,0
Extracties	0,1	0,0	0,0
Röntgenfoto's	0,1	0,0	0,0
Sealings	0,3	0,0	0,0
Totaal bedrag	1,4	0,2	0,0

Tabel 34. Overzicht van de te veel vergoede bedragen bij anomalieën in jeugd basisverzekering, 2012

De resultaten spreken voor zich. Wij achten het referentiepunt van 99% het meest reëel, zie hiervoor tabel 36.

Jaartal	Totaal bedrag vergoed in miljoenen euro A	Te veel vergoed op basis van overtreding declaratieregels B	Te veel vergoed op basis van anomalieën (95%) C	Totaal te veel vergoed (95%) D = B+C	Totaal te veel vergoed als percentage van totaal vergoed E = D/A
2010	276	0,1	0,9	1,0	0,4
2011	291	0,1	0,9	1,0	0,3
2012	320	0,4	1,4	1,8	0,6

Tabel 35. Overzicht van de totalen te veel vergoede bedragen in jeugd basisverzekering, 2010 – 2012, 95% referentiepunt

Jaartal	Totaal bedrag vergoed in miljoenen euro A	Te veel vergoed op basis van overtreding declaratieregels B	Te veel vergoed op basis van anomalieën (99%) C	Totaal te veel vergoed (99%) D = B+C	Totaal te veel vergoed als percentage van totaal vergoed E = D/A
2010	276	0,1	0,1	0,2	0,1
2011	291	0,1	0,1	0,2	0,1
2012	320	0,4	0,2	0,6	0,2

Tabel 36. Overzicht van de totalen te veel vergoede bedragen in jeugd basisverzekering, 2010 – 2012, 99% referentiepunt

Jaartal	Totaal bedrag vergoed in miljoenen euro A	Te veel vergoed op basis van overtreding declaratieregels B	Te veel vergoed op basis van anomalieën (99,9%) C	Totaal te veel vergoed (99,9%) D = B+C	Totaal te veel vergoed als percentage van totaal vergoed E = D/A
2010	276	0,1	0,0	0,1	0,04
2011	291	0,1	0,0	0,1	0,03
2012	320	0,4	0,0	0,4	0,1

Tabel 37. Overzicht van de totalen te veel vergoede bedragen in jeugd basisverzekering, 2010 – 2012, 99,9% referentiepunt

Tabel 38 geeft een overzicht van het percentage te veel vergoed ten opzichte van het totaal toegekende bedrag voor de jaren 2010 – 2012 uitgaande van het 99%-referentiepunt (zie de tabellen 22, 29 en 36). Uit die tabellen is ook af te lezen dat het bedrag gemoeid met anomalieën in de mondzorg als totaal in absolute zin hoger is dan in de basisverzekering (Mondzorgtotaal: 0,9, 1,0 en 1,2 miljoen; basisverzekering 0,2, 0,1 en 0,2 miljoen). Bij de jeugd in de basisverzekeringen bedragen de anomalieën respectievelijk 0,1, 0,1 en 0,2 miljoen euro.

Uit onderstaande tabel blijkt dat de relatieve verschillen klein zijn.

Jaartal	Percentage te veel vergoed (99%) Mondzorg totaal	Percentage te veel vergoed (99%) Basisverzekering	Percentage te veel vergoed (99%) Jeugd in basisverzekering
2010	0,1	0,1	0,1
2011	0,1	0,1	0,1
2012	0,2	0,1	0,2

Tabel 38. Overzicht van de percentages totaal te veel vergoede bedragen bij referentiepunt 99% voor de mondzorg totaal, basisverzekering en jeugd in basisverzekering, 2010 – 2012

3.4 Resultaten voor farmacie

3.4.1 De beschikbare data

Een overzicht van de beschikbare variabelen wordt in Appendix 6 gegeven.

Er is een stapsgewijze aanpak: eerst een beschrijving van de data gevolgd door de resultaten van een exploratieve analyse, vervolgens de toetsen op overtreding van declaratieregels en daarna anomalie detectie. De paragraaf wordt afgesloten met een overzicht van te veel vergoede bedragen.

Het gaat om declaratierecords dat wil zeggen dat deze declaraties vergoed zijn. In een klein aantal gevallen worden negatieve bedragen aangetroffen, dit zijn correctieboekingen. Door te aggregeren kan inzicht per verzekerde worden verkregen, net als inzicht per apotheek, per praktijk, per voorschrijver, per artikel, per kwartaal, per jaar, etc., etc.

Voor het begrijpen van het data bestand farmacie is het belangrijk om een onderscheid te maken tussen AGB-voorschrijver, AGB-afleveraar en AGB-praktijk. De laatste twee begrippen lopen soms door elkaar. De analyses zijn voor elk van de drie categorieën uitgevoerd. De AGB-afleveraar is de AGB-code van degene onder wiens verantwoordelijkheid de aflevering heeft plaatsgevonden. Dit is niet noodzakelijkerwijze degene die de vergoeding ontvangt.

Het aantal records voor dit onderzoek is groot, in totaal 619 miljoen. Het betreft vergoedingen voor een totaal van 15,5 miljard euro in de jaren 2010 – 2012.

De verdeling van het aantal records over deze jaren wordt in tabel 1 weergegeven. Het totale bedrag dat de afgelopen jaren is vergoed is gestaag dalende. De stijgende aantallen in de kolom AGB-praktijken rechtvaardigen niet de conclusie dat het aantal praktijken is gestegen: foute codes in het betreffende data veld kunnen een goede verklaring hiervoor zijn.

Jaartal	Aantal records	Totaal bedrag vergoed in miljoenen euro	Aantal AGB-afleveraars	Aantal AGB-voorschrijvers	Aantal AGB-praktijken
2010	191.744.461	5.531	6.790	45.466	5.942
2011	208.515.505	5.280	6.909	48.138	5.782
2012	219.200.187	4.669	6.779	51.813	8.267
Totaal	619.460.153	15.480			

Tabel 1. Aantal records betreffende vergoedingen farmacie in de jaren 2010 – 2012

Voor een aantal key-variabelen zijn er veel onbekende en ontbrekende waarden. Bijvoorbeeld voor de variabele AGB-afleveraar zijn er in 2010 1,1 miljoen records met onbekenden. Het wordt niet beter in 2011 en 2012: 1,9 miljoen records³³. Voor AGB-voorschrijvers is dit aantal onbekenden nog groter, 2010: 10,2 miljoen, 2011: 11,6 miljoen en 2012: 10,1 miljoen³⁴. De code AGB-praktijk ontbreekt in 2010 bij ongeveer 36 miljoen records³⁵.

³³ Met name voor in het buitenland geleverde zorg is de AGB-code onbekend.

³⁴ Dat de AGB code van de voorschrijver relatief vaak niet bekend is, is niet zo vreemd: een apotheek heeft veelal alleen maar de AGB code van veel voorkomende voorschrijvers in het systeem staan. Elektronisch voorschrijven waarbij de AGB code van de voorschrijver (en de instelling waarvan uit de voorschrijver voorschrijft) als wettelijk verplichte velden zou hier verbetering in kunnen aanbrengen.

³⁵ De AGB code praktijk wordt niet altijd direct door de indiener van de declaratie gevuld.

Ook in dit segment komen veel correctierecords voor. Zie hiervoor ook hoofdstuk 2. Daar is onderstaande tabel 2 opgenomen die een overzicht geeft van de correctierecords voor de jaren 2010 - 2012.

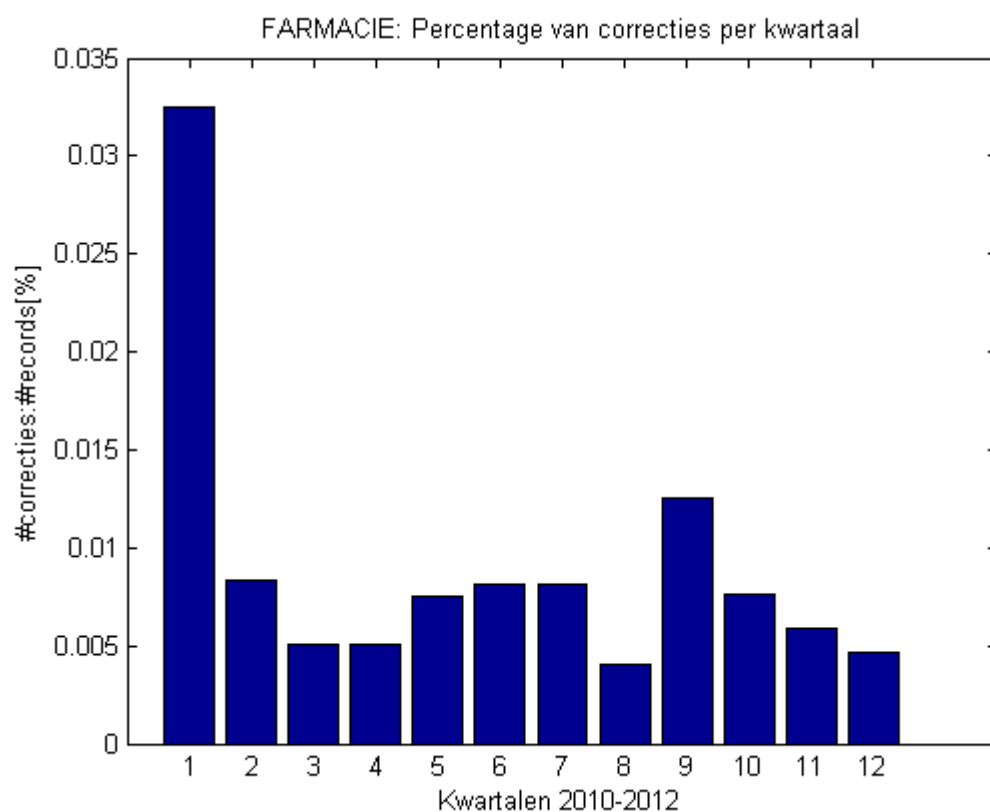
De wijze waarop correcties in de bestanden van farmacie verwerkt zijn, verschilt van de andere zorgsegmenten. Bij farmacie heeft Vektis de creditrecords met de corresponderende debetrecords gesaldeerd. Wanneer het resulterende vergoedingsbedrag gelijk was aan nul is het record vervallen. Dit is de verklaring waarom het aantal correctierecords bij farmacie zo laag is.

Hoewel nu relatief klein in aantal, zorgen de resterende correctierecords voor een neerwaartse correctie van het totaal van de ingediende declaraties in het segment farmacie van circa 1 miljoen euro per jaar.

Jaartal	Farmacie				
	Aantal correctie records	Netto aantal records	Totale bedrag correcties in miljoen euro	Percentage van de correcties in totaal aantal records	Percentage van het bedrag van de correcties in het totaal gedeclareerd bedrag
2010	23.727	191.636.481	1,1	0,0	0,0
2011	14.427	208.431.051	0,8	0,0	0,0
2012	16.634	219.113.995	0,7	0,0	0,0

Tabel 2. Correctieboekingen in het segment farmacie 2010-2012

In tegenstelling tot huisartsen en mondzorg, laat de verdeling van het percentage correctieboekingen in de farmacie een ander beeld zien, zie hiervoor figuur 1. Opvallend is het eerste kwartaal van 2010, en daarna in mindere mate, het eerste kwartaal van 2012.



Figuur 1. Overzicht van het percentage correctieboekingen in de loop van de tijd, 2010 – 2012

3.4.2 Overzicht van frequentie van declaraties bij farmacie

Een illustratief voorbeeld is de frequentie waarmee verzekerden voorkomen in het bestand van de apothekers. We geven de top 9 van deze verzekerden met het aantal netto records waarmee we ze tegen kwamen in het farmaciebestand³⁶. Zie tabel 3. De frequentie is veel hoger dan bij huisartsen, mondzorg en GGZ. Zo is te zien dat er één verzekerde is die in 2010 voor maar liefst 3.890 records bij de apotheek zorgde. Nader onderzoek leert dat de top 9 van twee opeenvolgende jaren een grote overlap kent, zie tabel 4.

2010	2011	2012
3.890	3.182	2.977
3.289	3.070	2.687
2.773	2.358	2.503
2.630	2.188	2.163
2.489	2.063	2.149
2.401	1.987	1.812
2.380	1.940	1.795
2.332	1.902	1.790
2.297	1.859	1.741

Tabel 3. Top 9 van de frequentie van verzekerden, farmacie, jaren 2010 – 2012

2010	2011	2012
3.890	3.182	2.977
3.289	3.070	2.687
2.773	2.358	2.503
2.630	2.188	2.163
2.489	2.063	2.149
2.401	1.987	1.812
2.380	1.940	1.795
2.332	1.902	1.790
2.297	1.859	1.741

Tabel 4. Top 9 van de frequentie van verzekerden, farmacie, jaren 2010 – 2012; de kleuren geven dezelfde verzekerde aan.

³⁶ Ook hier komen ontbrekende BSN voor:, in 2010 15.578, in 2011 13.532 en in 2012 7.978.

De twee nummer 1 cases in 2010 en 2011 zijn aan een expert voorgelegd. Hierbij is onder andere een nadere analyse gemaakt voor het eerste kwartaal van 2010. De expert³⁷ reageert als volgt:

“Opmerkelijke casus en voor een apotheek ook zeer discutabel; er is mogelijk sprake van:
- misbruik van iemands BSN door op deze manier verslaafden van een portie te voorzien;
- declaratiefraude door op 1 persoon veel te veel geneesmiddelen te boeken;
- een handel van de betreffende patiënt en apotheker in de geneesmiddelen zoals in de bijlage vernoemd.

Voor een specificatie zie hieronder:

Artikelcodes, de frequentie, merknaam en omschrijving voor deze patiënt in Q1 2010:

14691728	83	paroxetine	antidepr
13920626	63	oxazepam	kalmeringsmiddel
12138592	61	orap	antipsychotica
13920782	61	oxazepam	kalmeringsmiddel
14170183	61	carbasalaatcalcium cardio	antistollingsmiddel
14206994	61	Micardis	bloeddrukverlagers
14691604	61	Metformine	bloedglucoseverlagend
14696053	61	Detrusitol	overactieve blaassyndroom
14760606	61	Buspiron	kalmeringsmiddel
14850982	61	Simvastatine	cholesterolverlagers
14969424	61	Omeprazol	maagbeschermer
15094448	61	Glimepiride	Bloedglucoseverlagend
15346943	61	Eucreas	Bloedglucoseverlagend
15078884	60	Glimepiride	Bloedglucoseverlagend
15279073	41	Ferrogard	IJzertabletten
13889176	28	Dormicum	kalmeringsmiddel
13889184	11	Dormicum	kalmeringsmiddel
14689871	11	Zopiclon	Kalmeringsmiddel
13920456	9	Chloordiazepoxide	Kalmeringsmiddel
14603829	9	Oxycontin	Pijnstiller, opiaat
14964473	7	Lyrca	Pijn, angst, epilepsie
13908200	6	Paracetamol/Codeine	Pijnstiller
14784254	6	Oxycontin	Pijnstiller, opiaat
14838273	5	Oxynorm	Pijnstiller, opiaat
13813161	2	Arthrotec	Ontstekingsremmer
14278685	2	Amoxicilline	Antibioticum
14577674	2	Codeinefosfaat	Demping bij hoest, (of bij pijn)
14649209	2	Diclofenac	Ontstekingsremmer
14700875	2	Tramadol	Pijnstiller, opiaat
12564176	1	Dormicum	kalmeringsmiddel
13722255	1	Primperan	Misselijkheid
13730355	1	Legendal (lactulose)	Obstipatie
14127202	1	Ventolin	Luchtwegverwijder (astma)
14603837	1	Oxycontin	Opiaat, pijnstiller
14694964	1	Amoxicilline/clavulaanzuur	Antibioticum
14918714	1	Microlax	Obstipatie
15254194	1	aerius	bij hooikoorts bijv.
15265722	1	?	
15290808	1	Nitrofurantoin	bij blaasontsteking,
15458083	1	Foster	Luchtwegverwijder en ontstekingsremmer bij astma o.a.

³⁷ Een andere, onafhankelijke expert kwam na lezing van de reactie van deze expert tot een mogelijke verklaring van dit patroon waarbij hij wees op de mogelijkheid van een patiënt met suicide risico's die slechts een beperkte hoeveelheid per dag meekrijgt of onder toezicht van de apotheker zijn medicijnen in moet nemen. Dit voorbeeld laat goed zien hoe groot het belang is dat experts naar gevonden merkwaardige cases kijken.

3.4.3 Overzicht van declaratiebedragen

Op basis van de databestanden kan een goed beeld worden geschetst van de verschillen tussen de jaren. Zo is uit tabel 5 te lezen dat het totale bedrag aan farmacie in deze drie jaren gestaag daalt. We zien verder dat het gemiddelde bedrag per declaratie flink daalt. Ook is te zien dat de verdeling van de declaraties minder scheef is geworden: het verschil tussen het rekenkundig gemiddelde en de mediaan is gedaald van 18 in 2010 naar 15 in 2012.

Jaartal	Totaal bedrag toegekend in miljoenen euro	Gemiddeld toegekend bedrag van een declaratie in euro	Mediaan van het toegekende bedrag van een declaratie in euro
2010	5.531	28,85	10,89
2011	5.280	25,32	9,24
2012	4.669	23,30	8,05

Tabel 5. Overzicht van de totale vergoeding, gemiddelde en mediaan, farmacie, 2010 – 2012

3.4.4 Extreem hoge gemiddeld bedrag per declaratie

Voor elke declaratie is bekend hoeveel de zorgverzekeraar heeft vergoed. Wanneer we per apotheek kijken naar het gemiddelde bedrag per declaratie dan zien we grote verschillen. Sommige van die verschillen zijn toe te schrijven aan lokale omstandigheden, aan grootte van de praktijk of aan specialisatie. Maar soms is het gemiddelde bedrag per declaratie wel erg hoog. Zie hiervoor tabel 6.

Gemiddeld bedrag per declaratie 2010 €	Gemiddeld bedrag per declaratie 2011 €	Gemiddeld bedrag per declaratie 2012 €
2.627	3.159 ³⁸	3.357
2.611	2.345	2.537
2.299	2.327 ³⁹	2.294
2.218	2.303	2.222
2.013	2.083	1.904
1.918	2.065	1.853
1.498	2.021	1.753
1.467	1.958	1.656
1.433	1.953	1.537
482 ⁴⁰	1.595	1.300

Tabel 6. Top 10 van het gemiddelde bedrag per declaratie voor praktijken (met meer dan 40 patiënten waarvoor gedeclareerd is), farmacie, 2010 – 2012.

³⁸ Deze praktijk staat in 2010 en 2012 op nummer 1 van de top10 maar heeft daar slechts 36 en 39 patiënten waarvoor vergoed is. In 2011 gaat het om 46 patiënten. Het is een merkwaardige case waar nader gekeken moet worden. Wellicht gaat het om een gespecialiseerde apotheek/afdeling/samenwerkingsverband.

³⁹ In 2012 op de 11-de plaats.

⁴⁰ Het opvallend grote verschil tussen nummer 9 en nummer 10 op de lijst van 2010 is een illustratie van hoe groot de scheefheid van de verdeling is. Specialisatie zal hier een grote rol kunnen spelen.

Gemiddeld bedrag per declaratie 2010 €	Gemiddeld bedrag per declaratie 2011 €	Gemiddeld bedrag per declaratie 2012 €
2.627	3.159	3.357
2.611	2.345	2.537
2.299	2.327	2.294
2.218	2.303	2.222
2.013	2.083	1.904
1.918	2.065	1.853
1.498	2.021	1.753
1.467	1.958	1.656
1.433	1.953	1.537
482	1.595	1.300

Tabel 7. Top 10 van het gemiddelde bedrag per declaratie voor praktijken (met meer dan 40 patiënten waarvoor gedeclareerd is), farmacie, 2010 – 2012. Gelijke kleuren geven dezelfde praktijk aan.

Dat sommige patiënten veel medicijnen (moeten) nemen en dat daar soms hoge kosten mee gemoeid zijn, is bekend. Tabel 8 geeft een overzicht van de 10 patiënten die het grootste totale bedrag per jaar vergoed hebben gekregen aan medicijnen. Merk op dat ziekenhuisapotheken niet in deze tabel vervat zijn. In 2010 staat een patiënt bovenaan waarvoor in dat jaar voor circa 550.000 euro aan medicijnen is vergoed. Dit is een voorbeeld van een exploratieve analyse waarbij daarna een domeinexpert bekijkt wat de totale kosten zo hoog maakt.⁴¹

⁴¹ In ander onderzoek werden op die manier verschrijvingen gevonden in verstrekte aantallen en doseringen.

Totale bedrag vergoed aan medicijnen in 2010 €	Totale bedrag vergoed aan medicijnen in 2011 €	Totale bedrag vergoed aan medicijnen in 2012 €
550.081	586.173	498.142
518.823	519.597	482.416
489.780	453.196	474.474
484.475	452.662	474.057
414.092	452.043	457.355
402.829	405.941	456.725
387.339	401.085	417.089
386.365	398.838	405.244
362.987	397.622	396.617
346.531	380.483	393.500

Tabel 8. Top 10 van het totale bedrag vergoed aan medicijnen voor patiënten, farmacie, 2010 – 2012

De 1000 duurste patiënten zorgen per jaar voor circa 100 miljoen euro aan vergoedingen voor medicijnen die via de apotheken in dit databestand worden verstrekt. Ziekenhuisapotheken bevinden zich in het databestand van medisch specialistische zorg waarover afzonderlijk wordt gerapporteerd.

3.4.5 Verzekerden die langs veel apotheekpraktijken gaan

Binnen het data bestand treffen we veel verzekerden aan die een groot aantal apotheken bezoeken en daar medicijnen krijgen. Zo is er in 2012 een verzekerde die van 87 verschillende apotheken medicijnen kreeg (in totaal 284 vergoedingen!). Overigens trok deze persoon ook al in 2011 de aandacht hiermee: in dat jaar bezocht hij/zij 20 verschillende apotheken. Hij/zij wordt op de voet gevolgd door een verzekerde die ditzelfde gedrag liet zien: in 2011: 37 verschillende apotheken, in 2012: 74 apotheken. Soortgelijke gevallen treffen we aan in 2010 en 2011.

De omzet die verzekerden genereerden die 19 of meer verschillende apotheken bezochten in 2012 bedraagt 127.000 euro.

Aantal verschillende praktijken bezoekt in 2010	Aantal verschillende praktijken bezoekt in 2011	Aantal verschillende praktijken bezoekt in 2012
26	37	87
22	37	74
18	29	42
16	28	41
16	21	41
15	20	32
14	18	26
14	18	21
14	15	20
13	14	19

Tabel 9. Top 10 verzekerden die van veel verschillende apotheken medicijnen kregen, 2010 - 2012.

Even opmerkelijk is het aantal verschillende voorschrijvers per verzekerde. Zie hiervoor tabel 10. In 2012 is er één verzekerde die door 103 verschillende voorschrijvers medicijnen is voorgeschreven. Hij/zij wordt gevolgd door iemand met 78 verschillende voorschrijvers. Het totale bedrag dat gemoeid is met de verzekerden met meer dan 20 voorschrijvers bedraagt in 2012 445.200 euro.

Aantal verschillende voorschrijvers in 2010	Aantal verschillende voorschrijvers in 2011	Aantal verschillende voorschrijvers in 2012
29	51	103
28	39	78
27	31	71
26	31	57
26	31	53
26	30	35
26	29	33
26	27	31
25	27	30
24	26	28

Tabel 10. Top 10 van verzekerden met een hoog aantal verschillende voorschrijvers, 2010 - 2012.

