Rapport

Datamining bij fraudebestrijding door zorgverzekeraars

Datum : 15 februari 2012

Dit rapport heeft 34 pagina's (inclusief 2 bijlagen)

Inhoudsopgave

0. Samenvatting

In opdracht van het Ministerie van Volksgezondheid Welzijn en Sport heeft Grant Thornton Forensic & Investigation Services B.V een onderzoek uitgevoerd naar de verschillende aspecten van datamining bij de fraudebestrijding door zorgverzekeraars. In het onderzoek komen, naast een meer algemene inventarisatie, de commerciële pakketten die worden aangeboden en/of in gebruik zijn bij zorgverzekeraars aan bod. De gegevens zijn verzameld door middel van documentstudie, interviews met medewerkers van zorgverzekeraars, softwareleveranciers en instanties met een belang bij fraudebestrijding in de zorg.

Datamining is het door middel van geavanceerde methoden zoeken naar voorheen onbekende verbanden en patronen in grote verzamelingen gegevens en wordt onder meer toegepast in de (semi)geautomatiseerde detectie van fraude bij financiële instellingen en (zorg)verzekeraars. Alle zorgverzekeraars die aan het onderzoek hebben meegewerkt, maken gebruik van (semi)geautomatiseerde detectie van fraude in zorgdeclaraties, zij het niet allemaal op dezelfde wijze en in dezelfde mate.

Diverse softwarepakketten voor datamining zijn bruikbaar ten behoeve van fraudedetectie (bij zorgverzekeraars). In dit onderzoek zijn vijf softwarepakketten van vier leveranciers vergeleken. Twee van deze softwarepakketten zijn daadwerkelijk in gebruik bij zorgverzekeraars in Nederland. Op hoofdlijnen bieden al de vergeleken softwarepakketten dezelfde functionaliteiten en de eventuele keuze voor een specifiek pakket ten behoeve van fraudedetectie zal vooral ingegeven worden door omstandigheden en voorkeuren die per organisatie zullen verschillen.

Hoewel de bevindingen van het onderzoek laten zien dat datamining kan bijdragen aan een efficiëntere bestrijding van fraude binnen de zorg, blijken de omstandigheden waaronder dit plaatsvindt van (even) groot belang. Het is van primair belang dat de gegevens zo gedetailleerd mogelijk zijn en dat daar in de standaarden voor gegevensuitwisseling rekening mee wordt gehouden. Ook is het noodzakelijk dat de detectie en aanpak van fraude in de gehele keten goed worden afgestemd om daadwerkelijk effect te sorteren. Daarbij moet naast de detectie gedacht worden aan fraudebestendige regelgeving, robuuste procedures voor onder meer erkenningen en indicatiestellingen, en voldoende capaciteit - zowel bij zorgverzekeraars als bij de overheid - voor onderzoek en aanpak van fraudeurs.

Een belangrijke bevinding is dat datamining voor fraudedetectie in de zorg het meeste effect zou kunnen sorteren wanneer dit centraal over een zo groot mogelijke dataset wordt uitgevoerd. Dit is dan ook de eerste van de in totaal zes aanbevelingen die op basis van dit onderzoek worden gemaakt.

1. Aanleiding onderzoek, opdracht en methode

Op 26 september 2011 heeft het Ministerie van Volksgezondheid Welzijn en Sport (hierna: het ministerie) aan Grant Thornton Forensic & Investigation Services B.V. (hierna: Grant Thornton) de opdracht verstrekt om een onderzoek in te stellen naar de verschillende aspecten van datamining bij de fraudebestrijding door zorgverzekeraars. De achtergrond van het onderzoek is door het ministerie in de offerteaanvraag als volgt omschreven:

'Zorgverzekeraars / zorgkantoren gaan in hun formele en materiële controles na of de geleverde zorg binnen de verzekerde prestatie valt, of de zorg geleverd is en of het tarief dat in rekening is gebracht klopt. Dit geldt zowel voor de Zvw als voor de AWBZ zorg.

Controle is een essentiële stap in de beheersing van fraude: zonder een goede controle zal er geen fraude kunnen worden gedetecteerd. Er zijn echter meer verbanden tussen controle en fraudebeheersing: voor beide zijn eenduidige regelgeving, risicoanalyses en een goede vastlegging van onder meer declaratiegegevens van groot belang. Controle en fraudebeheersing dienen dan ook in samenhang bekeken te worden.