3.4.6 Verslavende medicijnen

Verzekerde	Aantal records opiaten in 2012	Totaal bedrag dat vergoed is €
1	28	27.010
2	15	23.282
3	34	16.733
4	8	13.269
5	53	12.436
6	41	11.624
7	16	9.180
8	58	8.927
9	7	8.513
10	53	8.408

Tabel 11. Overzicht van de top 10 van verzekerden met de vergoeding in 2012

We presenteren hier alleen de tabel voor 2012. Een eerste reactie van de medische expert is dat het hier om verslaving gaat. Ook blijkt dat in de top 10 van 2010, 2011 en 2012 dezelfde BSN's voor komen. Zoals eerder geschetst is bij de paragraaf over het bezoeken van een groot aantal verschillende apotheken, vraagt de top 10 van deze categorie verzekerden de aandacht, net als de gedetecteerde cases voor benzodiazepinen en middelen tegen slapeloosheid (niet in dit rapport opgenomen, wel beschikbaar).

3.4.7 Overtarifiering door toeslag eerste ter handstelling

Eén van de componenten in het vergoedingstelsel is de toeslag eerste ter handstelling. Dit is onderzocht door per apotheek, per jaar de verhouding van het aantal unieke patiënten en artikelcombinaties ten opzichte van het aantal declaraties met een toeslag eerste ter handstelling te beoordelen. Men kan daarbij kijken naar het artikelnummer maar ook naar de combinatie van ATC-code en de toedieningscode. Het idee hierachter is dat slechts eenmaal per jaar die toeslag mag worden gedeclareerd.⁴²

In de beleidsregel van de NZa is opgenomen dat een eerste terhandstellingsgesprek kan worden gedeclareerd indien een geneesmiddel met dezelfde werkzame stof, sterkte en toedieningsvorm niet eerder aan de patiënt ter hand is gesteld in een periode 12 maanden voorafgaand aan de Verstrekking. De werkzame stof, sterkte, toedieningsvorm en hoeveelheid zijn af te leiden uit het artikelnummer. Wanneer er binnen een periode van 12 maanden voor dezelfde patiënt door dezelfde apotheek eenzelfde artikelnummer met een eerste terhandstelling in rekening wordt gebracht dan spreken we van een onterecht in rekening gebrachte eerste terhandstelling.

Uit de analyse bleek dat in 2010 3,4 miljoen euro te veel gedeclareerd is door deze toeslag ten onrechte op te voeren. Voor 2011 gaat het om 2,3 miljoen euro en voor 2012 om 1,9 miljoen euro.

Met de voorgaande analyse sporen we één geval waarin onterecht een eerste terhandstelling in rekening wordt gebracht echter niet op. Dat is het geval waarin werkzame stof, sterkte en toedieningsvorm gelijk blijven, maar de hoeveelheid verandert (bijvoorbeeld doordat er niet 40, maar 20 pillen in het doosje zitten). Hiervoor toetsen we op ATC-code. Dit heeft echter ook als gevolg dat we veranderingen in werkzame stof, sterkte en toedieningsvorm niet meewegen. We nemen daarom ook de toedieningsweg-code mee. Het bedrag dat we met deze toets vinden bevat dan ook terecht in rekening gebrachte eerste terhandstellingen wanneer werkzame stof en sterkte zijn veranderd. De beschikbare data maakt het niet mogelijk om dit verder uit te filteren.

We toetsen het aantal declaraties met een eerste ter handstelling terwijl een medicijn met dezelfde ATC-code en toedieningsweg-code in de 12 maanden⁴³ daarvoor al een keer is verstrekt aan dezelfde persoon. Het teveel vergoede bedrag hiervoor was 17,5 miljoen (2010), 14,1 miljoen (2011) en 16,7 miljoen (2012).

Het daadwerkelijk teveel vergoed bedrag ligt tussen de getallen uit de eerste en tweede toets in. We nemen hierbij de conservatieve resultaten uit de eerste toets mee in het totaal, dat wil zeggen voor 2010 – 2012 respectievelijk 3,4 miljoen, 2,3 miljoen en 1,9 miljoen.

⁴² Hier is coulantie getoond door 330 dagen te hanteren.

⁴³ Hier is coulantie getoond door 330 dagen te hanteren.

3.4.8 Overtarifiering weekterhandstelling

Met de regels die gelden voor de eerste terhandstelling correspondeert een regel die weekterhandstelling betreft. Er is sprake van weekterhandstelling indien het geneesmiddel in samenspraak met voorschrijver en patiënt wordt ter hand gesteld in een weekdoseerverpakking voor een week of voor meerdere weken tegelijk. Onder een weekdoseerverpakking wordt verstaan de verpakking van geneesmiddelen, in gescheiden eenheden per dag of dagdeel, ten behoeve van patiënten die niet in staat zijn zonder een weekdoseerverpakking de terhand gestelde geneesmiddelen verantwoord bij zichzelf toe te dienen, dan wel te laten toedienen.

We hebben getoetst of het voorkomt dat er voor een patiënt binnen de periode van een week⁴⁴ voor hetzelfde geneesmiddel opnieuw een weekterhandstelling in rekening wordt gebracht. Men kan daarbij kijken naar het artikelnummer maar ook naar de combinatie van ATC-code en de toedieningscode. Het idee hierachter is dat slechts eenmaal per week die toeslag mag worden gedeclareerd.

We voeren deze toets eerst uit op basis van het artikelnummer. Deze regel wordt erg veel overtreden. In 2010 troffen we 1.923.300 overtredingen aan, in 2011 1.801.053 en in 2012 2.332.236. Dit betekent dat in 2010 7,5 miljoen euro te veel vergoed is, in 2011 6,1 miljoen euro en in 2012 7,6 miljoen euro.

Vervolgens voeren we de toets ook uit op basis van de combinatie van ATC-code en de toedieningscode. In 2010 troffen we 2.987.907 overtredingen aan, in 2011 3.167.889 en in 2012 4.014.295. Dit betekent dat in 2010 11,6 miljoen euro te veel vergoed is, in 2011 10,7 miljoen euro en in 2012 13,0 miljoen euro.

Het daadwerkelijk te veel vergoed bedrag ligt tussen de getallen uit de eerste en tweede toets in. We nemen hierbij de conservatieve resultaten uit de eerste toets mee in het totaal, dat wil zeggen voor 2010 – 2012 respectievelijk 7,5 miljoen, 6,1 miljoen en 7,6 miljoen.

3.4.9 Dubbele declaraties

Verder is er gekeken in hoeverre er dubbele declaraties voorkwamen, dat wil zeggen declaraties met het zelfde BSN, prestatie en datum maar *door verschillende indieners*. Met andere woorden: meerdere personen in de farmacie hebben een factuur voor deze unieke combinatie van BSN, prestatie en datum ingediend.

In dit onderzoek zijn er in 2010 64.955 van zulke declaraties gevonden, in 2011 69.759 en in 2012 84.001. Het te veel gedeclareerde bedrag bedraagt voor 2010 1,5 miljoen euro, voor 2011 1,6 miljoen euro en voor 2012 eveneens 1,6 miljoen euro.

Het te veel toegekende bedrag per case dat hiermee is gemoeid is als volgt berekend: het hoogste bedrag dat toegekend is aan een indiener is aangemerkt als het juiste bedrag. De som van alle andere toegekende bedragen voor deze case wordt beschouwd als te veel vergoed.

⁴⁴ Uit coulance is uitgegaan van vijf dagen.

3.4.10 Overschrijdingen WMG tariefcode⁴⁵

Tot en met 2011 was iedere tariefcode gekoppeld aan een limiet (MAX1) die slechts in een klein aantal specifieke gevallen overschreden kon worden. Echter het bedrag mag nooit boven een bovengrens limiet (MAX2) komen. Er zijn veel records die de MAX1 limiet overschrijden (circa 30%); een veel kleiner deel overschrijdt de MAX2 limiet (na 2009: bijna 0%). In 2012 waren er geen limieten gesteld. Om een inzicht te krijgen in 2012 hebben we voor dat jaar de limieten die golden in 2011 gebruikt.

Jaartal	MAX1	MAX1 [%]	Totale overschrijding	MAX2	MAX2 [%]	Totale overschrijding
2010	57.969.154	32,1	€ 30.524.155	1179	0,0	€ 4.140
2011	68.309.745	34,9	€ 26.525.310	15.387	0,0	€ 506.095

Tabel 12. Overzicht van het aantal records dat MAX1 en MAX 2 overschrijdt, 2010 - 2011

Ook de variabele WMG-tariefcode is een punt van aandacht: deze wordt niet altijd goed ingevuld.

⁴⁵ Deze tabel is afkomstig uit het FDEC-rapport voor de NZa inzake verstrekkingen, augustus 2013.

3.4.11 Overzicht van enkele anomalieën

3.4.11.1 Mirena-spiraal

Er zijn verschillende manieren om een spiraaltje aan te schaffen en te laten plaatsen. Eén van de mogelijkheden is dat dit spiraaltje door de verzekerde wordt gekocht bij de apotheek en zij het vervolgens mee neemt naar het ziekenhuis of naar de huisarts. Echter, bij het ziekenhuis zitten de kosten van het spiraaltje in de behandeling. Hierdoor wordt het spiraaltje eigenlijk twee maal vergoed.

Dit kan eigenlijk pas goed onderzocht worden als naast de declaratiebestanden van farmacie en huisarts ook de declaraties van medische specialistische zorg beschikbaar zijn.

Nu is onderzocht hoeveel patiënten (BSN) hun huisarts hebben bezocht binnen 15, 30 en 60 dagen nadat een Mirena-spiraal is gekocht. De logica hierachter is dat de patiënten die níet naar de huisarts zijn gegaan, het spiraaltje in een ziekenhuis hebben laten plaatsen en het spiraaltje dus hebben meegenomen naar de specialist.

Dit is een interessante case omdat het twee data bestanden overlapt: apotheken en huisartsen⁴⁶. Wanneer vrouwen het spiraaltje meenemen naar het ziekenhuis om het daar te laten plaatsen, wordt er 150 euro per case te veel vergoed aan het ziekenhuis omdat in de DBC de kosten van de Mirena-spiraal zijn vervat. Nader onderzoek hiernaar is vereist.

Jaartal	Bezoek aan huisarts binnen 15 dagen	Bezoek aan huisarts binnen 30 dagen	Bezoek aan huisarts binnen 60 dagen
2010	75,1%	82,1%	88,0%
2011	78,1%	85,2%	91,0%
2012	81,2%	87,7%	92,4%

Tabel 13. Overzicht van het percentage vrouwen dat binnen een zekere periode de huisarts bezoekt nadat een Mirena-spiraal is gekocht, 2010-2012.

Opvallend om te zien dat ook mannen dit Mirena-spiraal vergoed krijgen, zie tabel 14. Dit kan te wijten zijn aan administratieve fouten; niet is onderzocht of het wijst naar één apotheek of huisarts.

Jaartal	Spiraaltje vergoed aan mannen	Spiraaltje vergoed aan vrouwen	Spiraaltje vergoed aan geslacht onbekend
2010	36	100.001	5
2011	61	59.402	2
2012	57	55.454	1

Tabel 14. Overzicht van het aantal verkochte Mirena-spiraal naar geslacht en jaar, 2010-2012

In de lijst van apotheken met het hoogste aantal verkochte Mirena-spiralen vallen sommige apotheken erg op. Deze lijst moet door domeinexperts worden gerelateerd aan het aantal verzekerden dat bij de apotheek “hoort” en kenmerken van het verzorgingsgebied van deze apotheken.

⁴⁶ Wanneer ook de data bestanden van de ziekenhuizen beschikbaar is, kunnen deze cases helemaal worden gecompleteerd. Voor de huisartsen geldt dat de dekking van de declaratiegegevens in het bestand voor 2010, 2011 en 2012 respectievelijk 95%, 98% en 100% bedraagt. Dit vertekent het beeld voor 2010 en 2011 enigszins.

3.4.12 Schatting van het bedrag gemoeid met het overtreden van regels en anomalieën

3.4.12.1 Overtreden van harde regels

Zoals bekend is er een onderscheid gemaakt tussen het (schijnbaar) overtreden van regels enerzijds en het vaststellende cases anderzijds. De eerste categorie zijn harde regels, bij de tweede categorie moeten experts zich buigen over de duiding van de merkwaardige cases.

De berekening van het te veel vergoede bedrag bij het overtreden van de declaratieregels is relatief eenvoudig. Om vast te stellen of er sprake is van een overtreding, is “soepel” beoordeeld, dat wil zeggen dat de criteria zijn verzacht om twijfel uit te sluiten. Als voorbeeld: bij toets 12 en 13 die gaan over de toeslag voor een eerste ter handstelling (ETH) geldt dat dit per jaar geldt; in het onderzoek is dit criterium van 365 dagen verlaagd naar 330 dagen om twijfel uit te sluiten.

Dit betekent dat de bedragen in de onderstaande tabel relatief hard zijn.

Toets	Te veel vergoede bedrag 2010 in miljoen	Te veel vergoede bedrag 2011 in miljoen	Te veel vergoede bedrag 2012 in miljoen
Eerste terhandstelling	3,4 ⁴⁷	2,3	1,9
Week terhandstelling	7,5 ⁴⁸	6,1	7,6
Dubbele declaraties	1,5	1,6	1,6
Overschrijding WMG	0,0	0,5	nvt
Totaal	12,4	10,5	11,1
Totaal vergoed	5.531	5.280	4.669
Als percentage van het totale bedrag in dat jaar vergoed	0,2	0,2	0,2

Tabel 15. Overzicht van de te veel vergoede bedragen in 2010 - 2012

⁴⁷ Het betreft hier de conservatieve schatting

⁴⁸ Het betreft hier de conservatieve schatting

3.4.12.2 Anomalieën

In paragraaf 1.4 is beschreven hoe er bij anomalie detectie gekeken wordt naar de rechterstaart van de verdeling van enkele interessante variabelen. Vaak zijn dit meta-variabelen die betrekking hebben op de verdeling van werkzaamheden en diensten van de zorgaanbieder.

Bij het beoordelen van de toeslagen eerste ter handstelling en week terhandstelling hebben we gezien dat veel overtredingen plaatsvonden. Ook de aandacht op magistrale verstrekkingen, en het aandeel ervan in het totale aantal verstrekkingen, is getrokken. Deze verhouding verschilt sterk van apotheek tot apotheek en ook als er naar artikelnummer wordt gekeken, worden bijzonder merkwaardige resultaten gevonden.

Behalve de verhouding magistrale verstrekkingen ten opzichte van het totaal, zijn er drie andere meta-variabelen geanalyseerd. Bij deze anomalie detectie hebben we ons ook gericht op enkele specifieke onderdelen namelijk opiaten, benzodiazepinen en middelen tegen slapeloosheid. De resultaten worden in tabel 16 – 18 gegeven. Een uitleg over de referentiepunten 95%, 99% en 99,9% is in paragraaf 1.4 aangereikt. Ook is daar beschreven dat door domeinexperts nader onderzoek naar de anomalieën moet worden uitgevoerd om te bezien of het hier echt om heel merkwaardige cases gaat of, indien een zinnige verklaring kan worden gevonden, de anomalie detectie verder kan worden verfijnd.

Farmacie 2010	Schatting van het te veel vergoede bedrag, in miljoen euro		
Variabele	95% niveau	99% niveau	99,9% niveau
Verhouding magist./totaal	1,3	0,2	0,05
Opiaten	0,6	0,1	0,0
Benzodiazepinen	0,2	0,0	0,0
Slapeloosheid	0,1	0,0	0,0
Totaal	2,2	0,3	0,1

Tabel 16. Overzicht van de te veel vergoede bedragen bij anomalieën, 2010

Farmacie 2011	Schatting van het te veel vergoede bedrag, in miljoen euro		
Variabele	95% niveau	99% niveau	99,9% niveau
Verhouding magist./totaal	1,8	0,8	0,8
Opiaten	0,6	0,1	0,0
Benzodiazepinen	0,2	0,0	0,0
Slapeloosheid	0,1	0,0	0,0
Totaal	2,7	0,9	0,8

Tabel 17. Overzicht van de te veel vergoede bedragen bij anomalieën, 2011

Farmacie 2012	Schatting van het te veel vergoede bedrag, in miljoen euro		
Variabele	95% niveau	99% niveau	99,9% niveau
Verhouding magist./totaal	2,5	1,9	1,9
Opiaten	0,4	0,1	0,0
Benzodiazepinen	0,2	0,0	0,0
Slapeloosheid	0,1	0,0	0,0
Totaal	3,2	2,0	1,9

Tabel 18. Overzicht van de te veel vergoede bedragen bij anomalieën, 2012

3.4.13 Overzicht farmacie

Deze slotparagraaf brengt alle deelresultaten samen voor het farmaciesegment. De resultaten spreken voor zich. Wij achten het referentiepunt van 99% het meest reëel, zie hiervoor tabel 20.

Jaartal	Totaal bedrag vergoed in miljoenen euro A	Te veel vergoed op basis van overtreding declaratieregels ⁴⁹ B	Te veel vergoed op basis van anomalieën (95%) C	Totaal te veel vergoed (95%) D=B+C	Totaal te veel vergoed als percentage van totaal vergoed E=D/A
2010	5.531	12,4	2,2	14,6	0,3
2011	5.280	10,5	2,7	13,2	0,3
2012	4.669	11,1	3,2	14,3	0,3

Tabel 19. Overzicht van de totalen te veel vergoede bedragen, 2010 - 2012, 95% referentiepunt

Jaartal	Totaal bedrag vergoed in miljoenen euro A	Te veel vergoed op basis van overtreding declaratieregels ⁵⁰ B	Te veel vergoed op basis van anomalieën (99%) C	Totaal te veel vergoed (99%) D = B+C	Totaal te veel vergoed als percentage van totaal vergoed E = D/A
2010	5.531	12,4	0,3	12,7	0,2
2011	5.280	10,5	0,9	11,4	0,2
2012	4.669	11,1	2,0	13,1	0,3

Tabel 20. Overzicht van de totalen te veel vergoede bedragen, 2010 - 2012, 99% referentiepunt

Jaartal	Totaal bedrag vergoed in miljoenen euro A	Te veel vergoed op basis van overtreding declaratieregels ⁵¹ B	Te veel vergoed op basis van anomalieën (99,9%) C	Totaal te veel vergoed (99,9%) D = B+C	Totaal te veel vergoed als percentage van totaal vergoed E = D/A
2010	5.531	12,4	0,1	12,5	0,2
2011	5.280	10,5	0,8	11,3	0,2
2012	4.669	11,1	1,9	13,0	0,3

Tabel 21. Overzicht van de totalen te veel vergoede bedragen, 2010 - 2012, 99,9% referentiepunt

⁴⁹ Het betreft hier de conservatieve schatting

⁵⁰ Het betreft hier de conservatieve schatting

⁵¹ Het betreft hier de conservatieve schatting

3.4 Resultaten voor GGZ

3.5.1 Overzicht van de variabelen

Een overzicht van de variabelen wordt in Appendix 6 gegeven. Er is een stapsgewijze aanpak: eerst een beschrijving van de variabelen en de databestanden, daarna de resultaten van de exploratieve analyse en vervolgens toetsen op overtreding van declaratieregels. Voor de GGZ is nog geen anomalie detectie verricht zoals beschreven in 1.4. Deze paragraaf wordt afgesloten met een overzicht van te veel vergoede bedragen.

Het gaat hier om declaraties die vergoed zijn. De bedragen zijn door de zorgverzekeraar aan de zorgaanbieder vergoed, dat wil zeggen inclusief een eventuele opslag of afslag in verband met verrekenpercentages. Er is niet gecorrigeerd voor opbrengstverrekening. Eigen bijdragen van verzekerden zijn niet opgenomen in het vergoede bedrag, eventuele bedragen voor eigen risico wel.

In een klein aantal gevallen worden negatieve bedragen aangetroffen, dit zijn correctieboekingen. Door te aggregeren kan inzicht per verzekerde worden verkregen, net als inzicht per behandelaar, instelling, per verblijf, per prestatiecode, per kwartaal, per jaar, etc., etc. Het totaal bedrag dat is vergoed is in 2010 en 2011 is door afronding op miljoenen euro's gelijk.

In tegenstelling tot de huisartsen, mondzorg en farmacie beschikken we niet over de gegevens van 2012 en zijn 2009, 2010 en 2011 de drie meest recente jaren waarover kan worden beschikt.

Jaartal	Aantal records	Totaal bedrag vergoed in miljoenen euro	Aantal AGB-behandelaars	Aantal AGB-indieners
2009	1.165.353	3.990	2.758	2.915
2010	1.218.992	3.980	3.506	3.349
2011	1.251.854	3.980	4.119	3.660
Totaal	3.636.199	11.950		

Tabel 1. Overzicht van de declaraties GGZ, 2009 – 2011

Het aantal AGB-behandelaars lijkt de laatste jaren sterk te zijn toegenomen. Dat hoeft niet noodzakelijkerwijs in de praktijk zo te zijn: fouten in het betreffende data veld kunnen een goede verklaring daarvoor zijn.

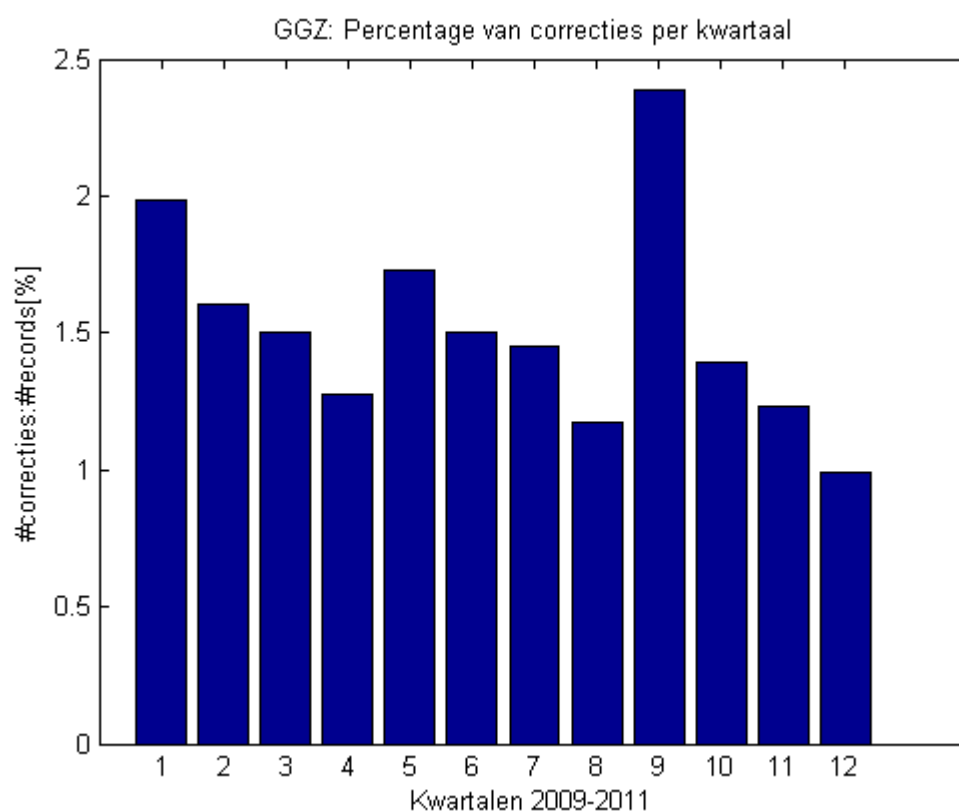
Er is sprake van een flinke concentratie: zo zorgt één instelling voor 13% van alle declaraties in 2009 (ter waarde van in totaal 397 miljoen) en neemt de top 10 35% van alle declaraties voor zijn rekening ten bedrage van 1,4 miljard (2009). Dit zijn enkele zeer grote instellingen.

Ook in het databestand van GGZ 2009 – 2011 komen veel correctieboekingen voor, zie ook hoofdstuk 2. Hoewel kleiner in aantal, zorgen zij voor grotere correctiebedragen. In tabel 2 wordt een overzicht gegeven. De correctierecords zorgen ervoor dat het toegekende bedrag aan GGZ in 2009 105 miljoen euro lager is dan gedeclareerd door dit segment, in 2010 is dit 94 miljoen en in 2011 is dit een bedrag van 81 miljoen euro. Zoals eerder beschreven is de partij die de correctieboeking initieert niet bekend en kan dit zowel de GGZ-instelling als de zorgverzekeraar zijn.

Jaartal	GGZ				
	Aantal correctie records	Netto aantal records	Totale bedrag correcties in miljoen euro	Percentage van de correcties in totaal aantal records	Percentage van het bedrag van de correcties in het totaal gedeclareerd bedrag
2009	18.679	1.120.268	104,5	1,7	2,6
2010	17.919	1.176.607	93,7	1,5	2,4
2011	19.211	1.206.324	80,7	1,6	2,0

Tabel 2. Correctieboekingen in het segment GGZ, inclusief verrekenpercentages, 2009 - 2011

Bij de andere segmenten (huisartsen, mondzorg en farmacie) zagen we dat er nog een aantal correctieboekingen onderweg is voor 2012. Voor GGZ is 2011 het meest recent beschikbare databestand en gezien de hoogte van het percentage van de correctieboekingen is het onduidelijk of nog meer correcties met terugwerkende kracht zullen volgen. Zie hiervoor ook figuur 1. Opvallend is het hoge aantal correctieboekingen dat in het eerste kwartaal plaatsvindt; de uitschieter in het eerste kwartaal van 2011 (kwartaal 9) valt daarbij het meeste op.



Figuur 1. Overzicht van het percentage correctieboekingen in de loop van de tijd, 2009 – 2011

3.5.2 Overzicht van frequentie van declaraties in de GGZ

Een illustratief voorbeeld is de frequentie waarmee verzekerden voorkomen in het bestand van de GGZ. We geven de top 9 van deze verzekerden met het aantal netto records. Zie tabel 3. Het gaat in deze tabel om 26 patiënten; één patiënt komt tweemaal in de top 9 voor en zorgde in 2010 voor 48 declaraties en in 2011 voor 39 declaraties. Een mogelijke verklaring van deze hoge frequenties kan methadonverstrekking aan deze patiënten zijn.

2009	2010	2011
146	65	60
146	59	49
144	51	41
139	51	39
136	50	38
135	48	35
134	48	45
128	48	34
125	46	33

Tabel 3. Top 9 van de frequentie van verzekerden, netto records 2009 – 2011

Opvallend in de databestanden van GGZ is dat in 220.059 gevallen onbekend is wat de prestatiecode is. Dat is in 6% van de cases het geval. Ook is in 90.559 gevallen de AGB-naam niet bekend.

Uit tabel 4 kan men lezen dat, terwijl de totale vergoeding in die jaren nagenoeg constant is, het gemiddelde bedrag per declaratie daalt. Dat geldt echter niet voor de mediaan, die is stijgende.

Jaartal	Totaal bedrag gedeclareerd in miljoenen euro	Gemiddeld bedrag van een declaratie in euro	Mediaan van het bedrag van een declaratie in euro
2009	3.990	3.426	966
2010	3.980	3.265	975
2011	3.980	3.176	993

Tabel 4. Overzicht van het totale bedrag, het rekenkundig gemiddelde en de mediaan, 2009-2011

Eén van de variabelen die gebruikt is, is de gemiddelde omzet per patiënt. Dit is voor de drie segmenten Psychologen, Psychiaters en Instellingen berekend. In tabel 5 wordt de gemiddelde omzet per patiënt (in euro) gegeven voor de top 10 uit deze groep voor het jaar 2010.⁵²

Opvallend is natuurlijk de topper van € 113.427 als gemiddelde omzet per patiënt van een psycholoog in 2010. Wanneer we de resultaten nader bekijken dan zien we dat de top bij de psychologen sterk wisselt: slechts één psycholoog zit in de top 10 voor alle jaren. Anders ligt het met de instellingen: daar wijzigt de top 10 slechts weinig; er zijn zes instellingen die daar stevast in voorkomen.

⁵² Deze overzichten zijn ook beschikbaar voor 2009 en 2011 maar worden hier niet gepresenteerd.

Psychologen	Psychiaters	Instellingen
113.427	17.706	41.016
37.612	13.807	28.196
16.965	13.541	25.359
14.210	13.105	24.941
12.749	9.250	16.298
12.328	8.024	13.310
11.862	6.871	12.709
10.928	6.636	11.974
10.843	6.371	11.744
10.465	6.227	11.138

Tabel 5. Top 10 van de hoogste gemiddelde omzet per patiënt, in euro's, 2010

Nader deskresearch bracht het volgende aan het licht over de psycholoog met een gemiddelde omzet van € 113.427 in 2010. Deze case lijkt dus verklaard.

Het gaat hier om een behandelaar met maar 1 gedeclareerde behandeling in 2010. In 2009 waren dit er 0 en in 2011 juist 133. De data is verrijkt met de huisartsendata waardoor kan worden geconstateerd dat het om een behandeling gaat van een jongetje van 8.

Aan de declaratiecode is tevens te zien dat het om een voortgezette behandeling voor gedragsstoornissen gaat en er sprake was van een vrij lang verblijf. In eerste instantie verwachtten we een gemiddeld bedrag. Om te kijken of dit inderdaad zo is, is gekeken naar vergelijkbare declaraties door psychologen. In het data bestand is deze declaratiecode verder niet terug te vinden voor psychologen. Dat kan komen doordat het gaat om een volgnummer (202). Daarom is de zoekterm afgezwakt naar alleen de combinatie van de verblijfscode en de behandeling (nog steeds voor psychologen). Dit leverde niets op.

Dan blijft eigenlijk alleen de behandelcode over. Dit levert 9 resultaten op met een gemiddeld declaratiebedrag van 44.056 euro. De psycholoog in kwestie heeft 113.427 euro gedeclareerd.

Een andere variabele in de exploratieve analyse is het aantal declaraties per patiënt per jaar. Wanneer we de uitkomsten daarvan vergelijken met tabel 5, zien we niet dezelfde behandelaars terug. De behandelaars die ook in die top 10 zitten, zijn met geel weergegeven, zie tabel 6. Nagegaan moet worden wat deze hoge gemiddelden heeft veroorzaakt.

Psychologen	Psychiaters	Instellingen
10	5	3
6	3	3
4	3	2
4	3	2
4	3	2
4	3	1
3	3	1
3	2	1
3	2	1
3	2	1

Tabel 6. Top 10 van het aantal declaraties per patiënt in 2010, afgerond, 2010

3.5.3 Dubbele facturen

Voor elk van de jaren 2009, 2010 en 2011 is gekeken naar dubbele facturen: facturen die dezelfde BSN hebben, dezelfde prestatiecode, en begin- en einddatum van de prestatie. In elk van die jaren werden driedubbele declaraties gevonden (2009: 116 stuks, 2010: 13 stuks en 2011: 2 stuks). Dit betekent dus dat deze declaraties driemaal zijn vergoed. In alle onderzochte jaren werden er vele dubbele facturen gevonden. Zie hiervoor tabel 7.

Het te veel toegekende bedrag per case dat hiermee is gemoeid is als volgt berekend: het hoogste bedrag dat toegekend is aan een indiener is aangemerkt als het juiste bedrag. De som van alle andere toegekende bedragen voor deze case wordt beschouwd als te veel vergoed.

Jaar	Aantal dubbele declaraties	Als percentage van het totaal aantal declaraties	Te veel vergoed in miljoen euro	Te veel vergoed als percentage van totaal vergoed
2009	11.303	1,0	3,6	0,09
2010	3.939	0,3	1,6	0,04
2011	7.041	0,6	2,0	0,05

Tabel 7. Overzicht van dubbele facturen in de jaren 2009 - 2011

3.5.4 Meerdere DBC's tegelijkertijd geopend voor een patiënt

In het algemeen kunnen voor een patiënt niet meer dan 3 DBC's geopend zijn volgens de spelregels die gelden voor DBC's in de GGZ (zie toets 10 in Appendix 5). Om deze declaratieregel te onderzoeken, hebben we voor iedere patiënt geteld wat het maximale aantal open DBC's is op een bepaalde dag. Het bleek dat in de jaren 2009 – 2011 er 2.831 patiënten waren waarvoor op een zekere dag meer dan drie DBC's open waren⁵³.

Het te veel vergoede bedrag per patiënt is als volgt berekend. Laten we uitgaan van een patiënt met meer dan drie geopende DBC's op een dag. We berekenen eerst het totale bedrag van alle DBC's die voor deze patiënt zijn gedeclareerd. Daarna zoeken we - voor iedere dag van het jaar – het aantal DBC's dat voor die patiënt is geopend (we noemen dit N). Als dit aantal groter is dan 3, dan berekenen we het getal $T = N - 3$. Het getal T noemen we het Teveel. Nadat alle waarden van N en T voor deze patiënt bekend zijn, berekenen we het relatieve aandeel van de categorie Teveel ten opzichte van het totale aantal DBC's. Dit is gelijk aan de som van alle T gedeeld door de som van alle N. Het te veel vergoede bedrag voor deze patiënt is dan gelijk aan het totaal vergoede bedrag voor deze patiënt maal het relatieve aandeel van Teveel ten opzichte van het totaal aantal DBC's.

Het te veel vergoede bedrag is in totaal 4,8 miljoen euro; in 2009: 0,6 miljoen, in 2010 2,0 miljoen en in 2011 2,2 miljoen.

Maximaal aantal open DBC's op een dag	Aantal patiënten 2009	Aantal patiënten 2010	Aantal patiënten 2011
1	776086	974316	1021143
2	70825	108372	112341
3	6233	10948	11374
4	550	1115	1172
5	65	119	123
6	0	9	19
7	2	2	0
8	0	0	0
9	1	0	0
Totaal aantal met meer dan drie DBC's open	618	1.245	1.314

Tabel 8. Maximaal aantal geopende DBC's op één dag in 2009, 2010 en 2011

⁵³ De correcties zijn als volgt behandeld: meerdere records met hetzelfde BSN, prestatiecode, datum opening en datum sluiting zijn samengevoegd tot een enkel record en zijn beschouwd als een enkele prestatie, onder de voorwaarde dat het totale bedrag dat aan deze records was toegekend positief is.

De toets die hiervoor beschreven is, gaat er van uit dat er voor een patiënt niet meer dan drie DBC's tegelijkertijd geopend mogen zijn. Het is ook mogelijk dat de spelregel uitgelegd kan worden als: het is niet toegestaan dat voor één patiënt bij één zorgindieners er tegelijkertijd drie DBC's open staan. Met andere woorden, wanneer een patiënt van meerdere zorgindieners zorg ontvangt, dan is het mogelijk dat er in totaal bijvoorbeeld 4 DBC's open staan (bijvoorbeeld 3 bij één zorgindieners en 1 bij een andere zorgindieners).

Deze interpretatie is onderzocht op grond van de data van 2009, 2010 en 2011. Voor 2009 en 2010 komen we uit op verwaarloosbare bedragen; voor 2012 gaat het om € 81.000.

Voor de totaalberekeningen gaan we uit van de oorspronkelijke toets, nummer 10.

3.5.5 Perioden met overlappende verblijf-DBC's

Andere opmerkelijke cases betreffen het verblijf van een patiënt: het is niet toegestaan dat twee verblijf-DBC's elkaar overlappen⁵⁴. Echter, bij het analyseren van de gegevens over de periode 1 januari 2009 tot 31 december 2011 werden 16.414 patiënten (BSN) aangetroffen waarbij dit wél het geval bleek te zijn.⁵⁵ Tabel 9 geeft de frequenties.

Gevonden verblijf-DBC's die elkaar overlappen	Aantal patiënten
2	15.307
3	1.054
4	48
5	4
7	1
Totaal	16.414

Tabel 9. Aantal patiënten waarvoor er perioden zijn dat verblijf-DBC's elkaar overlappen 2009 - 2011

Zo is te lezen dat er bij 15.307 patiënten dagen/weken werden gevonden met tegelijkertijd twee verschillende verblijf-DBC's.⁵⁶

Het te veel vergoede bedrag dat hiermee over deze drie jaren is gemoeid, bedraagt (berekend op dezelfde wijze als beschreven in paragraaf 3.5.4) 189 miljoen euro. Dit betekent gemiddeld 63 miljoen per jaar. Een verklaring hiervoor kan wellicht zijn dat een patiënt eerder wordt ontslagen dan de DBC aangeeft. Nader onderzoek door domeindeskundigen moet meer inzicht geven in a. waarom deze dagen meerdere malen zijn gedeclareerd en vergoed, en b. of er een redelijke verklaring voor kan worden gevonden. De uitkomsten worden nu niet meegenomen in de berekeningen van het totaal te veel vergoede bedrag.

⁵⁴ Ook hier spelen correctierecords een rol.

⁵⁵ Deze analyse is dus niet per jaar uitgevoerd maar over de periode van drie jaar.

⁵⁶ Deze vorm van gedeclareerd overlappend verblijf werd in het verleden ook al bij declaraties van ziekenhuisopnamen gevonden: bijvoorbeeld een patiënt werd opgenomen van 1 december tot 30 december maar er was ook een factuur voor een opname van 5 december tot 21 december.

De gevonden cases met overlappende verblijf-DBC's zijn niet 1-op-1 vergelijkbaar met (in ander FDEC onderzochte) ziekenhuizen. In de GGZ kan men bijvoorbeeld een DBC zien van bijvoorbeeld 1 januari tot en met 31 maart met daarin 18 tot 24 verblijfsdagen. Stel daarnaast een DBC van 1 maart tot en met 15 juli met 30 tot 36 verblijfsdagen. Het kan dan zijn dat er verblijfsdagen overlappen in maart. Het kan echter ook zo zijn dat er geen enkele overlap in verblijfsdagen is. Het probleem zit hem erin dat de begin- en einddatum van een verblijfs-DBC niet automatisch de begin- en einddatum van verblijf zijn. Er is vaak een periode dat een patiënt wordt opgenomen die korter is dan de duur van de DBC. Dat zou betekenen dat deze toets minder hard is dan gesteld. Echter, het te veel vergoede bedrag dat er mee is gemoeid is wel erg hoog. Vanwege het nader onderzoek dat nodig is voor dit verschijnsel, zijn de geschatte bedragen niet meegenomen in de eindtabel over GGZ.

Hieronder wordt een voorbeeld van een patiënt gegeven en de bijbehorende dagen. Dit voorbeeld geeft niet noodzakelijkwijs een onjuiste wijze van declareren weer maar vraagt om aandacht van domeinexperts.

Voor deze patiënt is in totaal € 871.250 vergoed. Echter, wanneer we rekening houden met de dagen die overlappen (en waarvoor dus twee of meer DBC's actief zijn) komen we uit op een mogelijk te veel vergoed bedrag van € 398.229.⁵⁷

DBC	Nummer van de dag open	Nummer van de dag sluiten	Vergoed bedrag
101011155147	85	449	531.998
202011113076	179	543	30.326
202011135187	450	814	195.165
202011065079	815	1179	53.950
202011053077	544	908	26.818
202011065079	815	1179	0
202011063079	909	1273	32.993
Totaal			871.250

Tabel 10. Een voorbeeld van een patiënt met perioden waarbij meerdere verblijf-DBC's open waren.

⁵⁷ Conform de rekenwijze in 3.5.4: $N = 2190$, $T = 1001$, Totaal vergoed = 871.250 euro; aldus te veel vergoed = $871.250 \times 1.001 / 2.190 = 398.229$ euro. De zesde declaratie, met bedrag 0 (na correcties), telt niet mee.

3.5.6 Schatting van het bedrag gemoeid met het overtreden van regels en anomalieën

3.5.6.1 Overtreden van harde declaratieregels

Zoals bekend is er een onderscheid gemaakt tussen het overtreden van regels enerzijds en het vaststellen van merkwaardige cases anderzijds. De eerste categorie zijn harde regels, bij de tweede categorie moeten experts zich buigen over de duiding van de merkwaardige cases.

De berekening van het te veel vergoede bedrag bij het overtreden van de declaratieregels is relatief eenvoudig. Tabel 11 geeft het overzicht.

Overtreden van declaratieregels	Te veel vergoede bedrag 2009 in miljoen	Te veel vergoede bedrag 2010 in miljoen	Te veel vergoede bedrag 2011 in miljoen
Dubbele facturen	3,6	1,6	2,0
Meer dan drie tegelijk geopende DBC voor één patiënt	0,6	2,0	2,2
Totaal	4,2	3,6	4,2
Totaal vergoed	3.990	3.980	3.980
Als percentage van het totale bedrag vergoed	0,1	0,1	0,1

Tabel 11. Overzicht van de te veel vergoede bedragen in 2009, 2010, 2011

3.5.6.2 Anomalieën

Voor GGZ is geen anomalie detectie uitgevoerd zoals beschreven in 1.4. Door gebrek aan tijd is het niet mogelijk geweest de algoritmen voor de anomalie detectie goed te evalueren en de eruit voortvloeiende cases goed te beoordelen. Daarom ontbreken hier de resultaten van de anomalieën.

3.5.7 Overzicht GGZ

Tabel 12 geeft een overzicht van de resultaten.

Jaartal	Totaal bedrag vergoed in miljoenen euro A	Te veel vergoed op basis van overtreding declaratieregels in miljoen euro B	Totaal te veel vergoed als percentage van totaal vergoed in miljoen euro B / A
2009	3.990	4,2	0,1
2010	3.980	3,6	0,1
2011	3.980	4,2	0,1

Tabel 12. Overzicht van de te veel vergoede bedragen, 2009 - 2011

4. Overzicht van de resultaten

4.1 Inleiding

In korte tijd zijn de records van vier segmenten van de zorg in Nederland doorzocht, zowel gericht gezocht met behulp van de kennis van domeinexperts als met datamining methoden. Veel merkwaardige cases konden worden gevonden. Ook werd duidelijk dat in een aantal gevallen (declaratie-) regels werden geschonden. Daar waar de regels zijn overschreden, kan het hiermee gemoeide schadebedrag worden berekend.

Anders ligt het met merkwaardige en verdachte cases: nader onderzoek naar deze cases is dan nodig. Daaronder valt ook een aantal merkwaardigheden waarmee grote bedragen zijn gemoeid en die nu niet in de resultaten zijn meegenomen (zoals eerste ter handstelling bij farmacie en overlappende DBC's bij GGZ). Dat kan door nader inzoomen door de NZa, door de zorgverzekeraar of beiden.

In paragraaf 4.2 komen we tot een synthese over de overtredingen van declaratieregels en de gevonden anomalieën in deze segmenten in de onderzochte jaren.

De resultaten die in dit rapport zijn beschreven, vormen slechts een klein onderdeel van het totaal aan tabellen, lijsten en overzichten die zijn opgeleverd. Het onderzoek is in een korte tijd verricht en de verwachting is dat, zeker wanneer meer *feedback* op de gevonden cases wordt verkregen, een scherper beeld van de totale omvang van bedrag dat te veel is uitbetaald.

Dit onderzoek heeft ook inzicht gegeven in welke componenten binnen de zorgsegmenten vooral gevoelig zijn voor het optreden van merkwaardigheden.

In deze fase kan een indicatie worden gegeven over de effectiviteit van toetsen en methoden. Dit is nuttig en praktisch omdat sommige van deze toetsen en regels in het automatiseringsproces bij verzekeraars kunnen worden opgenomen. In dat geval wijzen we nogmaals dringend op de noodzaak om velden die bekend moeten zijn, ook met de juiste data te vullen.

De positieve resultaten die als combinatie van methoden en toetsen zijn gegenereerd, werken door in de effectiviteit van beiden.

4.2 Schatting van het bedrag gemoeid met het overtreden van regels en anomalieën

In deze paragraaf worden de deelresultaten van de vier verschillende zorgsegmenten samengebracht. Gebaseerd op onze ervaring met anomalie detectie in verschillende zorgsegmenten en bedrijfstakken achten het referentiepunt van 99% het meest reëel. De betreffende kolommen zijn cursief aangegeven.

Segment	Schatting van het te veel vergoede bedrag, in miljoen euro			
2010	Overtreding regels in miljoen euro	Anomalieën 95%- niveau in miljoen euro	Anomalieën 99%- niveau in miljoen euro	Anomalieën 99,9%- niveau in miljoen euro
Huisartsen	11,1	27,7	5,9	0,4
Mondzorg	0,8	4,8	0,9	0,1
Farmacie	12,4	2,2	0,3	0,1
GGZ	3,6	nvt	nvt	nvt
Totaal	27,9	34,7	7,1	0,6

Tabel 1. Overzicht van de resultaten over 2010.

Segment	Schatting van het te veel vergoede bedrag, in miljoen euro			
2011	Overtreding regels in miljoen euro	Anomalieën 95%- niveau in miljoen euro	Anomalieën 99%- niveau in miljoen euro	Anomalieën 99,9%- niveau in miljoen euro
Huisartsen	15,4	28,8	6,2	0,8
Mondzorg	0,7	4,8	1,0	0,1
Farmacie	10,5	2,7	0,9	0,8
GGZ	4,2	nvt	nvt	nvt
Totaal	30,8	36,3	8,1	1,7

Tabel 2. Overzicht van de resultaten over 2011.

Segment	Schatting van het te veel vergoede bedrag, in miljoen euro			
2012	Overtreding regels	Anomalieën 95%- niveau in miljoen euro	Anomalieën 99%- niveau in miljoen euro	Anomalieën 99,9%- niveau in miljoen euro
Huisartsen	11,6	26,9	5,4	0,5
Mondzorg	2,5	5,6	1,2	0,1
Farmacie	11,1	3,2	2,0	1,9
GGZ	n.a.	nvt	nvt	nvt

Tabel 3. Overzicht van de resultaten over 2012.

Omdat de gegevens van GGZ ontbreken over 2012 is geen totaalberekening voor dat jaar gemaakt.

Een totaalbeeld over de jaren 2010 en 2011 is voor de vier segmenten beschikbaar (zie tabel 1 en 2).

5. Conclusies, observaties en aanbevelingen

5.1 Conclusies

Het is relatief eenvoudig om na te gaan wie er wel of niet aan de declaratieregels heeft voldaan. Regels die door domeinexperts zijn opgesteld kunnen zonder veel moeite op de grote data bestanden worden toegepast. Deze benadering is praktisch en levert direct identificeerbare behandelaars en indieners op. Het mogelijk te veel betaalde bedrag is eenvoudig te berekenen en relatief hard.

Het detecteren van anomalieën met een breed scala van moderne technieken is eveneens goed mogelijk. Erg onwaarschijnlijke cases worden snel zichtbaar. Die zijn soms verklaarbaar, soms ook niet. Het vereist expertkennis, audits en dossiervorming om uit te zoeken of dit gevallen zijn van fraude, van slordigheid, van opzet of dat het een oprechte vergissing is. De technieken blijken effectief te zijn.

De omvang van deze grote data bestanden geeft weinig problemen bij de analyses. Grotere problemen ontstaan door de afwezigheid of onbetrouwbaarheid van *key*-variabelen voor detectie. Dat zijn AGB-behandelaar, AGB-indiener, prestatiecode, aantal prestaties en BSN. Deze variabelen zijn voor het werk van Vektis misschien minder relevant, voor het detecteren van merkwaardigheden zijn ze cruciaal.

Een bijzonder verschijnsel is het aantal creditrecords (correctierecords). De reden van de creditering kan erg divers zijn. Er kan sprake zijn van verrekening (tussen de zorgverzekeraar en de verzekerde) van niet onder de polis vallende verrichtingen maar het kunnen ook correcties in de aanlevering van de data bestanden aan Vektis betreffen. En een veelvoud van andere oorzaken is mogelijk.

Het ontbreken van circa 10% van de declaraties van huisartsen in 2010 heeft geen invloed op het berekende percentage anomalieën in dat segment, wel op de berekening van het totaal van het hiermee gemoeide bedrag. Voor die ontbrekende 10% is bijgeschat. In de data bestanden van 2011 en 2012 deed dit verschijnsel zich gelukkig niet voor.