Controle richt zich in de praktijk vooral op de vraag of de declaraties juist zijn en minder op de vraag of de geleverde én gedeclareerde zorg onder de verzekerde prestatie valt en dus rechtmatig is verstrekt. Door het verbeteren van controleprocessen zal meer fraude worden opgespoord. Hierdoor zal de schadelast verminderen. In een tijd waarin we zien dat de zorgkosten maar blijven stijgen is het realiseren van een betere beheersing daarvan beleidsprioriteit nummer één. Alom bestaat het idee dat een betere controle daaraan een belangrijke bijdrage kan leveren. Dit sluit ook goed aan bij de tendens binnen de rijksoverheid om het toezicht op de controle van collectieve middelen te intensiveren.

Om aan het bestrijden van fraude meer aandacht te besteden rijst in het bijzonder de vraag hoe softwaresystemen behulpzaam kunnen zijn om vreemde patronen die op fraude kunnen duiden, te signaleren. Zorgverzekeraars gebruiken instrumenten als datamining nog maar mondjesmaat.'

Op basis van de bovenstaande probleemomschrijving zijn door het ministerie de navolgende (vier clusters van) onderzoeksvragen geformuleerd:

1. *Welke softwaresystemen voor bijvoorbeeld datamining zijn er op de markt? Wat zijn de belangrijkste softwaresystemen in de sfeer van het bestrijden van zorgfraude?*

2. *Hoe simpel of ingewikkeld is het om met deze systemen aan de slag te gaan? Detecteren de systemen ook zwakheden in onderliggende wetgeving? Geven de systemen ook een verklaring voor afwijkende patronen? Zijn deze systemen gemakkelijk toepasbaar in de zorg?*

3. *Wat zijn de effecten van deze systemen? Welke voor- en nadelen hebben ze? Welke vaardigheden hebben de gebruikers nodig om ermee te kunnen werken? Wat zijn de kosten van deze systemen (aanschaf en licentie?) Hoe makkelijk / moeilijk zijn ze te implementeren en wat kost dat? Hoe makkelijk / moeilijk zijn ze uit te breiden? Hoe gebruiksvriendelijk zijn ze? Welke onderdelen van de zorg lenen zich het beste voor instrumenten als datamining?*

4. *Hoe komt het dat deze systemen nog niet op grote schaal worden gebruikt door zorgverzekeraars? Wat zijn hier kansen en belemmeringen? Met welke argumenten zijn verzekeraars over te halen op grote schaal gebruik te maken van deze systemen?*

Teneinde deze vragen te beantwoorden, is op 27 september 2011 een onderzoek aangevangen dat is afgesloten op 30 januari 2012. In het onderzoek is op drie wijzen informatie verzameld. In de eerste plaats is voor het onderzoek documentatie bestudeerd die ons is aangeleverd door het ministerie en geïnterviewden, of uit open bronnen is verkregen. De relevante gebruikte documentatie, met uitzondering van de productbrochures van de softwarepakketten, staat beschreven in bijlage 2.

Daarnaast zijn 22 interviews gehouden met medewerkers van zorgverzekeraars (fraudecoördinatoren, bestuurders), brancheorganisaties, overheid en softwareleveranciers. Een lijst van de organisaties waar de geïnterviewden uit afkomstig zijn, is bijgevoegd als bijlage 1. Met het oogmerk de geïnterviewden de gelegenheid te bieden vrijuit te spreken zonder dat bepaalde uitspraken aan een persoon of organisatie kunnen worden gekoppeld, is vooraf met het ministerie overeengekomen dat de resultaten van de interviews niet tot personen en/of organisaties herleidbaar in het rapport worden verwerkt. Het ministerie heeft tevens geen afschrift van, en/of inzage in de interviewverslagen gekregen.

In de derde plaats heeft met de fraudecoördinatoren van de zorgverzekeraars op 10 november 2011 een expertmeeting plaatsgevonden. Tijdens deze meeting is aan de hand van een groot aantal stellingen variërend van de detectie tot en met de bestrijding van fraude door zorgverzekeraars, met de fraudecoördinatoren gediscussieerd. De stellingen zijn geformuleerd op basis van de interviewresultaten en de bestudeerde documentatie. Het doel van de expertmeeting is primair een verificatie van de resultaten van interviews, met andere woorden, in hoeverre de uitspraken van een enkele geïnterviewde gedeeld worden door de overige fraudecoördinatoren. Daarnaast was het doel een verdere diepgang en detaillering van relevante aspecten van de (on)mogelijkheden van geautomatiseerde fraudedetectie bij zorgverzekeraars te verkrijgen.

Er heeft geen technisch onderzoek naar de (werking van de) in dit rapport beschreven softwarepakketten plaatsgevonden.

Na een bespreking van de conceptbevindingen met de opdrachtgever op 6 december 2011 heeft nadere informatieverzameling en analyse plaatsgevonden. De bevindingen zijn in het volgende hoofdstuk, per cluster onderzoeksvragen, weergegeven. Op 15 februari 2012 is dit rapport uitgebracht aan het ministerie.