In dit project zijn veel en concrete merkwaardigheden en afwijkingen gevonden. Deze cases verdienen aandacht omdat het *false positives* (fout positieven) kunnen zijn; in dat geval kunnen de detectieregels waarmee ze zijn opgespoord, worden verbeterd. Anderzijds kan de gerede twijfel over deze cases verder toenemen en komen audits en materiële controle in beeld. Nagegaan moet worden op welke wijze dit traject (samen) met de zorgverzekeraar kan worden opgepakt.

Op basis van de data bestanden is voor de huisartsen, mondzorg, farmacie en GGZ berekend wat het te veel vergoede bedrag in deze segmenten is. We baseren ons op de gevonden overtredingen van declaratieregels én op gevonden anomalieën. Wanneer het detecteren van anomalieën verder wordt doorgezet, zie ook 5.2, zullen er meer worden gevonden. Daarnaast geldt dat er niet-correcte declaraties mogelijk zijn die (nog) niet door de huidige detectiemethoden zijn gevonden. We noemen die niet-correcte declaraties *false negatives*: er is iets mis, het alarm moet af gaan maar dit gebeurt niet. Er zullen zeker meer van dergelijke declaraties worden gevonden wanneer de databestanden worden uitgebreid met meer gegevens waardoor aan detectiekracht wordt gewonnen. Daarom zijn de totaalstellingen van te veel vergoede bedragen een ondergrens.

De berekening van het te veel vergoede bedrag is gedegen. Zij is gebaseerd op aangetoonde overtreding van declaratieregels aangevuld met de resultaten van onderzoek van uitschieters en rechterstaarten van de diverse verdelingen van (meta)variabelen.

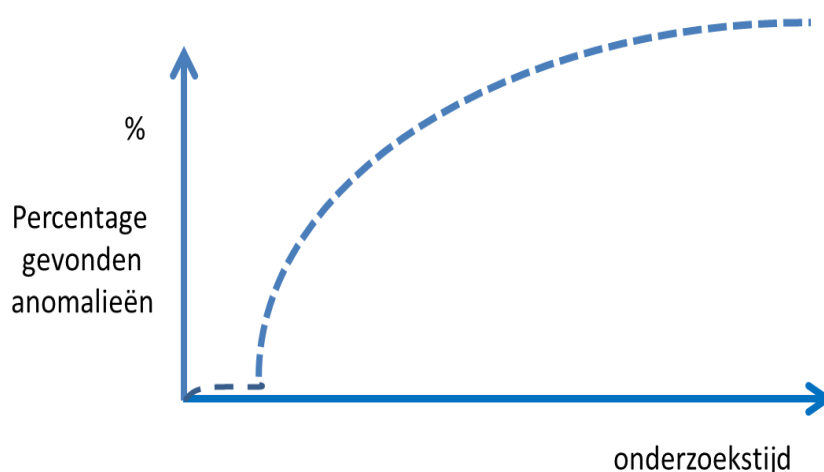
Nader onderzoek zal moeten worden verricht naar het grenspunt dat gehanteerd wordt. In dit rapport hanteren we drie grenspunten: het 95%-punt, het 99%-punt en het 99,9% punt. Wanneer de fraude in de zorg uiterst zeldzaam is, moeten we 99,9%-punt gebruiken. Hoe lager het grenspunt, hoe meer *false positives* (fout positieven, vals alarm) we in de gevonden cases zullen tegenkomen: dat zijn cases waar niets mee aan de hand is en die om een andere reden afwijken van wat van hen verwacht kan worden. De keerzijde is dat bij een hoger grenspunt meer kans is op *false negatives* (onterecht als juist beoordeelde declaraties). Slechts door audits en uitgebreid deskresearch over de betreffende behandelaar/verstrekker (door of met de zorgverzekeraar) kan bepaald worden welke grens het beste is. Op basis van onze ervaringen met het detecteren van anomalieën in de zorg en in andere bedrijfstakken, zijn wij van mening dat het 99%-referentiepunt het beste is met het oog op het effectief terugdringen van niet-correcte declaraties.

5.2 Observaties

De inbreng van domeindeskundigen in dit project is uitstekend geweest. Detectie van anomalieën kan alleen succesvol zijn als er een hecht team is van domeindeskundigen, informatici (die goed en snel met zeer grote data bestanden kunnen omgaan) en wiskundigen (voor het zoeken). Dat is in dit project goed verzorgd.

Kennis en contact met de zorgverzekeraars is een agendapunt: niet alleen wat betreft de opvolging van de gevonden anomalieën maar ook de inhoud van de data bestanden. Enkele variabelen die erg krachtig kunnen zijn voor de anomalie detectie, ontbreken in de data bestanden.

Dit onderzoek is een korte, intensieve tijdsperiode succesvol afgerond. Echter, het is goed mogelijk dat wanneer de periode waarin detectie heeft plaatsgevonden langer zou zijn, er meer is gevonden. In het eerder uitgebrachte rapport over verstrekkingen⁵⁸ is dit geschetst en figuur 1 geeft dit weer. Na een aanlooptijd waarin de voornaamste zorg is de juiste (kwalitatief goede) data bestanden en variabelen te krijgen, vindt intensief onderzoek naar anomalieën plaats. Gaandeweg zal men meer vinden. Inherent is dat het erg lastig te bepalen is wanneer men moet stoppen.⁵⁹



Figuur 1. De relatie tussen het percentage gevonden anomalieën versus de onderzoekstijd

⁵⁸ FDEC-rapport aan de NZa, 2013

⁵⁹ Dat is de reden dat organisaties waarvoor detectie cruciaal is (creditcardmaatschappijen, banken, credit management bedrijven) een eigen afdeling oprichten.

5.3 Aanbevelingen

In dit project zijn een groot aantal toetsen en detectieregels getest op de data bestanden. De meest succesvolle toetsen kunnen worden gedeeld met de zorgverzekeraars opdat zij die onregelmatigheden sneller en eenvoudiger kunnen signaleren. Daarmee worden onregelmatigheden al bij de bron aangepakt.

Dit onderzoek kan een uitstralende werking hebben op de beroepsgroepen. Uit Amerikaanse literatuur (agrarische oogstverzekeringen) is bekend dat *alleen al het feit dat boeren denken dat de verzekeraar luchtfoto's gebruikt om de oogstschade te schatten*, zowel het aantal claims als de gemiddelde hoogte van de claims dalen. Hoewel dit effect moeilijk in te schatten is in deze vier zorgsegmenten, is het overduidelijk dat de behandelaar/indiener meer op zijn hoede is. Dat daardoor de anomalieën kunnen verschuiven naar andere behandelingen/verrichtingen/medicijnen is een *fact of life*. Hiermee is de wedloop tussen overtreders en zij die de detectie doen, geschetst.

Omdat de schatting van de omvang van de onregelmatigheden door zoveel factoren wordt bepaald, stellen wij voor om de aandacht niet op het absolute percentage te richten maar op *de verschillen in gedetecteerde waarden per jaar*. Wanneer we daar een daling zien, kunnen we constateren dat we op de goede weg zijn.

Al eerder is gemeld dat de kwaliteit van sommige variabelen ontoereikend is. Bij vervolgstappen adviseren wij dringend om in het data bestand op te nemen: het rekeningnummer van de begunstigde en de geboortedag en geboortemaand. Ook de zescijferige postcode is nodig om de vergoeding in achterstandswijken te kunnen beoordelen.

Declaratiemonitoring (het volgen van een grote datastroom van declaraties) kan een aanvulling zijn van het toetsingskader. Het is dan een toezichtsinstrument om een zorgverzekeraar of zorgverlener te volgen. De voorkeur gaat uit naar enthousiaste samenwerking met een zorgverzekeraar in plaats van het afdwingen van een dergelijk toezicht.

Dit onderzoek betreft vier zorgsegmenten. Aggregatie van alle bestanden naar verzekerden (via het encrypte BSN) is een slag die gemaakt moet worden. We krijgen dan een compleet beeld van de verzekerde (uitgezonderd medisch specialistischer hulp, fysiotherapie, en enkele andere segmenten) wat de weg opent naar studie van meerdere partijen gezamenlijk in de zorg.

Appendix 1. Toelichting op methoden

Deze bijlage bevat een korte toelichting op de diverse methoden die zijn gebruikt om anomalieën⁶⁰ in de databestanden op te sporen. Een overzicht over de literatuur is te vinden op www.fdec.nl en bijvoorbeeld in Chandola, Banerjee en Kumar (2009): “Anomalie Detectie: A Survey”.

Het is belangrijk om op te merken dat wij met name die methoden en technieken kiezen die leiden tot een efficiënte en effectieve detectie van anomalieën; in de wetenschappelijke schoonheid, hoewel belangrijk, zijn wij in dit kader minder geïnteresseerd.

1. Exploratieve analyse

Bij exploratieve analyse wordt inzicht in de data verkregen door meerdere dwarsdoorsneden te maken, frequentieverdelingen, histogrammen, etc. teneinde overzicht te krijgen in de verdeling van de betreffende variabele. Uitschieters worden direct zichtbaar net als gegevens die ontbreken of niet volledig zijn. Volledigheidsanalyses worden uitgevoerd en ook wordt de data vergeleken met cijfers uit andere bronnen om de plausibiliteit te controleren.

Eenvoudige aggregatieberekeningen worden gemaakt, bijvoorbeeld per huisarts wordt de totale en gemiddelde vergoeding berekend op basis van zijn/haar patiënten. Andere voorbeelden zijn het percentage ongeldige prestatiecodes per indiener en het referentiecijfer van de indiener wat betreft de verhouding tussen lange en korte consulten, de verhouding tussen dag/weekend, etc., etc.

We kijken naar extreme waarden zoals bijvoorbeeld: hoe kunnen we de zorgaanbieders typeren die de hoogste vergoeding in Nederland hebben gekregen? Wat kunnen we zeggen over patiënten die meer dan 50 huisartsen hebben bezocht? Wat is er aan de hand met patiënten die meer dan 100 maal een zorgaanbieder per jaar bezoeken? Dit zijn slechts enkele eenvoudige voorbeelden.

De uitkomsten van deze exploratieve analyse worden voorgelegd aan experts op het terrein van deze betreffende zorg.

2. Expertkennis, common sense en beslisregels

Experts op het terrein van de zorg weten vaak waar interessante cases te vinden zijn, ongeacht of die onregelmatig, onrechtmatig of frauduleus zijn. Zij kunnen vrijwel direct aangeven of een case interessant is of niet. Voor een overzicht van dit soort regels/toetsen wordt verwezen naar de Appendix. Deze expertregels zijn bijzonder effectief gebleken.

Wanneer experts de resultaten van de exploratieve analyse beoordelen, ontstaat vrijwel direct de driedeling in onregelmatigheid, onrechtmatigheid en fraude. Voor sommige cases is echter meer nader onderzoek nodig en kan er over deze cases (nog) geen uitspraak worden gedaan.

Het oordeel van de expert kan eenvoudig worden meegenomen in het verdere zoekproces.

⁶⁰ Een *anomalie* is een opmerkelijke case in een zeer groot databestand die sterk afwijkt van hetgeen van die variabele(n) (dienstverlener, product, dienst, klant/patiënt/verzekerde, tijd, tijdstip, etc.) verwacht kan worden.

Voorbeelden zijn

- Niet-logische combinaties van behandelingen
- Speciale trucs/constructies die door sommige huisartsen worden gehanteerd
- Overlappende prestatiecodes
- Upgrading van prestaties
- Dubbele declaraties

3. Model based anomalie detectie

1. Statistische methoden

De veronderstelling die aan deze methoden ten grondslag ligt is dat “normale” gevallen te beschrijven zijn als een of ander statistisch proces met een waarschijnlijkheidsverdeling. We modelleren de doelvariabele met behulp van een parametrische of non-parametrische verdeling. Welke verdeling uiteindelijk als basis dient voor de anomalie detectie verschilt van variabele tot variabele en is minder interessant; belangrijk is dat het een voldoende stevige basis biedt én dat echte anomalieën worden gevonden.

We kijken vervolgens naar de linker en rechterstaarten van de berekende verdeling. Boven een gekozen kritische waarde (bijvoorbeeld het gemiddelde + 3 sigma) of beneden een gekozen kritische ondergrens (bijvoorbeeld het gemiddelde – 3 sigma) bevinden zich dan de cases die om aandacht schreeuwen. Naar deze cases wordt nader gekeken en wanneer geen logische verklaring gevonden kan worden, worden ze voorgelegd aan een expert op het terrein van zorg.

Voorbeeld:

We kunnen voor (bijv.) mondzorg nagaan welke praktijken een onverwacht hoog percentage drielaksvullingen hebben. Elke tandarts kan veel of weinig drielaksvullingen verrichtingen uitvoeren. Om een gemiddelde verdeling te vinden van de tandartsen, trekken we aselekt een (voldoende grote) steekproef uit alle tandartsen.

Op basis van die gevonden verdeling kunnen we nu uitspraken doen over het te verwachte aantal drielaksvullingen bij een zekere tandarts. Tevens kunnen we de waarschijnlijkheid of onwaarschijnlijkheid aangeven om meer dan X drielaksvullingen bij een tandarts aan te treffen.

Deze analysetechniek levert vaak goede, praktische resultaten. Immers, de tandartsen met een zeer hoge onwaarschijnlijkheid worden direct zichtbaar; een onwaarschijnlijkheid van 1:1.000.000.000 komt regelmatig voor. De volgende stap is na te gaan wat die betreffende tandartsen zo uniek maakt.

Ook wordt er gekeken naar de verdelingen van de residuen: het verschil tussen de werkelijke waarde en de door het model voorspelde waarde. Als de voorspelde waarde veel kleiner is dan de werkelijke waarde, dan is er blijkbaar iets aan de hand. De residuen hebben meestal, min of meer, een normale verdeling en de z-waarde is dan een goed handvat om uitschieters te vinden. Er zijn ook andere maatstaven waarmee we de residuen kunnen doorgronden.

Uitsluitend focussen op de rechterstaart van de verdeling is niet correct: het is mogelijk dat fraudeurs een wisselend patroon gebruiken (zoals bijvoorbeeld bekend is bij creditcard transacties) en slechts af en toe het “dure” segment betreden. Daarom is ook aandacht voor de linkerstaart nodig.

Er is een veelvoud van technieken die hier worden gebruikt. We gaan hier niet dieper op in.

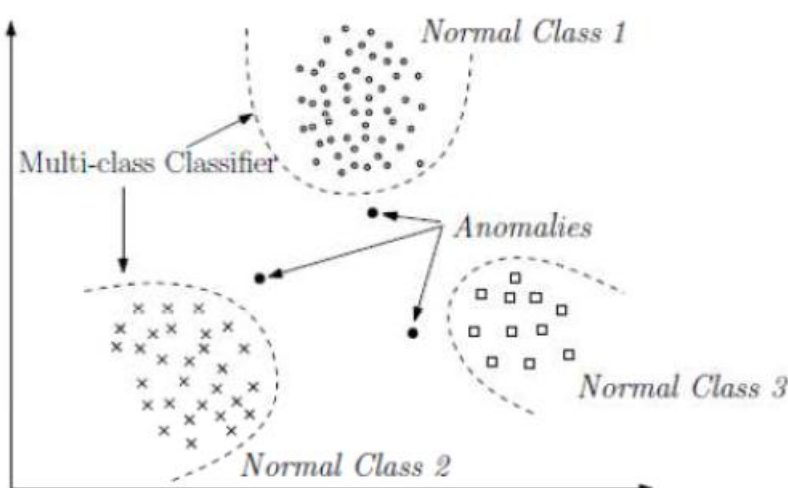
2. Regressie en classificatie technieken

Regressieanalyse is één van de vele statistische methoden die hiervoor al even zijn aangestipt. Omdat deze techniek vaak wordt gebruikt, beschrijven we het hier afzonderlijk.

De eerste stap is dat een multiple regressiemodel wordt opgesteld voor een variabele die we willen onderzoeken. De coëfficiënten van het regressiemodel worden geschat. Als voorbeeld: we willen de gemiddelde vergoeding van een huisarts voorspellen op basis van de verdeling van zijn/haar patiënten naar leeftijd en geslacht, alsmede het aantal artsen dat in een praktijk werkt.

Op basis van dit geschatte model (per jaar) berekenen we vervolgens voor iedere huisarts het verschil tussen de waarde die het model berekent voor dat jaar (dus hetgeen op basis van het model kan worden verwacht van deze arts rekening houdend met de variabelen in het model) en de gerealiseerde waarde van de huisarts in dat jaar. Hoe groter het verschil tussen verwachte vergoeding en de werkelijke vergoeding, des te interessanter de huisarts.⁶¹

Bij classificatie technieken vormen homogene groepen het uitgangspunt. Zie hiervoor figuur 1.



Figuur 1. Voorbeeld van anomalie detectie uitgaande van homogene groepen

De werkwijze is stapsgewijs:

1. Vind enkele “natuurlijke” klassen of categorieën van “normale” gevallen
2. Bepaal voor elke klasse een indicator of grenslijn die de klasse juist typeert en deze onderscheidt (afscheidt) van andere klassen
3. Pas deze indicator en grenslijnen toe op alle cases
4. Cases die – volgens de indicatoren of grenslijnen – tot geen van de onderscheiden klassen behoren, zijn anomalieën

⁶¹ Natuurlijk wordt er zorgvuldig gekeken of het model robuust genoeg is om op basis ervan naar anomalieën te speuren. Zaken als verklaarde variantie, de verdeling van residuen, heteroscedasticiteit en overfitting vormen hier een onderdeel van.

3. *Random Forests*

In een eerdere paragraaf hebben we uiteengezet hoe we met behulp van regressiemodellen merkwaardige personen of instellingen kunnen vinden in de data. In deze paragraaf zullen we in het kort een toelichting geven op een classificatie- en regressiemodel die door de jaren veel succes heeft genoten, het Random Forest model⁶².

Random Forests zijn een aggregaat van een verzameling modellen die beslisbomen heten. Het doel van beslisbomen is om een doelwaarde, bijvoorbeeld 'man' of 'vrouw', te voorspellen op basis van een nieuwe, niet eerder bekeken, data-instantie. Beslisbomen kunnen ook een regressie uitvoeren, waardoor ze geschikt worden om numerieke doelwaarden, zoals de gemiddelde kosten per patiënt, uit te rekenen. De beslisboom wordt dusdanig geconstrueerd dat de correctheid van de voorspelling wordt geoptimaliseerd. Echter kan de afwijking, d.w.z. het verschil in de daadwerkelijke en voorspelde waarden, in sommige situaties groot zijn.

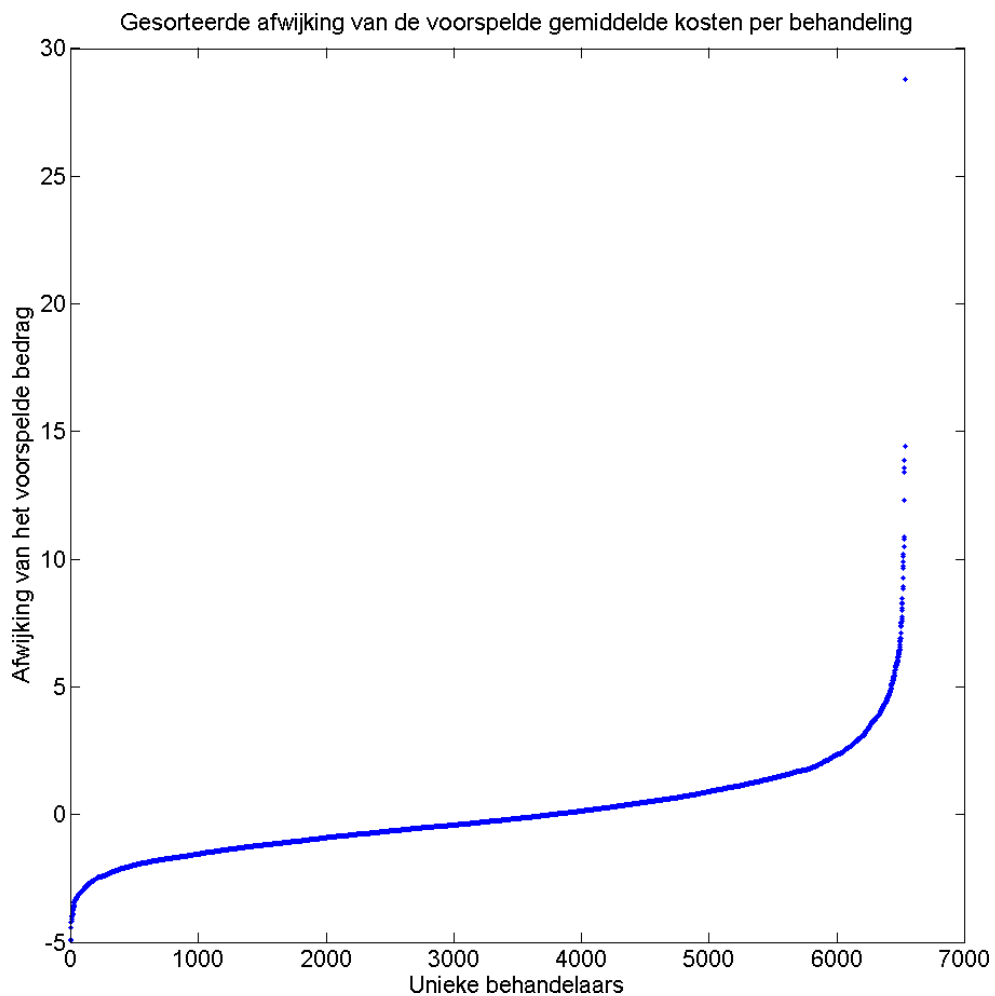
Een oplossing voor dit probleem is om niet één, maar meerdere ongecorreleerde modellen te vervaardigen en de voorspellingen van die modellen te middelen. We selecteren hiervoor ongeveer 2/3 van de data voor het opbouwen van elke boom door herhaaldelijk data, bijvoorbeeld declaratierecords, te selecteren met vervanging. Een bijkomend voordeel is dat we per beslisboom ongeveer 1/3 van de oorspronkelijk data overhouden. Wanneer we veel bomen, c.q. modellen, creëren kunnen we er zeker van zijn dat we de data, die gebruikt is om alle modellen te maken, tevens gebruikt kan worden in de classificatiestap, waarin we de doelwaarde voorspellen.

Om de procedure te verduidelijken bekijken we een Random Forest model zoals die is gemaakt voor de huisartsendata van 2010. We bekijken alleen de huisartsen die minstens 6532 verrichtingen hebben gedaan in dat jaar.