2. Bevindingen

2.1 Algemeen

In de vraagstelling van het onderzoek wordt de term 'datamining' gebruikt. Datamining wordt over het algemeen gedefinieerd als het *door middel van geavanceerde methoden zoeken naar voorheen onbekende verbanden en patronen in grote verzamelingen gegevens* (zie o.a. Seifert 2004). Dit betekent in de praktijk dat een grote hoeveelheid gegevens (semi) automatisch wordt bewerkt (gefilterd, geclusterd, samengevoegd, geclassificeerd) en wordt geanalyseerd, waarna eventueel gevonden verbanden en/of patronen worden beschreven en/of gevisualiseerd ten behoeve van nader onderzoek. Datamining wordt onder meer toegepast binnen de wetenschap voor het analyseren van onderzoekdata en binnen commerciële organisaties voor de analyse en verbetering van de bedrijfsvoering, waaronder de detectie van fraude.

Naast datamining bestaan ook andere methoden van (semi)geautomatiseerde fraudedetectie. In een recent artikel over geautomatiseerde fraudedetectie in de gezondheidszorg in de Verenigde Staten worden vier typen analyse van grote hoeveelheden gegevens onderscheiden die voor '*Electronic Fraud Detection'* in de zorgsector kunnen worden toegepast (Travialle *et al.* 2011). De vier genoemde typen analyse betreffen respectievelijk:

- *Supervised Classification Techniques* waarbij gegevens aan de hand van vooraf ingestelde patronen van 'legitiem' en 'frauduleus' handelen worden geanalyseerd;
- *Unsupervised Classification Techniques* waarbij gegevens worden geanalyseerd op afwijkingen zonder dat vooraf definities aan potentiële afwijkingen worden gegeven;
- *Statistical methods* waarmee gegevens worden geanalyseerd aan de hand van statistische modellen;
- *Rule based* technieken waarbij gegevens aan de hand van door domein experts opgestelde modellen worden geanalyseerd.

De eerste twee clusters van technieken worden over het algemeen gekwalificeerd als datamining (zie o.a. Mikut and Reischl 2011: 434) maar de beschrijvingen laten zien dat (semi)geautomatiseerde fraudedetectie breder is dan alleen datamining. Het onderscheid tussen de vier soorten technieken zal in dit rapport worden aangewend om waar nodig de bevindingen nader te duiden.

2.2 Beantwoording onderzoeksvraag 1

• *Welke softwaresystemen voor bijvoorbeeld datamining zijn er op de markt? Wat zijn de belangrijkste softwaresystemen in de sfeer van het bestrijden van zorgfraude?*

2.2.1 Beschrijving softwarepakketten

Er zijn tientallen commerciële en open source softwarepakketten voor datamining op de markt, variërend van relatief eenvoudig tot complex, gericht op een specifiek type van datamining, een specifieke sector of meer algemeen van aard (zie voor een overzicht Mikut and Reischl 2011).

Door middel van een marktscan en navraag bij de verschillende zorgverzekeraars is geïnventariseerd wat de bekende commerciële[1] softwarepakketten zijn die voor het detecteren en bestrijden van zorgfraude worden aangeboden op de Nederlandse markt. Zowel uit de marktscan als uit de bevraging van de fraudecoördinatoren van de zorgverzekeraars komen dezelfde vier systemen naar voren[2] waarbij in een later stadium van het onderzoek één van de leveranciers informatie verstrekte over een tweede softwarepakket dat zij op de markt brengen. De fraudecoördinatoren noemen daarnaast een casemanagementpakket dat niet direct een opzichzelfstaand fraudedetectiepakket betreft, maar wel een bijdrage aan (semi)geautomatiseerde fraudedetectie kan leveren. De softwarepakketten die hierna besproken worden, betreffen:

Leverancier[3]	**Functionaliteit**	**Productnaam**
FICO/Ordina	Detectie, analyse & casemanagement	FICO Insurance Fraud Manager
FRISS	Detectie, analyse en casemanagement	FRISS
IBM	Analyse	IBM Modeler
	Detectie, analyse en casemanagement	IBM *Fraude and Abuse Management System* (FAMS)
SAS	Detectie, analyse & casemanagement	SAS
Krammer Software	Casemanagement	*Facts!*

[1] Bij een aantal zorgverzekeraars is intern ontwikkelde software in gebruik zowel voor detectie en tegenhouden van onregelmatige declaraties als voor het casemanagement bij de bestrijding van zorgfraude. Deze software is niet in het onderzoek betrokken.

[2] Met de leverancier van een vijfde softwarepakket dat door één van de fraudecoördinatoren werd genoemd, is contact gezocht maar werd geen respons van ontvangen. Dit systeem wordt overigens bij geen van de verzekeraars die aan het onderzoek meewerkten gebruikt.

[3] Grant Thornton Forensic & Investigation Services heeft met geen enkele van de in dit rapport genoemde leveranciers enige zakelijke band of overeenkomst.

FICO

Kern van de systematiek van de FICO Insurance Fraud Manager is dat elke transactie als een frauduleuze transactie wordt beschouwd, tenzij door de kwaliteitsscoring de waarschijnlijkheid dat sprake is van fraude (0-1000 score) afneemt. Deze techniek is ontleend aan de fraudedetectie voor creditcardtransacties waar FICO wereldwijd naar eigen zeggen een marktaandeel heeft van ruim 70%. De standaardwijze van detectie is een wijze van *peer-reviewing*. Elke afwijking van het (verwachte) gedrag van de *peer group* wordt gedetecteerd en krijgt een bepaalde score. Het verwachte gedrag wordt zowel bepaald aan de hand van analyse van de data, als op basis van wat experts op het specifieke gebied aan het scoringsmodel hebben toegevoegd op basis van hun expertise en ervaring met ruim 200 zorgverzekeraars wereldwijd.

FICO Insurance Fraud Manager bestaat uit verschillende modules, te weten, *business rules* management voor de regelgeving van bijzondere gevallen, neurale netwerken voor patroon herkenning en gedragsherkenning, *predictive models* voor de voorspellende modellen en *fuzzy logic* voor de detectie van schijnbaar vergelijkbaar gedrag. FICO biedt voor verschillende soorten zorg verschillende modules aan die elk vooraf ingestelde patronen van 'legitiem' en 'frauduleus' handelen bevatten. FICO levert in Nederland momenteel alleen nog de mondzorgmodule. Andere zorgmodulen (o.a. huisartsen, specialisten, apothekers en ziekenhuizen) moeten nog voor gebruik in Nederland worden geconfigureerd.

Er worden door FICO twee verschillende wijzen van implementatie (configuraties) van het systeem aangeboden. De eerste betreft het providermodel waarbij specifiek wordt gefocust op het gedrag van de zorgaanbieder waarvan de declaratiegegevens *after-the-fact* batchgewijs worden geanalyseerd. De andere configuratie is het claimmodel waarbij binnen de ingediende claims *real time* detectie van eventuele fraude plaatsvindt. Voor de providermodellen wordt de applicatie niet lokaal bij de verzekeraar geïnstalleerd, maar worden de gegevens (maandelijks) naar FICO *gepushed* door middel van een versleutelde verbinding. FICO importeert de relevante datasets (claimdetails) van de betrokken verzekeraar in haar systeem in de Verenigde Staten en maakt bij de scoring van de gegevens gebruik van ervaringen bij andere zorgverzekeraars. Er wordt momenteel gewerkt aan een lokale server omgeving in Nederland.

Om een betrouwbare score te hebben dient een voldoende omvangrijke dataset te worden aangeleverd. Daarbij moet het, volgens de vertegenwoordiger van FICO, gaan om bijvoorbeeld minimaal 3,5 miljoen transacties op jaarbasis voor mondzorg of bijvoorbeeld minimaal 7 miljoen transacties voor medicijnverstrekkingen. Bij een kleinere dataset kan niet meer met het standaard *Insurance Fraud Manager* (IFM) product gewerkt worden. De oplossing word dan gecreëerd door middel van een combinatie van Business Rules en de standaard providermodellen (*Custom Model Approach*). Daarnaast is het van belang dat tenminste van 6 maanden vergelijkbare declaratiegegevens aanwezig dient te zijn, wil er een zeker mate van betrouwbaarheid van de scores kunnen worden gegarandeerd.

Nadat de scoring van de transactiedata heeft plaatsgevonden, krijgt de klant via een webapplicatie toegang tot de resultaten. Vanuit de resultaten kan worden gelinked naar de onderliggende transacties en met behulp van een casemanagement module verder worden onderzocht.

Tijdens het onderzoek bleek dat het FICO pakket (module mondzorg in de vorm van het providermodel) in Nederland bij één zorgverzekeraar is geïmplementeerd en dat een andere zorgverzekeraar een pilot met FICO overweegt.