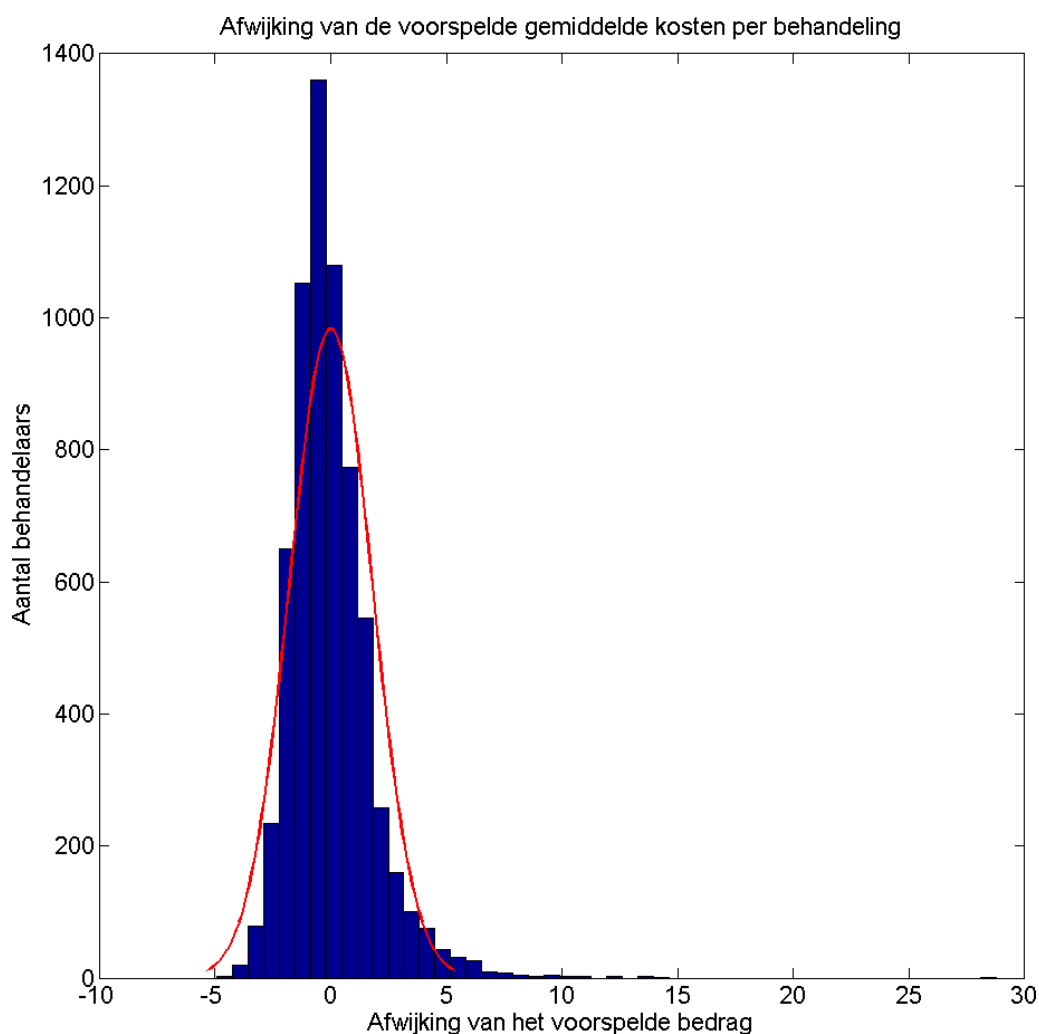
Allereerst moeten we een doelwaarde instellen. Hiervoor kiezen we de gemiddelde declaratiekosten per patiënt, zoals deze per behandelaar is uitgerekend. Ten tweede moeten we variabelen instellen die de doelwaarde kunnen voorspellen. Hiervoor kiezen we de leeftjidsverdeling in procenten voor zes leeftijdscategorieën. Vervolgens bouwen we 1000 beslisbomen op, gebruikmakende van de gekozen voorspelwaarden. Het resultaat van deze stap is dat we 1000 modellen hebben die ongecorreleerd op zijn gebouwd en allen in staat zijn de doelwaarde, de gemiddelde declaratiekosten per patiënt, te berekenen.

In figuren 1 en 2 is te zien dat het model voor de meeste behandelaars een redelijke voorspelling kan maken van de gemiddelde kosten per patiënt. Er is echter een aantal behandelaars waarvoor de afwijking, met een maximum van bijna 29 euro, fors te noemen is. Deze behandelaars zijn dan ook cases om nader te inspecteren.

⁶² Breiman, Leo (2001). "Random Forests". Machine Learning 45 (1): 5–32.



Figuur 1: Het Random Forest model voorspelt de gemiddelde kosten per patiënt. De behandelaars zijn vervolgens gesorteerd op de afwijking van de daadwerkelijke kosten per patiënt, waardoor deze S-curve ontstaat. Voor veel behandelaars komt het model aardig in de buurt, maar rechtsboven zijn een aantal outliers te zien.



Figuur 2: De distributie van de afwijking laat zien dat deze enigszins normaal verdeeld is. De behandelaars in het linkergedeelte zijn goedkoper dan voorspeld is, terwijl die in het rechter gedeelte duurder zijn. Dit zijn de behandelaars om nader te onderzoeken.

4. Methoden gebaseerd op afstandsmaten

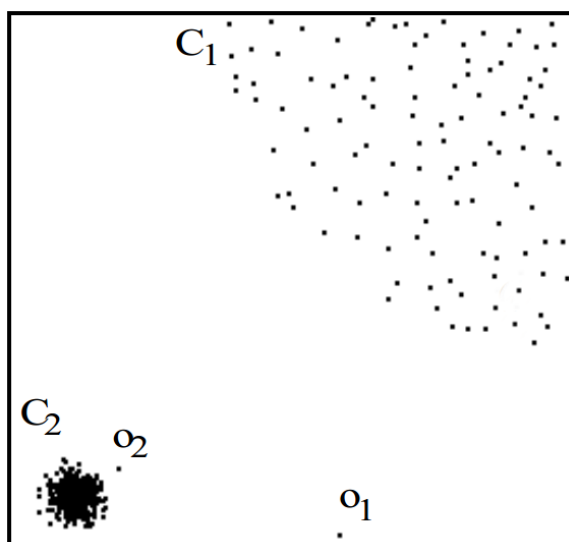
Entiteiten (huisartsen, apothekers, praktijken, patiënten, etc.) kunnen worden gerepresenteerd door hun profielen: vectoren van getallen, bijvoorbeeld de frequenties van de tien meest voorkomende declaraties, de frequenties van medicijnen, verhoudingen in consulten, etc., etc.

Deze vectoren kunnen beschouwd worden als punten in een meerdimensionale ruimte. In die meerdimensionale ruimte gaan we op zoek naar entiteiten die veel op elkaar lijken, dat wil zeggen: relatief dicht bij elkaar in de buurt liggen.

De anomalie-score van een case is dan de afstand tot zijn dichtstbijzijnde buur⁶³ of de gemiddelde afstand tot zijn meest nabije burens. Hoe groter de gemiddelde afstand, hoe meer deze case “apart” ligt. De top N van de meest geïsoleerde (ver weg gelegen) cases worden als afwijkend gezien.

Local outlier factor (LOF) is een algoritme om datapunten te vinden die abnormaal zijn wanneer ze vergeleken worden met de datapunten in de omgeving. Onder omgeving verstaan we de datapunten die het meest lijken op het datapunt in kwestie. Dit zijn de dichtstbijzijnde burens van een datapunt. Hoe meer datapunten op elkaar lijken hoe dichter ze bij elkaar liggen. Als de omgeving van een datapunt erg afwijkt van de omgeving van elk van de dichtstbijzijnde burens dan spreken we van een outlier.

Om dit te verduidelijken zie figuur 2. Hier zien we clusters C_1 en C_2 . Punten in cluster C_2 liggen dicht bij elkaar en zijn dus datapunten die erg op elkaar lijken. Ook cluster C_1 bevat datapunten die op elkaar lijken, hoewel in mindere mate. Dit is te zien aan de lagere dichtheid van cluster C_1 . Voor elk van de punten in C_1 en C_2 geldt dat de dichtheid (de afstand tot de dichtstbijzijnde burens) van een punt ongeveer gelijk is aan de dichtheid van zijn dichtstbijzijnde burens. Voor outliers o_1 en o_2 geldt dit niet. De dichtstbijzijnde burens van deze punten liggen veel verder weg dan dat van hun dichtstbijzijnde burens. Een belangrijke observatie hierbij is dat de afstand van o_2 naar zijn dichtstbijzijnde burens, die in C_2 liggen, ongeveer gelijk is als de afstand tussen punten binnen C_1 . Toch wordt outlier o_2 gevonden omdat we alleen op lokaal niveau naar verschillen in dichtheid kijken.



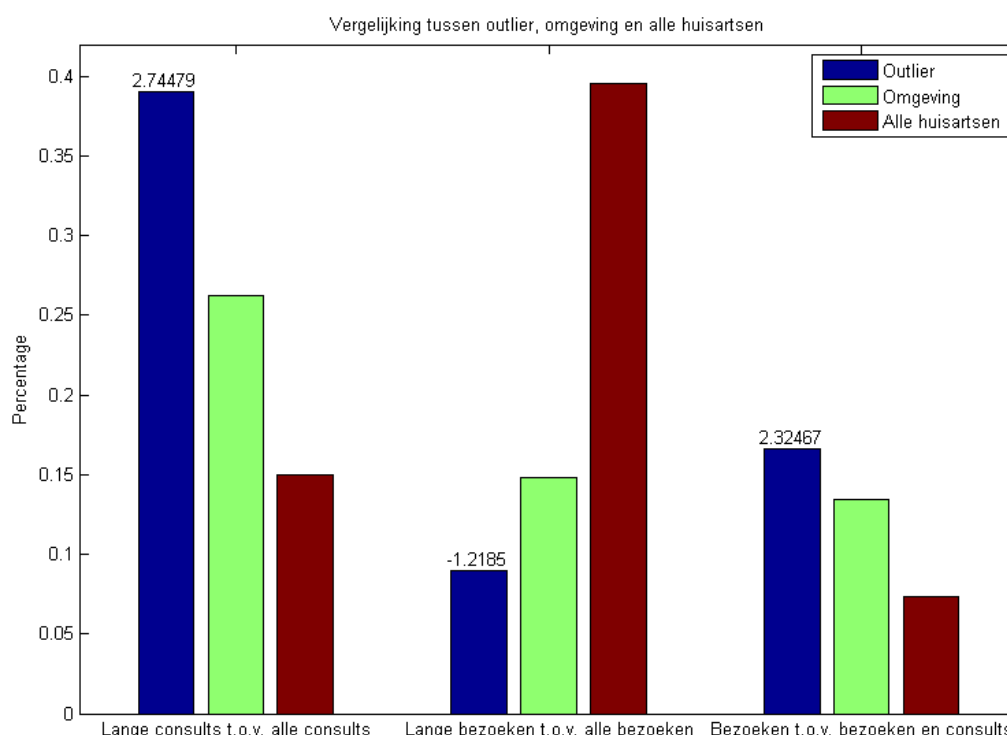
Figuur 2. Voorbeeld van anomalie detectie uitgaande van relatieve dichtheid

⁶³ Het gaat hier om k-th nearest neighbor methoden waarvan geprobeerd wordt die hier eenvoudig uit te leggen.

LOF geeft aan elk datapunt een score. Hoe lager de dichtheid van een punt ten opzichte van zijn burens is, hoe hoger de score en dus hoe meer een datapunt afwijkt van zijn dichtstbijzijnde burens. Een hogere score duidt dus op een grotere outlier. Een van de voordelen van LOF is dat het alleen lokaal punten vergelijkt waardoor outliers gevonden kunnen worden die niet gevonden zouden zijn wanneer er op globaal niveau gezocht zou worden.

Voorbeeld: LOF is gebruikt om te zoeken naar huisartsen die outliers zijn. Hiervoor is gekeken naar het percentage van consulten dat lang was, het percentage van bezoeken dat lang was en het percentage van het totaal van bezoeken en consulten dat een bezoek was. In alle drie de gevallen is het gewenst dat deze percentages zo laag mogelijk zijn omdat een lager percentage op minder kosten duidt. Zo kosten lange consulten bijvoorbeeld meer dan korte consulten. Wanneer een huisarts percentueel gezien veel meer lange consulten verricht dan andere huisartsen kan dit duiden op fraude.

Een voorbeeld hierbij is figuur 3. Deze huisarts (weergegeven in blauw) voert percentueel gezien veel meer lange consulten uit dan de huisartsen die qua verrichtingen het meest op hem lijken (groen). Bovendien verricht hij al helemaal veel meer lange consulten dan gemiddeld (rood): 2,7 standaarddeviaties. Ook voert deze huisarts veel huisbezoeken uit ten opzichte van het aantal consulten dat hij verricht, hoewel we uit het percentage lange bezoeken kunnen concluderen dat naar verhouding veel van deze bezoeken kort zijn.



Figuur 4. Voorbeeld van een gedetecteerde huisarts met opvallend veel lange consulten en veel huisbezoeken, huisartsen, 2010--2012

5. Subgroep Discovery en Cortana

Binnen de datamining verwijst de term *subgroup* naar een deel van de datapopulatie, waarbinnen het gedrag opvallend afwijkt van de norm (feitelijk de rest van de populatie). Deze subgroup wordt beschreven in termen van simpele condities op de beschikbare variabelen, zodat je niet alleen maar een lijst met records in de database hebt die afwijken (de subgroup), maar ook een simpele omschrijving van deze groep records (de *description*). Deze omschrijving kan door de domeinexpert gebruikt worden om inzicht te verkrijgen in de aard van de afwijking binnen de subgroup, of om een verklaring te vinden voor deze afwijking.

Subgroup Discovery – het vinden van één of meerdere subgroups – kan op verschillende manieren voor fraudedetectie ingezet worden. Hoewel de precieze inzet van SD een technische discussie is, gaat het over het algemeen om het vinden van groepen van afwijkende records. Uit de beschrijving van de groepen zullen experts moeten afleiden of het hier gaat om verklaarbare afwijkingen, bijvoorbeeld een behandelaar die zich specialiseert in bepaalde patiëntgroepen, of dat de records in kwestie afwijken in nieuwe manieren die opvallen of verdacht zijn. In dat geval zou het dus kunnen gaan om (verschillende gradaties van) onregelmatigheid of fraude. In de tekst hieronder worden details gegeven over twee manieren waarop SD in dit project ingezet is.

Merk op dat SD lijkt op de andere courante techniek voor fraudedetectie – Outlier Detection – die ook als doel heeft om opvallende of afwijkende records te herkennen. Maar waar het bij Outlier Detection gaat om een beoordeling per individueel record, gaat het bij SD vooral om afwijkend gedrag van *groepen* records. Het voordeel is dat het bij groepen van, stel, behandelaars kennelijk gaat om systematische afwijkingen die door meerdere behandelaars gehanteerd worden, terwijl bij een individuele outlier er niet (noodzakelijkerwijs) sprake hoeft te zijn van een structurele onregelmatigheid. Daarbij is er natuurlijk bij groepen potentieel een groter totaalbedrag gemoeid, zodat de impact van het gevonden patroon groter is. Als laatste is het beschikbaar zijn van een omschrijving van de subgroup natuurlijk erg praktisch voor het interpreteren van de onregelmatigheid.

Subgroup Discovery technieken zijn de afgelopen jaren in de wetenschappelijk wereld ontwikkeld, waarbij de Universiteit Leiden een grote rol heeft gespeeld in het standaardiseren en op nette wijze integreren van de verschillende ontwikkelingen. Dit heeft geresulteerd in een open source SD pakket, genaamd Cortana. Dit pakket is tijdens dit project gebruikt voor verschillende vormen van fraudedetectie in de zorg. De software en achtergronddocumentatie zijn te downloaden van <http://datamining.liacs.nl/cortana.html>.

Subgroup Discovery voor benchmarking

De eerste manier waarop SD voor fraudedetectie gebruikt wordt, gebruikt data over behandelaars op declaratie-niveau. Alle declaraties per markt (in deze tekst worden huisartsen als voorbeeld gebruikt) worden opgesomd over alle behandelaars. Dit geeft een impliciete groepering van de declaraties in plusminus 13 duizend brokken declaraties, voor de 13k huisartsen die voorkomen in deze data. Om nu te bepalen of een specifieke huisarts *H* afwijkt in declaratiegedrag van de andere huisartsen, kunnen alle declaratie van *H* gelabeld worden. SD is dan in staat om te bepalen wat een goede en compact beschrijving is, waarin de gelabelde declaraties afwijken van de rest van de data. De beschrijving van huisarts *H*, die natuurlijk maar in grove lijnen klopt, geeft dan een indicatie van opvallende zaken in de declaraties van de gelabelde huisarts. Hier moet natuurlijk nog een filteringslag plaatsvinden, waarbij onderscheid gemaakt wordt tussen specialisatie van een huisarts en echte onregelmatigheden in bijvoorbeeld de frequentie van bepaalde prestatiecodes. Dit proces van vergelijking en analyse wordt uiteindelijk herhaald voor elke afzonderlijke huisarts. Dit proces wordt ook wel benchmarking genoemd.

In het geval van de huisartsenmarkt gaat het in totaal om globaal 260 miljoen declaraties, zodat er gemiddeld zo'n 20k declaratie per huisarts te onderscheiden zijn (voor de periode waarover deze analyse gaat). In elke benchmarking-ronde worden dus dit aantal declaraties apart gezet, en probeert Cortana een zo exact mogelijke omschrijving te geven van deze groep declaraties. Dit proces wordt 13 duizend keer herhaald, wat dus een aanzienlijke rekenkundige uitdaging is.

Subgroup Discovery over geaggregeerde data

De volgende vorm van inzet van SD heeft betrekking op geaggregeerde data. Het valt te verwachten dat een frauderende huisarts (tandarts, apotheker, ...) niet te herkennen is aan één enkele opvallende declaratie, maar meer aan een opvallende declaratiepatroon dat alleen blijkt uit het totaalplaatje van declaraties van de huisarts. Om die reden zou je dus een huisarts graag beschrijven in termen van eigenschappen over een *verzameling* declaraties, in plaats van eigenschappen van individuele declaraties. In andere woorden, je zou graag de declaraties aggregeren naar het niveau van behandelaar, zodat je niet een database van 260m records hebt, maar één van 13k. De grote uitdaging hier is natuurlijk welke eigenschappen (variabelen) op huisartsenniveau relevant zijn. Voor deze analyse hebben we een flink aantal variabelen gedefinieerd die het declaratiegedrag vangen in termen van de verdeling van behandelingen (percentages per prestatiecode, bedrag, ...) en eigenschappen van de patiënten (verdeling geslacht, gemiddelde leeftijd, ...).

Deze aggregatiestap beschrijft elke huisarts op zinvolle wijze, maar zegt nog niets over de mate waarin een huisarts opvalt of afwijkt van de norm. Daarvoor voegen we met verschillende methodes⁶⁴ zogenaamde *outlier scores* toe. De uitdaging voor SD is nu: *vind subgroups van huisartsen waarvoor de outlier score opvallend hoog is*. In andere woorden, vind groepen van outliers die op een makkelijke manier te beschrijven zijn in termen van de geaggregeerde eigenschappen van huisartsen. Merk op dat we op deze manier voortbouwen op resultaten die door middel van andere analyses geproduceerd zijn.

⁶⁴ Bijvoorbeeld LOF (Local Outlier Factor) scores.

Literatuurverwijzingen

Interessante overzichtsartikelen (niet uitputtend) zijn:

Leo Breiman, (2001). "Random Forests". Machine Learning 45 (1): 5–32.

Richard J. Bolton and David J. Hand , 2002, Statistical Fraud Detectie: A Review, Statistical Science, Vol. 17, No. 3 (Aug., 2002), pp. 235-249.

Varun Chandola, Arindam Banerjee, and Vipin Kumar, 2009, Anomalie detectie: A survey. ACM Comput. Surv. 41, 3, Article 15 (July 2009).

H.M.P. Kersten en W.J. Kowalczyk, 2009, Transactiemonitoring: wie heeft er nog overzicht?. Jaarboek Compliance, 307-320.

E.W.T. Ngai, Yong Hu, Y.H. Wong, Yijun Chen b and Xin Sun, 2010, The application of data mining techniques in financial fraud detectie: A classificationframework and an academic review of literature.

A. Patcha, J-M. Park, 2007, An Overview of Anomalie Detectie Techniques: Existing Solutions and Latest Technological Trends, Computer Networks, doi: 10.1016/j.comnet.2007.02.001

Clifton Phua, Vincent Lee, Kate Smith en Ross Gayler, A Comprehensive Survey of Data Mining-based Fraud Detectie Research.

Appendix 2. Toetsen voor huisartsen

Algemene toetsen waarbij resultaat een overzicht met uitkomsten per AGB code is:

1. *Wat is de omzet*
2. *Wat is de omzet per patiënt*
3. *Wat is het aantal prestaties per patiënt*
4. *Wat is het aantal prestaties minus aantal inschrijftarieven (11000, 11100, 11101, 11102, 11103, 11104) per patiënt*

Inschrijftarief

Het inschrijftarief is een tarief per ingeschreven patiënt. Dit mag elk kwartaal in rekening worden gebracht.

Toets 5: Hoe vaak komt het voor dat een bepaalde AGB code meer dan 4 keer in 1 jaar de prestatie 11000⁶⁵ in rekening brengt voor 1 BSN nummer.

Toets 6: Hoe vaak komt het voor dat een bepaalde AGB code meer dan 4 keer in 1 jaar 1 of meerdere van de volgende prestaties (11000, 11100, 11101, 11102, 11103, 11104) in rekening brengt voor 1 BSN nummer.

Toets 7: Hoe vaak komt het voor dat een voor 1 BSN meer dan 4 keer in 1 jaar 1 of meerdere van de volgende prestaties (11000, 11100, 11101, 11102, 11103, 11104) in rekening worden gebracht. Vervolgens als dit het geval is een telling maken hoeveel dit er zijn per BSN en hoe vaak het voorkomt dat het er zoveel zijn.

Toets 8: Hoe vaak komt het voor dat een bepaalde AGB code 4 keer 1 van de volgende prestaties (11000, 11100, 11101, 11102, 11103, 11104) in rekening brengt op 1 datum voor 1 BSN nummer.

Toets 9: Hoe vaak komt het voor dat een aanbieder een inschrijftarief in rekening brengt voor iemand die in een achterstandswijk woont, maar daar in werkelijkheid niet woont. Toets dit door te kijken of bij het in rekening brengen van de prestatie 11102, 11103 en 11104 de patiënt een postcode heeft die valt onder postcodes achterstandswijken⁶⁶.

Ook voor mensen ouder dan 65 of 75 mag een hoger inschrijftarief in rekening worden gebracht.

Toets 10: Hoe vaak komt het voor dat een aanbieder een inschrijftarief in rekening brengt voor iemand die 65 jaar of ouder is, maar dat nog niet is.

Toets 11: Hoe vaak komt het voor dat een aanbieder een inschrijftarief in rekening brengt voor iemand die 75 jaar of ouder is, maar dat nog niet is.

⁶⁵ 11000 is het basisinschrijftarief. Er zijn er echter meer. In principe kunnen er voor 1 BSN maximaal 4 van deze codes per jaar in rekening worden gebracht. Hierbij kan de code wisselen door verhuizing of het passeren van een leeftijdsgrens. Daarom kan bovenstaande toets nogmaals worden uitgevoerd maar dan op meerdere codes.

⁶⁶ We kunnen deze analyse enkel doen op de 4-cijfer postcode omdat we geen andere gegevens hebben. De controle is dan of de postcode van de patiënt voorkomt in bijlage 2 van de beschikking voor huisartsen (<http://www.nza.nl/98174/139255/247115/TB-CU-7009-01-.pdf>) wanneer code 11102, 11103 of 11104 in rekening is gebracht.

Declareren van niet geleverde zorg

Toets 12: Hoeveel procent van de ingeschreven patiënten komt gedurende het jaar ook daadwerkelijk bij de aanbieder.

Upcoding

Ook in de huisartsenzorg zijn er verschillende gradaties in zorg. Deze toetsen richten zich op de verhouding tussen gradaties in zorg.

13. de verhouding (12001)/(12000+12001)

14. de verhouding (12003)/(12002+12003)

15. de verhouding (12111)/(12110+12111)

16. de verhouding (12113)/(12112+12113)

Daarnaast kan je ook van prestatie veranderen

Toets of je van telefonisch consult of emailconsult mogelijk een gewoon consult maakt

17. de verhouding (12000)/(12000+12004+12007)

Toets of je van vaccinatie mogelijk een gewoon consult maakt

18. de verhouding (12000)/(12000+12006)

Dubbele bekostiging

Wanneer er zorg wordt geleverd kan het zo zijn dat dit op meerdere manieren in rekening wordt gebracht. Daarom de volgende toetsen:

Toets 19: Het kan zo zijn dat een M&I verrichting wordt gedaan maar dat er twee dingen in rekening worden gebracht: namelijk een M&I verrichting en een consult. Toets per huisarts/AGB hoe vaak een consult (12000/12001) bij dezelfde patiënt op dezelfde dag in rekening wordt gebracht in combinatie met een prestatie 13xxx.