FRISS

Het fraudedetectie platform van FRISS detecteert afwijkend gedrag met gebruik van een combinatie van vier technieken: *rule based* (detectie op basis van kennisregels), profielen (detectie op basis van "normaal gedrag" in profielen), voorspellende modellen (detectie op basis van historisch gedrag) en link-analyse (detectie door netwerk analyse op basis van externe data). Daarmee kunnen zowel individuele declaraties als declaratieprofielen van zorgverleners worden getoetst. De detectietool maakt voor de link-analyse gebruik van vergelijkingsdata uit circa 60 beschikbare externe bronnen zoals handelsregister en kredietwaardigheidsgegevens. De profielen en regels zijn via een beheerstool voor meer ervaren gebruikers zelf aan te passen aan bijvoorbeeld veranderingen in de producten of geconstateerde risicowijzigingen. De detectie levert een score van de declaratie, verzekerde of zorgverlener op in een zogenoemd stoplichtenmodel (rood-oranje-groen). Naast de risicoscore zijn eveneens de argumentatie (onderbouwing van de risicoscore) en de onderliggende gegevens beschikbaar.

FRISS kan net als FICO de detectie-applicatie als '*software as a service*' leveren, het platform draait dan bij FRISS en de klanten leveren batchgewijs, danwel *real time* de gegevens aan. Naast de detectiemodules, beschikt het FRISS platform ook over een casemanagementmodule. Binnen dit deel van het platform kan de verzekeraar zijn fraude-incident of andere incidenten afhandelen. Op basis van de resultaten kunnen de indicatoren automatisch worden bijgesteld en kan deze module worden gebruikt om alle formele en materiele controles[4] te plannen en te monitoren.

Volgens opgave van de leverancier zou ten tijde van het onderzoek bij één zorgverzekeraar een *proof-of-concept* met het FRISS-platform gaande zijn en het platform bij een andere zorgverzekeraar al daadwerkelijk in gebruik zijn voor detectie van fraude in de declaratiestroom. Deze informatie werd door de fraudecoördinatoren van de beide zorgverzekeraars, ook na navraag in hun eigen organisatie, evenwel niet bevestigd[5]. Het is wel bekend dat het FRISS-platform en met name de casemanagement module bij een aantal schadeverzekeraars en financiële instellingen in gebruik is.

[4] Zie paragraaf 2.5.1 waar wordt ingegaan op de formele en materiële controles.
[5] Wel is bij een van beide verzekeraars FRISS in gebruik voor de controles bij het beheer van (nieuwe) klanten.

IBM

IBM Modeler (voorheen bekend onder de namen SPSS Clementine en SPSS Modeler) is een generiek datamining/analyse softwarepakket. Het pakket is niet specifiek toegesneden op fraudedetectie of op de zorgsector, maar kan in alle soorten dataverzamelingen de gegevens in relatie tot elkaar bekijken en afwijkingen laten zien. Bij gebruik bij een zorgverzekeraar kan vanuit een eerste analyse vervolgens worden ingezoomd op de afwijkingen, bijvoorbeeld individuele zorgaanbieders, en door middel van een zogenaamde *heatmap* kan/kunnen de oorza(a)k(en) van de afwijking gevisualiseerd worden. Het is mogelijk tot op regelniveau in de onderliggende data (bijvoorbeeld zorgdeclaraties) in te zoomen.

De analyses die op de gegevens worden losgelaten, zijn door de gebruiker te configureren op basis van een gekozen methode van aanpak. Dit vereist van de analist die hiermee gaat werken zowel grondige statistische kennis, als zorginhoudelijke kennis. Het pakket draait binnen de eigen organisatie van de gebruiker. Tijdens het onderzoek bleek dat IBM Modeler momenteel bij één zorgverzekeraar in gebruik is.

Daarnaast biedt IBM het *Fraude and Abuse Management System* (FAMS) aan waar specifiek op de gezondheidszorg toegesneden modellen voor zijn ontwikkeld. FAMS kan zowel binnen de eigen organisatie alsook als dienst worden geleverd. Volgens opgaaf van IBM is FAMS bij méér dan 40 organisaties wereldwijd en met name in Amerika in gebruik. Het stelt de gebruikers in staat om zorgverleners onderling te vergelijken op meer dan 8000 voorgedefinieerde zorg specifieke indicatoren en fraudeschema's. Op basis van statistische formules worden afwijkingen in declaratiegedrag geautomatiseerd per zorgaanbieder in beeld gebracht. Gebruikers kunnen vervolgens beschikken over online rapporten met betrekking tot het fraude-potentieel per beroepsgroep en kunnen via een web-interface inzoomen op afwijkingen per zorgaanbieder en individuele declaratie. Nieuw ontdekte fraudeschema's kunnen in het systeem worden opgenomen. Het systeem kan volgens opgave van IBM worden toegesneden op het Nederlandse zorgstelsel maar is op dit moment nog niet bij een Nederlandse zorgverzekeraar in gebruik.

SAS

SAS bouwt en levert diverse *business analytics* systemen; de SAS-applicaties zijn opgebouwd rond de functies 'analytics', ´datamodelling´ en ´reporting´ waarbij van alle data die uiteindelijk gerapporteerd worden, kan worden nagegaan waar deze vandaan komen en welke bewerkingen de data hebben ondergaan. Eén van de gebieden waar *business analytics* systemen van SAS gebruikt kunnen worden, is in de fraudebestrijding. Voor fraudebestrijding bij verzekeraars biedt SAS het *Fraud Framework for Insurance* aan.

Binnen dit *framework* vindt de analyse van de gegevens plaats aan de hand van drie typen analyse: allereerst aan de hand van *rules* die een weerslag van bekende fraudemethoden zijn. De tweede analyse vindt plaats met behulp van stochastische berekeningen (modellen) waarbij wordt bepaald in hoeverre observaties afwijken

van de populatie (anomalie detectie en voorspellende modellen). Afwijkende observaties kunnen duiden op onregelmatigheden, bijvoorbeeld fraude. In de SAS systematiek kunnen de fraudemethoden, die door middel van de stochastische analyse zijn ontdekt en zijn gevalideerd als fraude, aan de set met *rules* worden toegevoegd, waarbij de parameters variabel zijn in te stellen. De derde analyse, die SAS hanteert bij fraudedetectie, is een *network based* analyse. In deze analyse wordt naar de samenhang van declaraties gekeken, maar ook externe gegevens kunnen worden gebruikt om verbanden en patronen te herkennen.

Bovenop de fraudedetectie biedt SAS een casemanagementmodule aan, waarbij elke declaratie die als potentieel frauduleus wordt bestempeld, te volgen is. Deze casemanagement module heeft een eigen analysefunctie die alle soortgelijke cases kan identificeren, waarna deze kunnen worden gebundeld. Het *Fraud Framework for Insurance* is modulair opgebouwd, een organisatie kan bijvoorbeeld eerst alleen de *rule engine* afnemen en pas later de andere analysemodules. Er zijn verschillende wijzen waarop een SAS oplossing kan worden geïmplementeerd. De minst ingrijpende is een analyse *´after the fact´* waarbij periodiek een dump van de declaratiegegevens in een SAS omgeving wordt geladen, hetzij binnen de eigen organisatie, hetzij in een door SAS aangeboden *hosted solution*. Het is ook mogelijk om met SAS *real time* de declaratiegegevens op potentiële fraude te analyseren. Tijdens het onderzoek bleek dat vier van de zorgverzekeraars SAS als *business analystics* systeem in gebruik hebben, zij het (nog) zonder de fraudemodule.

Facts!

Facts! is geen fraude-detectiesoftware maar een casemanagement systeem waarin vermoedelijke fraude-incidenten uit een fraude-detectietool of uit andere bronnen kunnen worden ingelezen waarna deze met andere gegevens kunnen worden veredeld. *Facts!* kan vervolgens doormeldingen genereren aan (externe) systemen en de veredelde gegevens kunnen vervolgens weer nieuwe input zijn voor de fraude-detectietools. Tijdens het onderzoek bleek dat vrijwel alle zorgverzekeraars met dit casemanagementsysteem werken en dat het sinds kort ook op centraal niveau bij Zorgverzekeraars Nederland (hierna: ZN) gebruikt wordt. Nu nog in een pilot, maar in 2012 zullen alle zorgverzekeraars starten met het doorgegeven van hun incidenten vanuit hun eigen *Facts!* aan ZN. Daar kan vervolgens, indien nodig, coördinatie van de fraudebestrijding plaatsvinden. De gezamenlijke gegevens kunnen gebruikt worden als input voor centrale fraudedetectie. Gelet op de focus van dit onderzoek op systemen voor detectie van fraude zal *Facts!* in dit rapport verder niet besproken worden.

2.2.2 Vergelijking en analyse

De kenmerken van de in de vorige paragraaf beschreven softwarepakketten voor fraudedetectie zijn in de onderstaande tabel samengevat en zullen daarna nader worden besproken.[6]

Pakket	Modulair	Detectietechniek(en)	Case-management	Interactieve visualisatie	Hosted/ In company	Real time/ After-the-fact
FICO	Ja	Regels, neurale netwerken, voorspellende modellen, fuzzy logic. (Analyse type: *Rule based*, *Supervised Classification, Unsupervised Classification*)	Ja	Ja	Beide	Beide
FRISS	Ja	Regels, profielen, voorspellende modellen en link-analyse. (Analyse type : *Rule based*, *Unsupervised Classification*)	Ja	Ja	Beide	Beide
IBM Modeller	Nee	Statistische modellen (Analyse type: *Statistical methods*)	Nee	Ja	In company	After the fact
IBM FAMS	Ja	Regels, voorspellende modellen (Analyse type : *Rule Based*, *Unsupervised Classification*)	Ja	Ja	Beide	Beide
SAS *Fraud Framework*	Ja	*Rule based*, anomaliedetectie, voorspellende modellen, social network analysis. (Analyse type : *Rule Based*, *Supervised Classification* , *Unsupervised Classification*)	Ja	Ja	Beide	Beide