Toets 20: Het kan zo zijn dat iemand met meerdere zorgvragen bij de huisarts komt en dat deze hiervoor meerdere consulten in rekening brengt. Hij moet er echter 1 in rekening brengen. Toets per huisarts/AGB hoe vaak er meer dan 1 consult (12000/12001) bij dezelfde patiënt op dezelfde dag in rekening wordt gebracht.

Toets 21: In de ketenzorg kunnen patiënten met diabetes, COPD en CVR op meerdere manieren worden bekostigd. Elke patiënt mag echter maar op 1 manier bekostigd worden. Er kan daarom worden gecontroleerd of een bepaalde combinatie van prestatiecodes voorkomt.

- *Toets of code 13029 of 13030 voorkomt bij een BSN in combinatie met code 40001 of 40002. Geef aan voor hoeveel BSN's dit geldt. Tel vervolgens hoe vaak dit voorkomt per AGB code, waarbij de AGB code die de 13029/13030 in rekening brengt leidend is.*
- *Toets of code 13031 voorkomt bij een BSN in combinatie met code 40021 of 40022. Geef aan voor hoeveel BSN's dit geldt. Tel vervolgens hoe vaak dit voorkomt per AGB code, waarbij de AGB code die de 13031 in rekening brengt leidend is.*

Appendix 3. Toetsen voor mondzorg

Toetsen voor mondzorg 2010 & 2011

Vanwege aanpassing voor prestaties in 2012 zijn deze toetsen alleen geschikt voor data over 2010 en 2011. Voor 2012 zijn er specifieke toetsen, zie hieronder.

Algemene toetsen waarbij resultaat een overzicht met uitkomsten per AGB code is:

1. *Wat is de omzet*
2. *Wat is de omzet per patiënt*
3. *Wat is het aantal prestaties per patiënt*
4. *Wat is het aantal dagen waarop iets in rekening is gebracht per patiënt*

De kosten worden vooral gemaakt in een aantal specifieke codes. Daarom aanvullende toetsen:

5. *Wat is aantal V-codes per patiënt (Vxx)*
6. *Wat is aantal C-codes per patiënt (Cxx)*
7. *Wat is aantal M-codes per patiënt (Mxx)*
8. *Wat is aantal X-codes per patiënt (Xxx)*
9. *Wat is de omzet van V-codes per patiënt (Vxx)*
10. *Wat is de omzet van C-codes per patiënt (Cxx)*
11. *Wat is de omzet van M-codes per patiënt (Mxx)*
12. *Wat is de omzet van X-codes per patiënt (Xxx)*
13. *Wat is de omzet van V-codes per patiënt van 17 jaar of jonger (Vxx)*
14. *Wat is de omzet van C-codes per patiënt van 17 jaar of jonger (Cxx)*
15. *Wat is de omzet van M-codes per patiënt van 17 jaar of jonger (Mxx)*
16. *Wat is de omzet van X-codes per patiënt van 17 jaar of jonger (Xxx)*

Declareren van niet geleverde zorg

Geen aanvullende toetsen.

Upcoding

Ook in de mondzorg zijn er verschillende gradaties in zorg. Zo zijn er verschillende gradaties in gebitsreiniging. Het kan zo zijn dat je als aanbieder oneigenlijk vaak de duurste variant in rekening brengt. Vandaar een aantal toetsen voor de verhouding tussen gradaties in zorg.

Gebitsreiniging

Er zijn 3 prestaties

M50 - gebitsreiniging beperkt

M55 - gebitsreiniging gemiddeld

M59 - gebitsreiniging uitgebreid

Breng per aanbieder in kaart:

17. *de verhouding $M59/(M50+M55+M59)$*
18. *de verhouding $(M55+M59)/(M50+M55+M59)$*

Vullingen

Er zijn 6 prestaties: (V10, V11, V12, V13, V14, V15)

Breng per aanbieder in kaart:

- 19. de verhouding $(V15)/(som\ V10, V11, V12, V13, V14, V15)$
- 20. de verhouding $(V14)/(som\ V10, V11, V12, V13, V14, V15)$
- 21. de verhouding $(V14+V15)/(som\ V10, V11, V12, V13, V14, V15)$
- 22. de verhouding $(V13+V14+V15)/(som\ V10, V11, V12, V13, V14, V15)$
- 23. de verhouding $(V12+V13+V14+V15)/(som\ V10, V11, V12, V13, V14, V15)$

Wortelkanaalbehandelingen

Er zijn 4 prestaties: (E13, E14, E16, E17).

Breng per aanbieder in kaart:

- 24. de verhouding $(E17)/(som\ E13, E14, E16, E17)$
- 25. de verhouding $(E16)/(som\ E13, E14, E16, E17)$
- 26. de verhouding $(E16+E17)/(som\ E13, E14, E16, E17)$

Extracties

Er zijn vier prestaties (H10, H15, H30, H35)

- 27. de verhouding $(H30)/(som\ H10, H15, H30, H35)$
- 28. de verhouding $(H35)/(som\ H10, H15, H30, H35)$
- 29. de verhouding $(H30+H35)/(som\ H10, H15, H30, H35)$

Preventieve zorg

Een deel van de mondzorg is preventief. Een tandarts kan er voor kiezen om verhoudingsgewijs veel van deze preventieve zorg te leveren. Enerzijds kan je dit als zeer positief bestempelen, wanneer dit bij deze patiënten leidt tot minder curatieve ingrepen. Anderzijds kan een hoog aandeel preventieve zorg ook een signaal zijn dat een tandarts meer zorg levert dan nodig of dat een deel van de preventieve niet geleverd is. Dit laatste is niet direct uit de data af te leiden, maar het begint met het controleren van het aantal keer dat bepaalde prestaties zijn geleverd per patiënt (voor alle patiënten van 1 aanbieder, voor alle patiënten van 1 aanbieder van 17 jaar of jonger en voor alle patiënten van 1 aanbieder die minimaal 1 keer die prestatie hebben gehad)

- 30. Hoe vaak is er gemiddeld per patiënt per jaar een prestatie in rekening gebracht voor controle (code C11, C12, C13) (voor alle patiënten van 1 aanbieder)
- 31. Hoe vaak is er gemiddeld per patiënt per jaar een prestatie in rekening gebracht voor controle (code C11, C12, C13) (voor alle patiënten van 1 aanbieder die minimaal 1 keer een C11, C12 of C13 hebben gehad)
- 32. Hoe vaak is er gemiddeld per patiënt een prestatie Plaque-score (V31) in rekening gebracht (voor alle patiënten van 1 aanbieder)
- 33. Hoe vaak is er gemiddeld per patiënt een prestatie Plaque-score (V31) in rekening gebracht (voor alle patiënten van 1 aanbieder van 17 jaar of jonger)
- 34. Hoe vaak is er gemiddeld per patiënt een prestatie Plaque-score (V31) in rekening gebracht (voor alle patiënten van 1 aanbieder die minimaal 1 keer een V31 hebben gehad)
- 35. Hoe vaak is er gemiddeld per patiënt een röntgenfoto (X10) in rekening gebracht (voor alle patiënten van 1 aanbieder)
- 36. Hoe vaak is er gemiddeld per patiënt een röntgenfoto (X10) in rekening gebracht (voor alle patiënten van 1 aanbieder die minimaal 1 keer een VX10VX10 hebben gehad)
- 37. Hoe vaak is er gemiddeld per patiënt een fluoride applicatie (M10, M20, M21) in rekening gebracht (voor alle patiënten van 1 aanbieder)
- 38. Hoe vaak is er gemiddeld per patiënt een fluoride applicatie (M10, M20, M21) in rekening gebracht (voor alle patiënten van 1 aanbieder die minimaal 1 keer een M10, M20 of M21 hebben gehad)

39. *Hoe vaak is er gemiddeld per patiënt een sealing (V30, V35) in rekening gebracht (voor alle patiënten van 1 aanbieder)*
40. *Hoe vaak is er gemiddeld per patiënt een sealing (V30, V35) in rekening gebracht (voor alle patiënten van 1 aanbieder van 17 jaar of jonger)*
41. *Hoe vaak is er gemiddeld per patiënt een sealing (V30, V35) in rekening gebracht (voor alle patiënten van 1 aanbieder die minimaal 1 keer een V30V30 of V35 hebben gehad)*

Onwaarschijnlijk aantal vullingen / extracties / wortelkanaalbehandelingen / röntgenfoto's

De hoeveelheid behandelingen in de mond is gelimiteerd. Indien er heel vaak een bepaalde behandeling wordt gedaan voor een patiënt dan kan dat wijzen op het in rekening brengen van niet geleverde zorg. Daarom de vraag om het onderstaande in kaart te brengen:

42. *Breng per patiënt per datum het aantal keer dat een vulling (V10, V11, V12, V13, V14, V15) in rekening is gebracht in kaart in een lijst met het hoogste aantal bovenaan.*
43. *Breng per patiënt per datum het aantal keer dat een wortelkanaalbehandeling (E13, E14, E16, E17) in rekening is gebracht in kaart in een lijst met het hoogste aantal bovenaan.*
44. *Breng per patiënt per datum het aantal keer dat een extractie (H10, H15, H30, H35) in rekening is gebracht in kaart in een lijst met het hoogste aantal bovenaan.*
45. *Breng per patiënt per datum het aantal keer dat een röntgenfoto (X10, X21, X22, X24) in rekening is gebracht in kaart in een lijst met het hoogste aantal bovenaan.*

Vervolgens in de tweede stap willen we graag weten bij welke aanbieders deze patiënten dan zitten. Dit kan per record worden aangegeven, maar vervolgens is het ook goed om een telling te maken. Dit kan voor de top 100, voor de top 1000 of bijvoorbeeld voor de top x waarbij je x bepaalt met gemiddelde + 2 keer de standaarddeviatie. Het resultaat op basis van top 100 is in ieder geval als volgt

46. *Maak op basis van bovenstaande bevindingen van de top 100 een overzicht van het aantal keer dat een bepaalde agb code indiener voor komt in deze top 100.*

Hetzelfde kan worden gedaan per jaar ipv per dag

47. *Breng per patiënt per jaar het aantal keer dat een vulling (V10, V11, V12, V13, V14, V15) in rekening is gebracht in kaart.*
48. *Breng per patiënt per jaar het aantal keer dat een wortelkanaalbehandeling (E13, E14, E16, E17) in rekening is gebracht in kaart.*
49. *Breng per patiënt per jaar het aantal keer dat een extractie (H10, H15, H30, H35) in rekening is gebracht in kaart.*
50. *Breng per patiënt per jaar het aantal keer dat een röntgenfoto (X10, X21, X22, X24) in rekening is gebracht in kaart.*
51. *Maak op basis van bovenstaande bevindingen van de top 100 een overzicht van het aantal keer dat een bepaalde agb code indiener voor komt in deze top 100.*

Leeftijdsgrens

Er zijn een aantal behandelingen die je niet verwacht bij kinderen. Indien er toch vaak een declaratie is met deze prestatie voor een kind dan kan dat wijzen op fraude. Een kind definiëren we als iemand van 17 jaar of jonger. Hiervoor zijn de volgende toetsen:

- 52. Het aantal keer per jaar dat een aanbieder een wortelkanaalbehandeling (E13, E14, E16, E17) heeft uitgevoerd bij een patiënt van 17 jaar of jonger*
- 53. Het aantal keer per jaar dat een aanbieder een kroon (V14, R20, R25, R27, R26, R28, R29) heeft uitgevoerd bij een patiënt van 17 jaar of jonger*
- 54. Het aantal keer per jaar dat een aanbieder een prothese heeft gedaan (P10, P15, P21, P25, P30, P34, P35) heeft uitgevoerd bij een patiënt van 17 jaar of jonger*
- 55. Het aantal keer per jaar dat een aanbieder orale implantalogie (Jxx) heeft uitgevoerd bij een patiënt van 17 jaar of jonger*

Dubbele bekostiging

Wanneer er zorg wordt geleverd, kan het zo zijn dat dit meer dan 1 keer in rekening wordt gebracht. Daarom de volgende toetsen.

Hoe vaak komt het voor dat er op 1 dag voor patiënt meer dan 1 keer de volgende code(s) in rekening wordt/worden gebracht:

- 56. code C11, C12, C13*
- 57. code M50, M55, M59*

Toets 56 tot en met 61 zien er voor 2012 iets anders uit vanwege nieuwe prestatiecodes in 2012. In grote lijnen zijn de toetsen hetzelfde als in 2010 en 2011 maar met andere velden.

Het kan ook voorkomen dat een rekening op verschillende manieren wordt ingediend. Bijvoorbeeld door zowel aanbieder als patiënt, of door zowel factoringmaatschappij als aanbieder. Je kunt dit controleren door te kijken naar identieke rekeningen die enkel verschillen bij code indiener.

- 58. Hoe vaak komt het voor per aanbieder dat er 2 records met dezelfde BSN, prestatie en datum in de dataset zitten die verschillen in code indiener en/of agb code indiener?*

Ongeldige combinaties⁶⁷

Het is niet toegestaan bepaalde combinaties samen in rekening te brengen.

- 59. Hoe vaak per aanbieder krijgt een patiënt op 1 dag zowel een C22 als een C11 in rekening gebracht.*
- 60. Hoe vaak per aanbieder krijgt een patiënt op 1 dag zowel een C28 als een C11 in rekening gebracht.*
- 61. Hoe vaak per aanbieder krijgt een patiënt op 1 dag zowel een C29 als een C11 in rekening gebracht.*
- 62. Som van de drie hierboven.*

⁶⁷ Deze lijst van ongeldige combinaties is niet uitputtend. In de loop der jaren zal deze lijst groeien, wijzigen en aangepast worden.

Onterecht ANW toeslag

Wanneer een tandarts zorg levert in avond, nacht of weekend dan mag er een toeslag in rekening worden gebracht. Hierop de volgende toetsen:

- 63. *Hoe vaak heeft een aanbieder C85 in rekening gebracht in verhouding tot het totaal aantal prestaties?*
- 64. *Hoe vaak heeft een aanbieder C86 in rekening gebracht in verhouding tot het totaal aantal prestaties?*
- 65. *Hoe vaak heeft een aanbieder C87 in rekening gebracht in verhouding tot het totaal aantal prestaties?*
- 66. *Som van bovenstaande drie toetsen.*
- 67. *Hoe vaak heeft een aanbieder C85 in rekening gebracht op een datum die niet behoorde tot het weekend (weekend definiëren als enkel de zaterdag en zondag)*

Toetsen voor mondzorg 2012

Algemene toetsen waarbij resultaat een overzicht met uitkomsten per AGB code is:

1. *Wat is de omzet*
2. *Wat is de omzet per patiënt*
3. *Wat is het aantal prestaties per patiënt*
4. *Wat is het aantal dagen waarop iets in rekening is gebracht per patiënt*

De kosten worden vooral gemaakt in een aantal specifieke codes. Daarom aanvullende toetsen

5. *Wat is aantal E-codes (vullingen, kronen en bruggen) per patiënt (Exx)?*
6. *Wat is aantal A-codes (consulten en diagnostiek) per patiënt (Axx)?*
7. *Wat is aantal C-codes (preventieve mondzorg) per patiënt (Cxx)?*
8. *Wat is de omzet van E-codes per patiënt (Exx)?*
9. *Wat is de omzet van A-codes per patiënt (Axx)?*
10. *Wat is de omzet van C-codes per patiënt (Cxx)?*
11. *Wat is de omzet van E-codes per patiënt van 17 jaar of jonger (Exx)?*
12. *Wat is de omzet van A-codes per patiënt van 17 jaar of jonger (Axx)?*
13. *Wat is de omzet van C-codes per patiënt van 17 jaar of jonger (Cxx)?*

Declareren van niet geleverde zorg

Geen aanvullende toetsen tot dusver

Upcoding

Ook in de mondzorg zijn er verschillende gradaties in zorg. Zo heb je verschillende gradaties in gebitsreiniging. Het kan zo zijn dat je als aanbieder oneigenlijk vaak de duurste variant in rekening brengt. Vandaar een aantal toetsen voor de verhouding tussen gradaties in zorg.

Gebitsreiniging

Er zijn 2 prestaties

C112 Preventieve voorlichting en/of instructie standaard

C114 Preventieve voorlichting en/of instructie uitgebreid

Breng per aanbieder in kaart:

14. *de verhouding $C114/(C112+C114)$ in een lijst met het hoogste aandeel bovenaan*

Vullingen

Er zijn 4 prestaties: (E111, E112, E113, E114)

Breng per aanbieder in kaart:

15. *de verhouding $(E114)/(som\ E111, E112, E113, E114)$ in een lijst met het hoogst aandeel bovenaan*
16. *de verhouding $(E113)/(som\ E111, E112, E113, E114)$ in een lijst met het hoogst aandeel bovenaan*
17. *de verhouding $(E113+E114)/(som\ E111, E112, E113, E114)$ in een lijst met het hoogst aandeel bovenaan*
18. *de verhouding $(E112+E113+E114)/(som\ E111, E112, E113, E114)$ in een lijst met het hoogst aandeel bovenaan*

Wortelkanaalbehandelingen

Er zijn meerdere prestaties maar we focussen ons op 4 prestaties: (L311, L312, L313, L314).

Breng per aanbieder in kaart:

19. *de verhouding (L314)/(som L311, L312, L313, L314) in een lijst met het hoogst aandeel bovenaan*
20. *de verhouding (L313)/(som L311, L312, L313, L314) in een lijst met het hoogst aandeel bovenaan*
21. *de verhouding (L313+L314)/(som L311, L312, L313, L314) in een lijst met het hoogst aandeel bovenaan*

Extracties

Er zijn 2 prestaties (J311 en J315)

22. *de verhouding (J315)/(som J311, J315) in een lijst met het hoogst aandeel bovenaan*

Preventieve zorg

Een deel van de mondzorg is preventief. Een tandarts kan er voor kiezen om verhoudingsgewijs veel van deze preventieve zorg te leveren. Enerzijds kan je dit als zeer positief bestempelen, wanneer dit bij deze patiënten leidt tot minder curatieve ingrepen. Anderzijds kan een hoog aandeel preventieve zorg ook een signaal zijn dat een tandarts meer zorg levert dan nodig of dat een deel van de preventieve niet geleverd is. Dit laatste is niet direct uit de data af te leiden, maar het begint met het controleren van het aantal keer dat bepaalde prestaties zijn geleverd per patiënt (voor alle patiënten van 1 aanbieder, voor alle patiënten van 1 aanbieder van 17 jaar of jonger en voor alle patiënten van 1 aanbieder die minimaal 1 keer die prestatie hebben gehad)

23. *Hoe vaak is er gemiddeld per patiënt per jaar een prestatie in rekening gebracht voor controle (code A111, A211, A231, A261) (voor alle patiënten van 1 aanbieder)*
24. *Hoe vaak is er gemiddeld per patiënt per jaar een prestatie in rekening gebracht voor controle (code A111, A211, A231, A261) (voor alle patiënten van 1 aanbieder die minimaal 1 keer een A111, A211, A231 of A261 hebben gehad)*
25. *Hoe vaak is er gemiddeld per patiënt een röntgenfoto (A311) in rekening gebracht (voor alle patiënten van 1 aanbieder)*
26. *Hoe vaak is er gemiddeld per patiënt een röntgenfoto (A311) in rekening gebracht (voor alle patiënten van 1 aanbieder die minimaal 1 keer een A311 hebben gehad)*
27. *Hoe vaak is er gemiddeld per patiënt een fluoridebehandeling (C611, C811) in rekening gebracht (voor alle patiënten van 1 aanbieder)*
28. *Hoe vaak is er gemiddeld per patiënt een fluoride applicatie (C611, C811) in rekening gebracht (voor alle patiënten van 1 aanbieder die minimaal 1 keer een (C611, C811) hebben gehad)*
29. *Hoe vaak is er gemiddeld per patiënt een sealing (C511) in rekening gebracht (voor alle patiënten van 1 aanbieder)*
30. *Hoe vaak is er gemiddeld per patiënt een sealing (C511) in rekening gebracht (voor alle patiënten van 1 aanbieder van 17 jaar of jonger)*
31. *Hoe vaak is er gemiddeld per patiënt een sealing (C511) in rekening gebracht (voor alle patiënten van 1 aanbieder die minimaal 1 keer een C511 hebben gehad)*

Onwaarschijnlijk aantal vullingen / extracties / wortelkanaalbehandelingen / röntgenfoto's

De hoeveelheid behandelingen in de mond is gelimiteerd. Indien er heel vaak een bepaalde behandeling wordt gedaan voor een patiënt dan kan dat wijzen op het in rekening brengen van niet geleverde zorg. Daarom de vraag om het onderstaande in kaart te brengen:

- 32. Breng per patiënt per datum het aantal keer dat een vulling (E111, E112, E113, E114) in rekening is gebracht in kaart in een lijst met het hoogste aantal bovenaan.*
- 33. Breng per patiënt per datum het aantal keer dat een wortelkanaalbehandeling (L311, L312, L313, L314) in rekening is gebracht in kaart in een lijst met het hoogste aantal bovenaan.*
- 34. Breng per patiënt per datum het aantal keer dat een extractie (J311, J315) in rekening is gebracht in kaart in een lijst met het hoogste aantal bovenaan.*
- 35. Breng per patiënt per datum het aantal keer dat een röntgenfoto (A311, A321, A324, A327) in rekening is gebracht in kaart in een lijst met het hoogste aantal bovenaan.*

Vervolgens in de tweede stap willen we graag weten bij welke aanbieders deze patiënten dan zitten. Dit kan per record worden aangegeven, maar vervolgens is het ook goed om een telling te maken. Dit kan voor de top 100, voor de top 1000 of bijvoorbeeld voor de top x waarbij je x bepaalt met gemiddelde + 2 keer de standaarddeviatie. Het resultaat op basis van top 100 is in ieder geval als volgt

- 36. Maak op basis van bovenstaande bevindingen van de top 100 een overzicht van het aantal keer dat een bepaalde agb code indiener voor komt in deze top 100. Dit resulteert in een lijst met het hoogste aantal bovenaan.*

Hetzelfde kan worden gedaan per jaar ipv per dag

- 37. Breng per patiënt per jaar het aantal keer dat een vulling (E111, E112, E113, E114) in rekening is gebracht in kaart in een lijst met het hoogste aantal bovenaan.*
- 38. Breng per patiënt per jaar het aantal keer dat een wortelkanaalbehandeling (L311, L312, L313, L314) in rekening is gebracht in kaart in een lijst met het hoogste aantal bovenaan.*
- 39. Breng per patiënt per jaar het aantal keer dat een extractie (J311, J315) in rekening is gebracht in kaart in een lijst met het hoogste aantal bovenaan.*
- 40. Breng per patiënt per jaar het aantal keer dat een röntgenfoto (A311, A321, A324, A327) in rekening is gebracht in kaart in een lijst met het hoogste aantal bovenaan.*

Vervolgens in de tweede stap willen we graag weten bij welke aanbieders deze patiënten dan zitten. Dit kan per record worden aangegeven, maar vervolgens is het ook goed om een telling te maken. Dit kan voor de top 100, voor de top 1000 of bijvoorbeeld voor de top x waarbij je x bepaalt met gemiddelde + 2 keer de standaarddeviatie. Het resultaat op basis van top 100 is in ieder geval als volgt

- 41. Maak op basis van bovenstaande bevindingen van de top 100 een overzicht van het aantal keer dat een bepaalde agb code indiener voor komt in deze top 100. Dit resulteert in een lijst met het hoogste aantal bovenaan.*

Leeftijdsgrens

Er zijn een aantal behandelingen die je niet verwacht bij kinderen. Indien er toch vaak een declaratie is met deze prestatie voor een kind dan kan dat wijzen op fraude. De resultaten hiervan vragen we zowel in absolute aantallen als in percentage van het aantal kinderen dat een dergelijke behandeling heeft gehad ten opzichte van het totaal aantal kinderen dat door een tandarts is geholpen in een jaar. Een kind definiëren we als iemand van 17 jaar of jonger. Hiervoor zijn de volgende toetsen:

- 42. Het aantal keer per jaar dat een aanbieder een wortelkanaalbehandeling (L311, L312, L313, L314) heeft uitgevoerd bij een patiënt van 17 jaar of jonger*
- 43. Het aantal keer per jaar dat een aanbieder een kroon (E321, E331, E711, E721) heeft uitgevoerd bij een patiënt van 17 jaar of jonger*
- 44. Het aantal keer per jaar dat een aanbieder een prothese heeft gedaan (G111, G112, G113, G211, G251, G253) heeft uitgevoerd bij een patiënt van 17 jaar of jonger*
- 45. Het aantal keer per jaar dat een aanbieder orale implantologie (Pxx) heeft uitgevoerd bij een patiënt van 17 jaar of jonger*

Dubbele bekostiging

Wanneer er zorg wordt geleverd kan het zo zijn dat dit meer dan 1 keer in rekening wordt gebracht. Daarom de volgende toetsen.