Er zijn tientallen verschillende technieken die gebruikt kunnen worden voor datamining ten behoeve van fraudedetectie (zie Travaille *et al.* 2011:5 en Mikut and Reischl 2011: 434-435), onderstaand worden de verschillende detectietechnieken die door de leveranciers van de bekeken pakketten worden genoemd toegelicht.

Detectietechnieken

Bij *rule based* detectie van fraude worden de declaraties aan de hand van vooraf ingestelde regels geanalyseerd waarbij wordt bekeken of de declaraties aan bepaalde vereisten voldoen. Zo kunnen eenvoudig dubbele declaraties worden gefilterd maar ook declaraties die voldoen aan een patroon van eerder ontdekte fraude, bijvoorbeeld onlogische combinaties van verschillende behandelingen of onmogelijke combinaties van behandeling en patiënt, zoals bijvoorbeeld een wortelkanaalbehandeling bij driejarigen. Het nadeel van het gebruik van vaste regels bij fraudedetectie is dat fraudeurs hun *modus operandi* zullen aanpassen zodat deze niet meer door de vaste regels worden opgemerkt. Het voordeel van *rule based* detectie is dat het een relatief eenvoudig proces is waarmee grote hoeveelheden data op efficiënte wijze kunnen worden *gescreend* op het voldoen aan formele vereisten en eventuele afwijkingen. Als onregelmatig bestempelde declaraties kunnen zonder verdere tussenkomst van een fraudeonderzoeker worden afgewezen.

[6] De kenmerken zijn volgens opgave van de leveranciers. Het is niet mogelijk gebleken alle geclaimde functionaliteiten te verifiëren.

Bij detectie van potentiële fraude door middel van *neurale netwerken* (*anomalie detectie*) wordt de software aan de hand van bekende fraudepatronen 'geleerd' welke (input en output) kenmerken van de zorgaanbieder, het type zorg, de verzekerde, etc. aanwijzingen zijn voor fraude. Het gebruik van *neurale netwerken* valt onder de zogenoemde *Supervised Classification Techniques*. Een voordeel daarvan is dat door veel meer gegevens te analyseren dan bij detectie op basis van regels, veel nauwkeuriger kan worden bepaald of wel of geen sprake is van fraude. Een nadeel is dat nieuwe fraudeschema's pas worden ontdekt als voldoende referentiedata voorhanden is. Bovendien is het bouwen van modellen kostbaar; de geïnterviewde vertegenwoordiger van FICO schatte de kosten voor het maken van de FICO-service voor medicijnverstrekkingen (apotheekmodule) op € 750k tot € 1,5 miljoen.

Bij detectie van potentiële fraude door middel van *voorspellende modellen* wordt gebruik gemaakt van historische gedragsinformatie om afwijkende patronen te herkennen. Inzicht in het gedrag van zowel zorgaanbieders als verzekerden kan afwijkende patronen aan het licht brengen. Een voordeel van deze techniek, die onder de zogenoemde *Unsupervised Classification Techniques* valt, is dat snel kan worden gezien of een bepaalde declaratie afwijkt van historische patronen. Het nadeel van deze techniek is volgens Travaille *et al.* (2011: 8) het hoge aantal *false positives*.[7]

Fuzzy logic is een analysesystematiek waarbij de uitkomst meer waarden kan hebben dan alleen 'waar' of 'niet waar' en wordt bij de detectie van fraude toegepast om mogelijk vergelijkbare gevallen te detecteren die niet exact overeenkomen maar wel bijna. Het nadeel hiervan is dat het aantal *false positives* kan toenemen.

De detectie van potentiële fraude door middel van link-analyse (*social network analysis*) vindt plaats aan de hand van het gebruik van andere (interne en externe) datasets. Het grote voordeel is dat verbanden die niet uit de beschikbare declaratie-informatie kunnen blijken, toch gevonden worden. Een nadeel van deze techniek is dat relevante interne en externe databestanden moeten worden ontsloten en/of aangekocht.