Hoe vaak komt het voor dat er op 1 dag voor patiënt meer dan 1 keer de volgende code(s) in rekening wordt/worden gebracht:

- 46. code A111, A211, A231, A261*
- 47. code C112, C114*

Het kan ook voorkomen dat een rekening op verschillende manieren wordt ingediend. Bijvoorbeeld door zowel aanbieder als patiënt, of door zowel factoringmaatschappij als aanbieder. Je kunt dit controleren door te kijken naar identieke rekeningen die enkel verschillen bij code indiener.

- 48. Hoe vaak komt het voor per aanbieder dat er 2 records met dezelfde BSN, prestatie en datum in de dataset zitten die verschillen in code indiener en/of agb code indiener?*

Ongeldige combinaties

Het is niet toegestaan bepaalde combinaties samen in rekening te brengen.

Mogelijk zijn er nog meer ongeldige combinaties. Later kan deze lijst nog aangevuld worden.

Onterecht ANW toeslag

Wanneer een tandarts zorg levert in avond, nacht of weekend dan mag er een toeslag in rekening worden gebracht. Hierop de volgende toetsen:

- 49. Hoe vaak heeft een aanbieder X111 in rekening gebracht in verhouding tot het totaal aantal prestaties? Uitkomst lijst van verhouding per aanbieder met het hoogste aantal bovenaan*

Appendix 4. Toetsen voor farmacie

Algemene toetsen waarbij het resultaat een overzicht is met uitkomsten per AGB (praktijk, voorschrijver, afleveraar) code

1. Wat is de omzet
2. Wat is het aantal patienten
3. Wat is de omzet per patiënt
4. Wat is de omzet per declaratie
5. Wat is het aantal declaraties per patiënt

Algemene toetsen waarbij het resultaat een overzicht is met uitkomsten per patiënt (BSN)

6. Wat is de omzet
7. Wat is het aantal prestaties
8. Wat is het aantal unieke AGB-codes praktijken
9. Wat is het aantal unieke AGB-codes voorschrijvers
10. Wat is het aantal unieke AGB-codes afleveraars
11. Wat is het aantal unieke data van prestaties

Registratie en declaratie

Het kan voorkomen dat de toeslag voor eerste uitgifte in rekening wordt gebracht terwijl het middel in de 12 maanden daarvoor al is verstrekt. Ook voor hetzelfde middel maar gewijzigde dosering, dus een ander artikelnummer geldt dat een toeslag voor eerste uitgifte niet opnieuw in rekening gebracht mag worden. De lijst met artikelnummers per stofnaam biedt hiervoor de ruimte omdat alle doseringen daar in staan.

12. Per apotheek, het aantal declaraties met een eerste ter handstelling terwijl een medicijn met hetzelfde artikelnummer in de 12 maanden daarvoor al een keer is verstrekt aan dezelfde persoon. Declaraties met WMG-codes 7, 8, 9, 10, 11, 12, 19, 20, 21, 22, 23, 24, 31, 32, 33, 34, 35, 36, 43, 44, 45, 46, 47 en 48 zijn met een toeslag eerste te handstelling.

13. Per apotheek, het aantal declaraties met een eerste ter handstelling terwijl een medicijn met hetzelfde ATC-code en toedieningsweg-code in de 330 dagen daarvoor al een keer is verstrekt aan dezelfde persoon. Declaraties met WMG-codes 7, 8, 9, 10, 11, 12, 19, 20, 21, 22, 23, 24, 31, 32, 33, 34, 35, 36, 43, 44, 45, 46, 47 en 48 zijn met een toeslag eerste te handstelling.

Als er voor een medicijn met hetzelfde ATC-code meerdere keren in de week een week te handstelling wordt gedeclareerd kan het duiden op fraude.

14. Per apotheek, het aantal declaraties met een week ter handstelling terwijl een artikel met hetzelfde ATC-code en toedieningsweg-code binnen die week al een keer is verstrekt aan dezelfde persoon. Declaraties met WMG-codes 4, 5, 6, 10, 11, 12, 16, 17, 18, 22, 23, 24, 28, 29, 30, 34, 35, 36, 40, 41, 42, 46, 47, 48 zijn met een toeslag week ter handstelling.

Een apotheek kan een standaard uitgifte als uitgifte met magistrale bereiding declareren of reguliere magistrale bereiding als bijzondere bereiding om daarmee hoger tarief te declareren.

15. Verhouding geen magistrale bereiding, magistrale bereiding en bijzondere magistrale bereiding per AGB. WMG-codes 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32, 35, 38, 41, 44 en 47 zijn een magistrale bereidingen en WMG-code 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33, 36, 39, 42, 45 en 48 zijn bijzondere magistrale bereidingen.

Bij een magistrale bereiding gebruikt de apotheker grondstoffen om de medicijn te bereiden. Daarom moet de ATC-code van de desbetreffende declaratie een grondstof aangeven. Als dit niet zo is, zou het kunnen zijn dat de apotheker een magistrale bereiding declareert terwijl hij alleen een medicijn heeft overhandigd.

- 16. Per apotheek, per artikelnummer, het aantal declaraties met een magistrale bereiding relatief ten opzichte van de gemiddelde over alle apotheken.*

Een apotheek kan een standaard uitgifte als weekend uitgifte declareren om daarmee een hoger tarief te declareren.

- 17. Per apotheek, de verhouding van aantal met een toeslag ANZ (Avond Nacht Zondag) t.o.v. totaal aantal prestaties. Declaraties met WMG-codes 13 t/m 24 en 37 t/m 48 zijn met een toeslag ANZ.*

Dubbele bekostiging

Het kan voorkomen dat een rekening op verschillende manieren wordt ingediend. Bijvoorbeeld door zowel apotheker als patiënt, of door zowel factoringmaatschappij als apotheker. Je kunt dit controleren door te kijken naar identieke rekeningen die enkel verschillen bij code indiener.

- 18. Hoe vaak komt het voor per apotheker dat er 2 records met dezelfde BSN, artikelnummer en datum in de dataset zitten die verschillen in code indiener en/of agb code indiener.*

Het kan voorkomen dat als iemand een hormoon-spiraal wil laten zetten in het ziekenhuis dat de specialist de patiënt naar het ziekenhuisapothek stuur om een spiraal te halen. Dit is echter niet toegestaan, want de spiraal is ook vergoed door de DBC prestatie die de specialist declareert. De Z-index nummers voor de Mirena-spiralen zijn te vinden in het Z-index-bestand.

- 19. Per apotheek, het aantal declaraties van Mirena-spiralen. In het bijzonder het aantal van ziekenhuisapotheken.*

- 20. Het aantal declaraties van Mirena-spiraal die niet gevolgd zijn door een huisarts consult binnen 2 maanden.*

- 21. Neem de subset met alle declaraties waarvoor de AGB-code van de voorschrijver een medisch specialist (agb begint met 03) of een ziekenhuis (agb begint met 06) is. Bepaal de aantal declaraties van Mirena-spiraal per voorschrijver.*

Doorverkopen van medicijnen

Medicijnen in de categorie benzodiazepinen, opiaten en, hypnotica en slapeloosheid zijn middelen die makkelijk doorverkocht kunnen worden. Het kan voorkomen dat een patiënt met een kopie van een recept of in samenwerking met de huisarts vaak een medicijn krijgt en deze dan doorverkoopt. Het kunnen ook verslaafden zijn die steeds de medicijnen bij de apotheek haalt.

- 22. Per patiënt, het bedrag dat aan een specifieke categorie van verslavende medicijnen is gedeclareerd relatief ten opzichte van de gemiddelde over alle patiënten. De lijst met de artikelnummer van de verslavende middelen staat per categorie in het document "Geneesmiddel analyse chronisch en verslavend.xlsx".*

Ook het aantal unieke apotheken per patiënt kan duiden op personen die de medicijnen kopen om door te verkopen. (toets 6)

Appendix 5. Toetsen voor GGZ

Toets 5, 6, 7 en 9 moeten per soort behandeling apart berekend worden omdat de verhoudingen per instelling nogal kunnen verschillen door de samenstelling van de aandoeningen die worden behandeld bij de instelling.

Bij sommige toetsen is het handig om per productgroep de resultaten te berekenen, omdat de verhoudingen per instelling nogal kan verschillen door de samenstelling van de aandoeningen die worden behandeld bij de instelling. Bij het uitvoeren van de toetsen is het ook handig om de volgende categorieën te maken: psychologen (agb begint met 94 of 95), psychiaters (agb begint met 03), instellingen, (agb begint met 62).

Algemene toetsen waarbij resultaat een overzicht met uitkomsten per AGB code is:

1. *Wat is de omzet*
2. *Wat is de omzet per patiënt*
3. *Wat is het aantal prestaties per patiënt*

Zorglevering

Patiënten kunnen onnodig met verblijf worden behandeld.

4. *Verhouding van het aantal declaraties met verblijf t.o.v. alle declaraties per AGB. De 7 tot 9de cijfers van de prestatiecode geven aan of een prestatie met verblijf of zonder verblijf is. Als de cijfers 000 zijn dan is het een prestatie zonder verblijf anders is het een prestatie met verblijf.*

Voor bepaalde groepen, zoals de vrijgevestigde psychologen en psychiaters, verwachten we dat er minder declaraties met verblijf zijn, waardoor het goed is om naar de verhouding tussen het aantal met en zonder verblijf expliciet voor deze groep te bekijken.

5. *Dezelfde test als toets 4, maar specifiek voor vrijgevestigde psychologen en psychiaters.*

Patiënten kunnen onnodig in een te zware verblijfsafdeling geplaatst worden om een hogere vergoeding te krijgen.

6. *Verhouding gradaties verblijf per AGB. Er zijn 5 gradaties. De toets zou uitgevoerd kunnen worden op de hoogste 3 gradaties. De 9^e cijfer van de prestatiecode geeft de gradatie van het verblijf aan.*

Registratie en declaratie

Het opknippen van lange DBC's in kortere DBC's om een hogere vergoeding te krijgen.

7. *Per AGB-code, de gemiddelde aantal prestaties per jaar per patiënt per productgroep.*

Er kan meer gedeclareerd worden dan dat aan zorg geleverd is.

8. *Per AGB-code, het aantal declaraties waarvoor het aantal dagen dat zorg geleverd is (einddatum-begindatum) minder is dan de minimum aantal dagen van de gedeclareerde prestatie.*

Van dezelfde hoofddiagnose/productgroep mag er maar 1 DBC open staan per patiënt.

9. *Per AGB-code, het aantal patiënten met meer dan 1 DBC per productgroep gedurende dezelfde periode.*

Per patiënt mag er maar 3 DBC's open staan per patiënt.

10. *Per AGB-code, het aantal patiënten met meer dan 3 DBC's gedurende dezelfde periode.*

Het verkorten van DBC's en dan doorsturen naar aanbieders in een netwerk.

11. *Neem per jaar alle patiënten die bij meer dan 1 aanbieder een declaratie heeft. Bepaal voor deze subset voor alle aanbieders, de percentage patiënten die ook een declaratie heeft van ieder ander aanbieder.*

In overlappende periodes bij twee aanbieders in een netwerk dezelfde behandeling declareren.

12. *Dezelfde test als test 10. En bepaal ook het aantal patiënten die een overlappende behandelperiode hebben voor de zelfde hoofddiagnose/productgroep.*

Vanaf 2012 is de behandeling van aanpassingsstoornis onverzekerde zorg geworden. Hierdoor kan het gebeuren dat patiënten die voor angststoornissen behandeld werden nu onder een andere aandoening geregistreerd worden zodat de zorg wel via de zorgverzekeraar wordt vergoed.

13. *Neem alle patiënten die in 2011 een declaratie hadden met de hoofdproductgroep (4^{de} -6^{de} cijfer van de prestatiecode) aanpassingsstoornissen en geen andere declaraties had met een andere hoofdproductgroep. Bepaal voor deze groep patiënten, per AGB-code, voor 2012 het aantal patiënten per hoofdproductgroep.*

14. *Neem alle AGB-codes die in 2011 een declaratie hadden met de hoofdproductgroep aanpassingsstoornissen. Bepaal voor deze groep AGB-codes, per AGB-code, per hoofdproductgroep, de ratio tussen het aantal patiënten in 2012 en het aantal patiënten in 2011. En deze ratio vergelijken met de Nederlandse gemiddelde.*

Upcoding

Opzettelijk een uurtje meer leveren of registreren.

15. *Verhouding tijd categorieën per AGB per productgroep. De toets uitvoeren op de 3 categorieën met de meeste tijd.*

Leeftijdsgrens

Er zijn een aantal behandelingen die je minder verwacht bij kinderen. Indien er toch vaak een declaratie is met deze prestatie voor een kind dan kan dat wijzen op fraude.

16. *Maak een landelijke top 5 of top 10 van diagnoses die het minst bij kinderen voorkomen. Bepaal het aantal keer per jaar dat een aanbieder deze diagnoses heeft uitgevoerd bij een patiënt van 4 jaar of jonger.*
17. *Dezelfde test maar bij patiënten jonger dan 17 jaar.*

Appendix 6. Toelichting op de data

6.1 Algemeen

Door de NZa zijn gegevens opgevraagd bij Vektis over de jaren 2010 t/m 2012. Het ging hierbij om zeer grote aantallen records. De gegevens zijn afkomstig van de zorgverzekeraars in Nederland en betreffende de zorgsegmenten:

- Huisartsen
- Mondzorg
- Farmacie
- GGZ (periode 2009 – 2011)

Tabel 1 geeft een overzicht van de omvang van de bestanden per onderzoek segment.

De data zijn gecontroleerd op volledigheid en op de kwaliteit van de inhoud. Daaruit bleek dat er soms veel ontbrekende waarden zijn, onjuiste en/of ongeldige waarden zijn en niet consistente waarden. In een klein aantal gevallen ontbraken ook *key*-variabelen zoals prestatiecode. Onderzocht is of het ontbreken of de incorrectheid verdacht is en al gelijk leidt tot het aanmerken als een anomalie. Dat bleek niet het geval te zijn.

Jaar	Huisartsen		Mondzorg		Farmacie		GGZ	
	Aantal records	Aantal AGB-beh.	Aantal records	Aantal AGB-beh.	Aantal records	Aantal AGB-afl.	Aantal records	Aantal AGB-beh.
2009							1.165.353	2.758
2010	262.584.340	14.950	69.297.896	15.190	191.744.461	6.790	1.218.992	3.506
2011	304.654.670	17.052	68.449.999	15.550	208.515.505	6.909	1.251.854	4.119
2012	313.926.643	24.532	51.934.447	15.486	219.200.187	6.779		
Totaal	881.165.653		189.682.342		619.460.153		3.636.199	

Tabel 1. Overzicht van de data en de zorgsegmenten

Alles overziende vormen de data een basis voor het uitvoeren van anomalie detectie. Voor vervolgonderzoek is het verstandig een aantal punten van zorg die de kwaliteit van de data betreffen, op te lossen.

6.2 Huisartsen

De volgende dertien variabelen zijn in het data bestand aanwezig:

1. uzovi
2. bsn_encrypt
3. declajaar
4. declakwartaal
5. leeftijd_1jul
6. geslacht
7. srtvzk
8. openjaar
9. openmaand
10. opendag
11. agb_indiener
12. agb_behandelaar
13. prestcode
14. aantprest
15. declaratiebedrag
16. geen_vzk

Het gaat om declaratierecords dat wil zeggen dat deze declaraties vergoed zijn. In een aantal gevallen is er sprake van creditrecords, dit zijn correctieboekingen. Soms worden lege velden en een aantal missing values aangetroffen. Ook lijkt er bij de gegevens over 2010 een deel te ontbreken (circa 10%). Onderzocht is of de afwezigheid van deze gegevens argwaan oproept d.w.z. of er een mogelijk verband is met onrechtmatigheid en fraude; dit is niet gevonden. Op het grote aantal declaraties is de relatieve omvang van deze ontbrekende waarden echter beperkt.

Door te aggregeren kan inzicht per verzekerde worden verkregen, net als inzicht per huisarts, per praktijk, per prestatiecode, per kwartaal, per jaar, etc., etc.

Het aantal records voor dit onderzoek is groot, in totaal 881 miljoen. Het betreft vergoedingen voor een totaal van 7,5 miljard euro in de jaren 2010 – 2012.

De verdeling van het aantal record over deze jaren wordt in tabel 2.1 weergegeven. De indruk bestaat dat in 2010 een deel van de gegevens ontbreekt. Voor dit ontbrekende deel (het betreft één uzovi waarvan bekend is dat deze geen gegevens aan Vektis heeft aangeleverd) is bijgeschat bij de berekening van de totalen van het huisartsensegment in 2010.

Jaartal	Aantal records
2010	262.584.340
2011	304.654.670
2012	313.926.643
Totaal	881.165.653

Tabel 1. Aantal records betreffende vergoedingen aan huisartsen in de jaren 2010 – 2012

Resumerend:

Aandacht is nodig

- voor het ontbrekende deel in het data bestand van 2010, circa 30 miljoen records.
- voor ontbrekende waarden op *key*-variabelen zoals prestatiecode, AGB-behandelaar en BSN
- voor het kunnen beschikken over meer gedetailleerde postcodedata om de toets over achterstandswijken uit te kunnen voeren. Daarvoor ontbreken nu de benodigde gegevens.

6.3 Mondzorg

De volgende variabelen zijn beschikbaar voor de jaren 2010, 2011 en 2012.

1. UZOVICOD
2. bsn_encrypt
3. declajaar
4. declakwartaal
5. leeftijd_1jul
6. GSL
7. SRTVZK
8. openjaar
9. openmaand
10. opendag
11. agb_indiener
12. INDIENCOD
13. agb_behandelaar
14. SPECMEBEHAND
15. agb_voorschrijver
16. SPECMEVSR
17. SRTPREST
18. PRESTLST
19. PRESTCOD
20. INDBIJZTAND
21. SRTBIJZTAND
22. BOVENONDER
23. ELEMENTCOD
24. VLAKCOD
25. DIAGLST
26. DIAGBIJZTAND
27. INDONGEVAL
28. AANTPREST
29. BERBDRGBTW
30. DECLBDRGBTW
31. TOEGBDRGBTW
32. BTWPERC
33. TARPRESTBTW
34. MINBDRGBIJZTAND
35. REFNR
36. REFNRVRG
37. elis_tarprestbtw
38. elis_berek_bedrag
39. INDDBTCDTBEREKEN
40. INDDBTCDTDECLA
41. INDDBTCDTTOEG
42. ELIS_INDDBTCDTTAR
43. ELIS_INDDBTCDTBEREKEN
44. geen_vzk

Het aantal onbekenden op de variabele AGB-behandelaar is wel erg hoog: de code "0" werd 2010 3.128.009, in 2011 3.365.689 en in 2012 3.516.310 maal aangetroffen. Hoewel deze aantallen procentueel klein zijn, betreft het hier een key-variabele. Wanneer mondzorg heeft plaatsgevonden in het buitenland zal deze variabele doorgaans onbekend zijn.

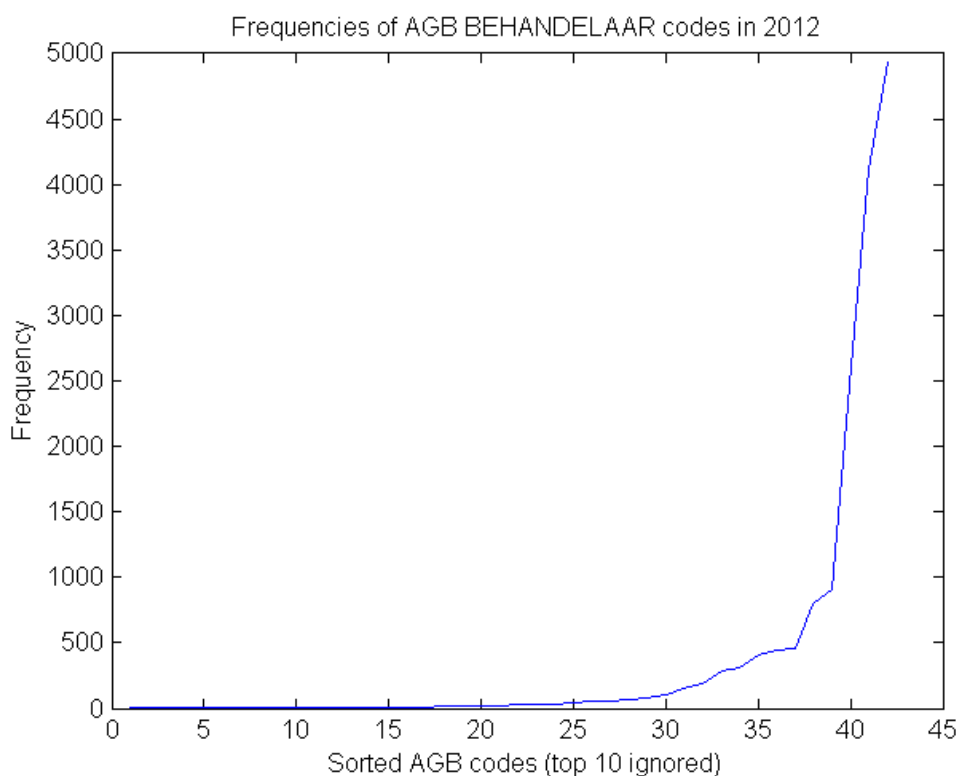
Bij de variabele "indiener" treffen we nog veel meer missing values (waarde "0", "12345678" en "99999999") aan, circa 20 miljoen records voor elk van de jaren. Wanneer de verzekerde de indiener is, ontbreekt als vanzelf de variabele AGB-indiener.

In de bestanden zijn de volgende categorieën te onderscheiden:

- 11: Tandartsspecialisten mondziekten en kaakchirurgie
- 12: Tandartsen, algemeen practicus
- 13: Tandartsspecialisten dentomaxillaire orthopaedie
- 38: Tandheelkundige centra
- 39: Regionale instellingen voor jeugdtandverzorging
- 77: Niet in de lijst, onbekend
- 87: Mondhygiënist
- 93: Tandtechnici/tandprothetici

Uit de data is duidelijk af te lezen dat de grootste tandartsen/centra zorgen voor 6% van alle records (debetrecords én creditrecords) in 2012. Zo zijn er tandartsen met 222.578 records, 219.010, 184.490, enzovoorts.

Wanneer we die top 10 weglaten, zien we dat er nog steeds grote verschillen in de frequentie records bestaan. Zie hiervoor figuur 1. Er zijn dus veel behandelaars met meer dan 4000 records in 2012.



Figuur 1. Aantal records (debet en credit) per tandarts, de top 10 uitgesloten, 2012.

De verdeling van het aantal records over deze jaren wordt in tabel 2.1 weergegeven.

Jaartal	Aantal records	Aantal AGB-behandelaars	AGB-indieners
2010	69.297.896	15.190	8.439
2011	68.449.999	15.550	8.325
2012	51.934.447	15.486	8.289
Totaal	189.682.342		

Tabel 1. Aantal records (debet en credit) betreffende vergoedingen voor mondzorg in de jaren 2010 – 2012

De prestatiecode kan worden aangemerkt als een *key*-variabele in deze analyse omdat het rechtstreeks verwijst naar de verrichte prestatie. In een klein percentage van de records ontbrak deze en dit is nader onderzocht (het veld 'prestatiecode' is helemaal leeg). Zie hiervoor tabel 2.

Jaartal	Percentage ontbrekende prestatiecode	Gemiddelde vergoeding bij de ontbrekende prestatiecode in euro	Gemiddelde vergoeding bij bekende prestatiecode in euro	Totaal bedrag bij ontbrekende prestatiecode in miljoenen euro
2010	0,18	77,60	31,37	9,3
2011	0,22	83,28	31,73	12,7
2012	0,02	95,06	41,94	1,4

Tabel 2. Overzicht van de ontbrekende prestatiecode, mondzorg, 2010 – 2012

Het betekent dat in 2010 er 9,3 miljoen is uitbetaald terwijl onbekend is waarvoor; in 2011 is dit 12,7 miljoen en in 2012 gaat het om 1,4 miljoen euro.

Resumerend:

Aandacht is nodig voor ontbrekende waarden op *key*-variabelen zoals prestatiecode, AGB-behandelaar en BSN

6.4 Farmacie

De volgende variabelen zijn beschikbaar voor de jaren 2010, 2011 en 2012.

45. indicator_AV
46. hoeveelheid_afgeleverd_middel
47. vergoedingsbedrag
48. WMG_opslag_gedeclareerd
49. WMG_opslag_NZa_normaal_maximum
50. WMG_opslag_NZa_verhoogd_maximum
51. GVS_bijdrage
52. voorgeschreven_dagdosering
53. leeftijd_op_1_juli
54. toegekend_bedrag
55. prestatiedatum
56. prestatiecodelijst
57. artikelnummer
58. WMG_tariefcode
59. DWH_burger_id
60. BSN_encrypt
61. geboortejaar
62. geboren_voor_1_juli
63. geslacht
64. postcode_cijfers
65. UZOVI
66. Voorschrijver_AGB_code
67. Voorschrijver_naam
68. Voorschrijver_postcode_cijfers
69. Voorschrijver_postcode_letters
70. Voorschrijver_soort
71. Voorschrijver_soort_instelling
72. Voorschrijver_soort_AGB_code
73. Voorschrijver_begindatum
74. Voorschrijver_einddatum
75. Voorschrijver_specialisme_code
76. Voorschrijver_specialisme_omschrijving
77. Afleveraar_AGB_code
78. Afleveraar_naam
79. Afleveraar_postcode_cijfers
80. Afleveraar_postcode_letters
81. Afleveraar_soort
82. Afleveraar_soort_instelling
83. Afleveraar_soort_AGB_code
84. Afleveraar_begindatum
85. Afleveraar_einddatum
86. Praktijk_AGB_code
87. Praktijk_naam
88. Praktijk_postcode_cijfers
89. Praktijk_postcode_letters
90. Praktijk_soort
91. Praktijk_soort_instelling
92. Praktijk_soort_AGB_code
93. Praktijk_begindatum
94. Praktijk_einddatum

Z-INDEX

Daarnaast is er een bestand waarmee meer informatie kan worden verkregen.

1. artikelcode
2. jaar
3. maand
4. handelsproduktcode
5. generiekeproductcode
6. ATC_code
7. productgroep
8. eenheid
9. naam_volledig
10. merkstamnaam
11. etiketnaam
12. ingegeven_sterke_stofnamen
13. memocode
14. GVS_clustercode
15. GVS_vergoedingslimiet
16. aantal_DDD_per_basiseenheid_product
17. aantal_DDD_per_verpakking
18. inkoopprijs
19. verkoopprijs_exclusief_BTW
20. referentieprijs
21. vergoedingsprijs
22. verpakkingshoeveelheid
23. inkoopkanaal
24. code_vervaldatum
25. WTG_code
26. PRK_code
27. farmaceutische_vorm_code
28. RZV_voorwaarden_bijlage2
29. RZV_verstrekking
30. registratienummer_1
31. registratienummer_2
32. registratienummer_3
33. toedieningsweg_code
34. regeling_farmaceutische_hulp
35. reden_regeling_farmaceutische_hulp
36. land_memocode
37. naam_ondernemer
38. memocode_ondernemer_2_posities_numeriek
39. memocode_ondernemer_3_posities_alfanumeriek
40. nummercode_ondernemer
41. BTW_code
42. RVG_nummer_1
43. RVG_nummer_2_teken
44. RVG_nummer_3
45. aantal_deelverpakkingen
46. hoeveelheid_per_deelverpakking
47. deelverpakking_omschrijvingscode
48. aantal_hoofdverpakkingen
49. hoofdverpakkingen_omschrijvingscode
50. leverancier_code
51. U_N_code
52. maximum_prijs_VWS
53. kortingspercentage
54. U_N_datum
55. ZN_code

Resumerend:

Aandacht is nodig voor grote aantal ontbrekende waarden op *key*-variabelen.

6.5 GGZ

De volgende variabelen zijn beschikbaar voor de jaren 2009, 2010 en 2011.

1. FullYear
2. begindatum
3. einddatum
4. decla_jaar
5. decla_kwartaal
6. UZOVI
7. UZOVI_hist
8. BSN_encrypt
9. declaratiecode
10. declagroep
11. verblijf
12. duurverblijf
13. decla_omschrijving
14. prestatiecode
15. prest_omschrijving
16. beginjaar
17. toegekendbedrag
18. aantal_schades_gecorrigeerd
19. oorspronkelijkbedrag
20. overdekkingbedrag
21. toegbedrag_excl_verrekenperc
22. tarief_TOG
23. instttype
24. behandelaarcode
25. indienercode
26. srtvzk
27. naam_AGB
28. financiering

Resumerend:

De data bestanden worden gedomineerd door een aantal zeer grote instellingen/behandelaars. Daardoor is het lastig om “binnen” die grote instellingen te zoeken. De inbreng van meer domeinexperts hierbij zal het aantal anomalieën sterk kunnen vergroten.

Appendix 7 Resultaten op basis van Cortana, huisartsen 2010

1. Inleiding

Zoals in Appendix 1 is beschreven, is Cortana een hulpmiddel voor het duiden van anomalieën en het – op basis daarvan – vinden van *look-alikes*. We geven in deze appendix een aantal voorbeelden van huisartsen zodat de lezer ziet hoe via een stepping stone aanpak een klein aantal leads leidt tot gericht verder onderzoek. De grootste uitdaging bij Cortana is de duiding/omschrijving van de gevonden subgroep. Dit probleem komen we ook tegen bij andere statistische technieken zoals clusteranalyses en principale componentenanalyse.

Voor de analyses die in deze appendix worden beschreven, geldt dat ze alleen uitgevoerd zijn op 2010 data van huisartsen.

LOF scores uitleg:

Age-money: Om LOF-scores te berekenen is gekeken naar de leeftijd van patiënten en de bedragen die verdiend worden aan patiënten. Wanneer er buitenproportioneel veel geld verdiend wordt aan een bepaalde leeftijdsgroep kan dat verdacht zijn. Om verdachte gevallen te vinden zijn de patiënten van een huisarts verdeeld in 6 leeftijdscategorieën (0-17, 18-34, 35-47, 48-60, 61-79, 80 en ouder). Vervolgens is berekend hoeveel procent van de patiënten van een huisarts in elke categorie valt. Het gedeclareerde bedrag van een huisarts is in dezelfde 6 groepen opgesplitst en vervolgens is berekend hoeveel procent van het totaal gedeclareerde bedrag in elke groep valt. Door nu de percentages met leeftijdsgroepen te delen door de percentages gedeclareerd geld kunnen opvallende gevallen gedetecteerd worden.

Long-short ratio: Om LOF-scores te berekenen is gekeken naar de verhoudingen tussen verschillende prestaties, om zo duur declaratiegedrag te vinden. De volgende drie verhoudingen zijn gebruikt om scores te berekenen:

- tussen lange en korte consults ($\text{prest_lang_consult_abs} / \text{prest_kort_consult_abs}$)
- tussen lange en korte bezoeken ($\text{prest_lang_bezoek_abs} / \text{prest_kort_bezoek_abs}$)
- tussen bezoeken en consults ($(\text{prest_lang_bezoek_abs} + \text{prest_kort_bezoek_abs}) / (\text{prest_lang_consult_abs} + \text{prest_kort_bezoek_abs})$)

Hoge verhoudingen duiden op hoge kosten, die het gevolg kunnen zijn van fraude.

2. Resultaten ‘Huisarts’ groep

2.1 Deelanalyse: vijf verdachte huisartsen aangeleverd door NZa

Onderzoeksvraag: hoe zijn de 5 voorbeeldhuisartsen te karakteriseren in termen van hun aggregaat beschrijvingen, en zijn er andere huisartsen aan te wijzen met vergelijkbare kenmerken?

Met behulp van Cortana is de subgroup gevonden, waarin alle 5 huisartsen bevat zijn. De geselecteerde groep van huisartsen voldoet aan de volgende eisen:

$\text{prest_incorrect_abs} \geq 20$
 $\text{declaratie_avg} \geq 10.744685$
 $\text{declaratie_48_60_rel} \geq 0.181643$
 $\text{lang_kort_ratio} \leq 0.215206$
 $\text{prest_inschrijf_abs} \leq 32505$

In totaal gaat het hier om **266 huisartsen** uit de lijst van 12682 huisartsen in de deelverzameling ‘Huisarts’, wat neer komt op een selectie van 2,1%. De exacte samenstelling van deze groep wordt

hier vanwege privacy-overwegingen niet gegeven. Deze subgroup kan een goede start zijn voor het napluizen van deze artsen.

2.2 Deelanalyse: 407 gevallen resulterende van 'test 21a'

Het gaat hier om resultaten van test 21a. De gevonden lijst bestaat in dit geval uit 407 leads.

Een groep van **3282 huisartsen** heeft een verhoogde kans om te scoren op test 21a. Van de 407 gevallen valt 81,6% in deze groep, terwijl de overige huisartsen (een groep van 9400 huisartsen) slechts 18,4% van de verdachte gevallen bevat. Anders omschreven, gemiddeld scoort 3,2% van de huisartsen op test 21a, terwijl dit in deze selectie meer dan verdrievoudigd is naar 10,1%.

`prest_incorrect_rel >= 0.000111`

`prest_haphds_abs = 0.0`

`prestcode_valid_rel >= 0.99864`

2.3 Deelanalyse: 3 gevallen resulterende van 'test 21b'

Deze analyse is in opzet gelijk aan die hierboven beschreven. Het gaat in dit geval echter om resultaten van test 21b. Slechts drie gevallen blijken te scoren op test 21b, wat het vinden van subgroups bemoeilijkt.

De volgende subgroup van **50 huisartsen** (0,39%) bevat deze drie huisartsen:

`prest_lang_bezoek_abs >= 218.0`

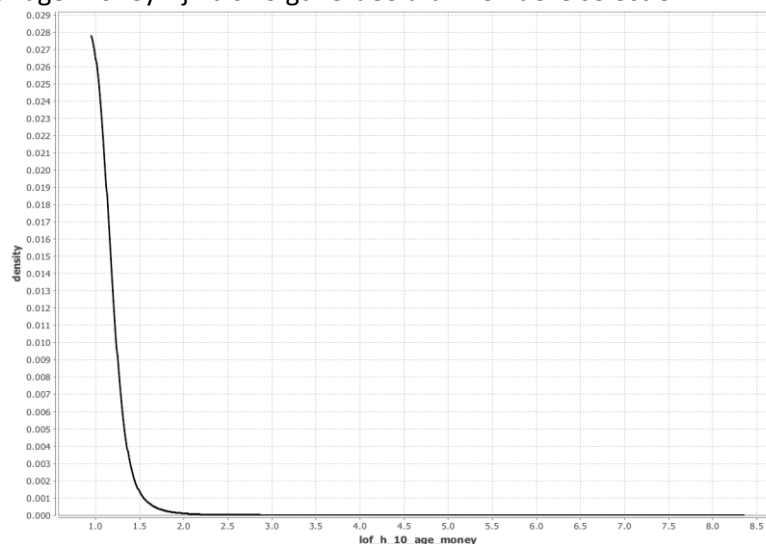
`age_18_34_rel >= 0.172472`

`patients_visiting_rel >= 0.812335`

2.4 Deelanalyse: outlier score LOF age-money

Voor deze analyse is er een pre-selectie gemaakt, waarbij alle huisartsen die 1000 bezoekende patiënten of minder per jaar hebben, weggelaten zijn uit de analyse. De resultaten zijn dus geproduceerd op de 6451 huisartsen met minstens 1001 bezoekende patiënten op jaarbasis.

De LOF scores voor age-money zijn als volgt verdeeld binnen deze selectie:

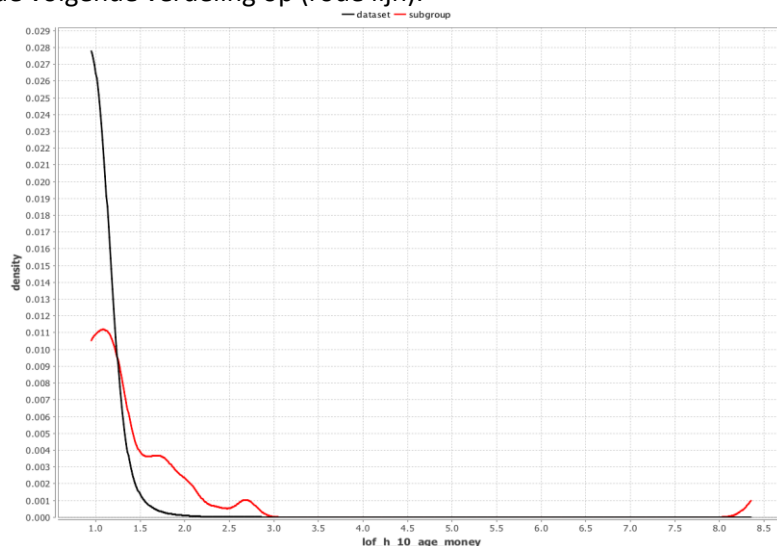


Figuur 1. LOF-scores voor de variabele age-money, huisartsen, 2010

De gemiddelde score is 1,1. Er is een subgroup van **50 huisartsen** die relatief hoge LOF scores hebben voor age-money, wat wijst op een afwijkende verdeling in termen van leeftijd en bedragen. Deze groep is als volgt te omschrijven:

patients_visiting_abs $\leq$ 1353
prest_mi_verrichtingen_rel $\geq$ 0.035407
prest_lang_consult_abs $\geq$ 262

Voor de specifieke groep huisartsen geldt een gemiddelde outlier score van 1,57, dus zo'n 50% hoger. Dit levert de volgende verdeling op (rode lijn).



Figuur 2. LOF-scores voor de variabele age-money en LOF-score voor de subgroup, huisartsen, 2010

2.5 Deelanalyse: opvallend hoge gemiddelde declaratiekosten

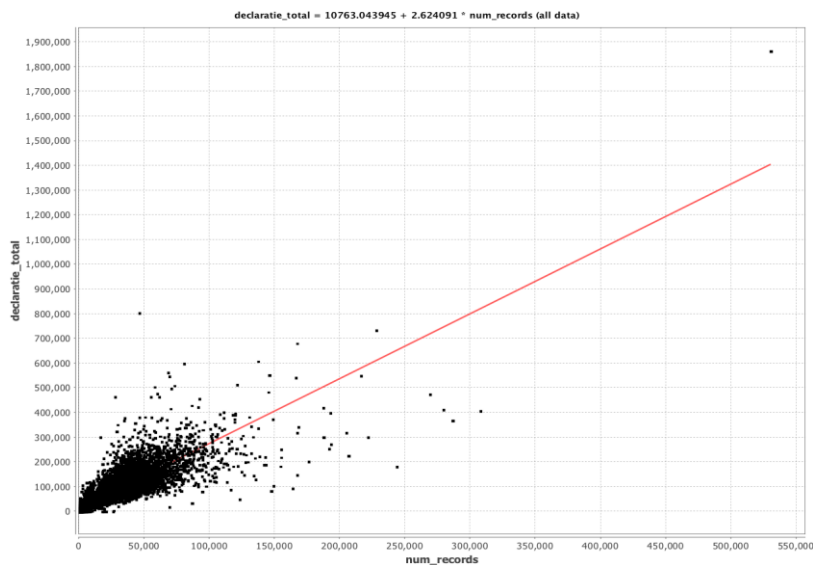
Met behulp van Cortana is de volgende subgroup van **52 huisartsen** gevonden. De geselecteerde groep van huisartsen voldoet aan de volgende eisen:

lang_kort_ratio = 0.0
declaratie_61_79_abs $\geq$ 234
prestcode_valid_abs $\leq$ 8

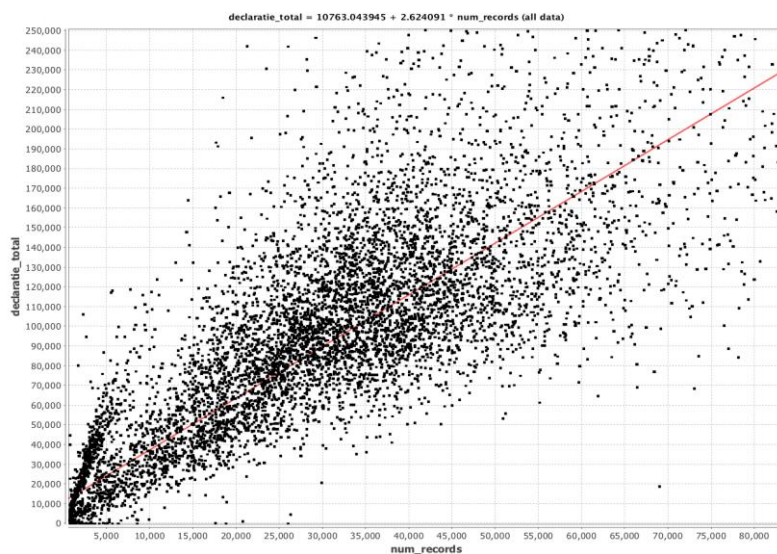
In totaal gaat het hier om 52 huisartsen uit de lijst van 12682 huisartsen in de deelverzameling 'Huisarts', wat neer komt op een selectie van 0,41%. Deze huisartsen hebben een gemiddeld declaratiebedrag van 138,37 euro, vergeleken met een overall gemiddelde van 14,02 euro. De AGB-codes van de huisartsen in deze groep zijn vanwege privacy-overwegingen niet opgenomen in dit rapport.

2.6 Deelanalyse: afwijkende relatie tussen aantal records en declaratietotaal

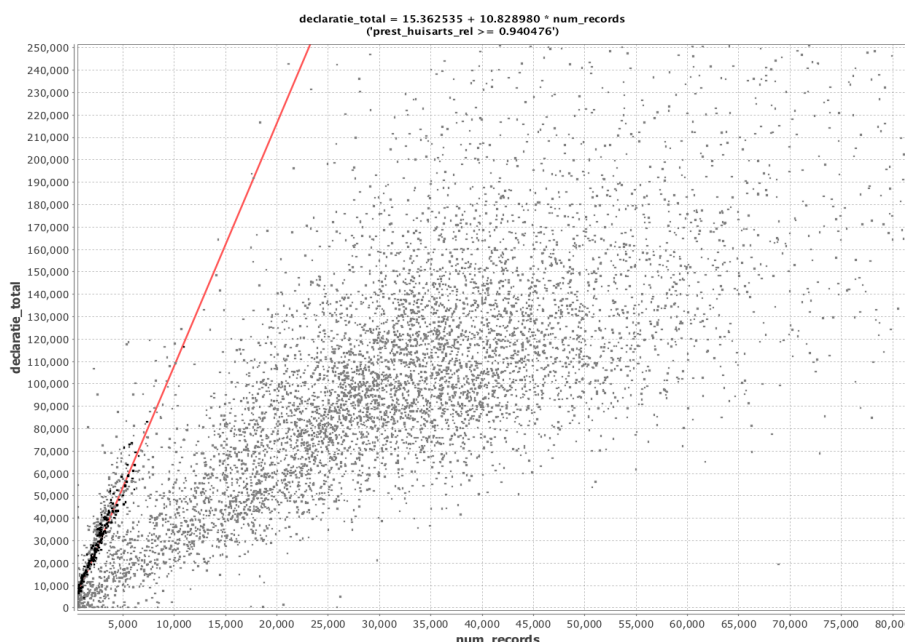
Een eenvoudig lineair model die dit beschrijft is: $\text{declaratie_totaal} = 10763 + 2,62 * \text{aantal records}$ (dus 2,62 euro per record).



Figuur 1. Globale verdeling van aantal records versus declaratietotaal, 2010



Figuur 2. Detail van deze verdeling, 2010



Figuur 3. Subgroup van huisartsen, waarvoor een hoger gemiddeld bedrag per record geldt, 2010

Voor deze onderzoeksvraag vindt Cortana een subgroup van **1070 huisartsen**, waarvan de verhouding tussen aantal records en declaratietotaal duidelijk afwijkt van de norm. Deze huisartsen declareren gemiddeld 10,82 euro per record, vergeleken met de 2,62 euro per record overall. Deze huisartsen zijn gekenmerkt door de volgende simpele conditie:

$$\text{prest_huisarts_rel} \geq 0.940476$$

2.7 Deelanalyse: opvallend hoge gemiddelde declaratiekosten in samenwerkingsverbanden

Met behulp van Cortana is een subgroup van **5 samenwerkingen** gevonden. De geselecteerde groep van samenwerkingen voldoet aan de volgende eisen:

$$\text{declaratie_0_17_abs} \geq 92.13$$

$$\text{patients_total} \leq 455$$

In totaal gaat het hier om 5 samenwerkingen uit de lijst van 237 samenwerkingen in de deelverzameling 'Samenwerking', wat neer komt op een selectie van 2,1%. Deze samenwerkingen hebben een gemiddeld declaratiebedrag van 277,45 euro, vergeleken met een overall gemiddelde van 52,09 euro. De AGB-codes worden uit privacy-overwegingen hier niet gegeven.

2.8 Deelanalyse: outlier score LOF age-money voor gezondheidscentra

Voor deze analyse is er een pre-selectie gemaakt, waarbij alle gezondheidscentra die 1000 bezoekende patiënten of minder per jaar hebben, weggelaten zijn uit de analyse. De resultaten zijn dus geproduceerd op de 46 centra met minstens 1001 bezoekende patiënten op jaarbasis.

Voor de outlier score *age-money* is de volgende subgroup **van 5 gezondheidscentra** gevonden die in de bijbehorende dimensies (leeftijd en bedrag dus) opvallend afwijken van de overige centra. Deze subgroup is te kenmerken met de volgende condities:

$$\text{prest_lang_consult_abs} \leq 534$$

$$\text{prest_telefonisch_rel} \geq 0.251834$$

De gemiddelde LOF-score voor deze subgroup is 2,27 vergeleken met 1,344 voor alle gezondheidscentra met meer dan 1000 bezoekers.

2.9 Deelanalyse: opvallend hoge gemiddelde declaratiekosten

Met behulp van Cortana is de volgende subgroup **van 10 gezondheidscentra** gevonden. De geselecteerde groep van gezondheidscentra voldoet aan de volgende eisen:

$\text{prest_verbruiksmaterialen_abs} \leq 2$

$\text{prestcode_missing_abs} \leq 1$

$\text{num_records} \leq 11200$

In totaal gaat het hier om 10 gezondheidscentra uit de lijst van 237 huisartsen in de deelverzameling 'Gezondheidscentrum', wat neer komt op een selectie van 4,2%. Deze huisartsen hebben een gemiddeld declaratiebedrag van 17,13 euro, vergeleken met een overall gemiddelde van 10,85 euro. De AGB-codes betreffende deze gezondheidscentra worden hier niet gegeven.