Bij detectie van fraude door middel van *statistische methoden* wordt onderzocht door middel van statistische berekeningen welke verbanden of afwijkingen in de dataset te vinden zijn aan de hand van een hypothese. Bijvoorbeeld een zorgaanbieder die per patiënt gemiddeld meer declareert dan collega's in vergelijkbare situaties. Het nadeel van deze techniek is dat elk gevonden verband en/of afwijking op relevantie onderzocht moet worden met het gevaar op een hoog aantal *false positives*. Ook zal de gebruiker al zelf een goed beeld (hypothese) van potentiële fraude moeten hebben die vervolgens wordt getest. Het voordeel is dat met de

[7] Een *false positive* wil zeggen dat een geval ten onrechte als onregelmatig/onrechtmatig wordt aangemerkt terwijl een legitieme verklaring voor de gevonden afwijking bestaat. Bij een *false negative* daarentegen is een daadwerkelijke onregelmatigheid /onrechtmatigheid als niet afwijkend geclassificeerd. Grofweg kan gesteld worden dat de aantallen *false positives* en *false negatives* in fraude detectiesystemen aan elkaar zijn gerelateerd: hoe strakker de regels en detectiemodellen staan ingesteld, hoe groter de kans op *false positives*, immers de tolerantie voor afwijkingen is kleiner. Omgekeerd geldt hoe minder strak de regels en modellen zijn afgesteld, hoe groter de kans op *false negatives*, immers er worden meer afwijkingen geaccepteerd en dus ook daadwerkelijke fraude. Over het algemeen geldt daarbij dat teveel *false positives* resulteren in een inefficiënt gebruik van onderzoekscapaciteit en teveel *false negatives* resulteren in een laag detectieniveau.

juiste inhoudelijke kennis heel specifiek en diepgaand onderzoek kan worden gedaan op één enkel aspect in de dataset.

Andere kenmerken

De besproken pakketten van FICO, FRISS, IBM en SAS zijn allen *modulair* hetgeen wil zeggen dat de verschillende aangeboden (detectie)functies los van elkaar verkrijgbaar zijn. Daar zit vaak wel een bepaalde volgorde in, maar na afname van een basisconfiguratie kunnen andere modules, naar gelang de behoefte, worden toegevoegd. Ook is het bij al deze pakketten mogelijk te kiezen voor een *hosted* / *in company* oplossing, hetgeen betekent dat het pakket ofwel binnen de eigen organisatie draait, ofwel in een datacentrum van de leverancier. Ook kan gekozen worden voor analyse achteraf (*after-the-fact*) waarvoor de gegevens batchgewijs worden verwerkt of een *real time* oplossing waarbij de analyse op potentiële fraude plaatsvindt in het declaratieproces. Tenslotte beschikken alle pakketten over (interactieve) visualisatie van de output.

Samenvattend kunnen de hier beschreven softwarepakketten voor datamining / fraudedetectie worden gekenmerkt als de belangrijkste softwaresystemen, in de sfeer van het bestrijden van zorgfraude die momenteel op de Nederlandse markt worden aangeboden. In een studie van Forrester Research worden FICO, IBM en SAS tot de top van *predictive analysis* en *datamining* systemen gerekend (zie Kobielus 2010)[8] en in de praktijk worden oplossingen soms naast elkaar gebruikt.[9]

2.3 Beantwoording onderzoeksvraag 2

Hoe simpel of ingewikkeld is het om met deze systemen aan de slag te gaan? Detecteren de systemen ook zwakheden in onderliggende wetgeving? Geven de systemen ook een verklaring voor afwijkende patronen? Zijn deze systemen gemakkelijk toepasbaar in de zorg?

2.3.1 Complexiteit systemen

De vraag '*Hoe simpel of ingewikkeld is het om met deze systemen aan de slag te gaan*' hebben wij benaderd vanuit de implementatie van het systeem, het onderhoud van het systeem en het feitelijk gebruik door de eindgebruiker.

Ten aanzien van de complexiteit van de implementatie van de diverse softwaresystemen is het allereerst van groot belang te benadrukken dat deze softwarepakketten geenszins te vergelijken zijn met, bijvoorbeeld, een Microsoft Office applicatie op een desktop computer. Er is geen sprake van software die eenvoudigweg vanaf een cd-rom geïnstalleerd kan worden; het betreft maatwerk en de complexiteit is daarom ook sterk afhankelijk

[8] Wij merken op dat daar wat betreft FRISS geen conclusie aan mag worden gehecht, gelet op de wijze van selecteren van de door Forrester Research in het betreffende onderzoek geëvalueerde systemen.
[9] Zie Conz 2009.

van de gekozen configuratie. Met name indien gekozen wordt voor een *real time* oplossing die draait binnen de eigen organisatie zal de implementatie een aanzienlijke inspanning vergen van de organisatie alsook mogelijke aanpassingen van, en investeringen in, de hardware. De software zal veelal draaien op een (*dedicated*) server of zelfs meerdere servers en connecties zullen moeten worden gebouwd tussen de bestaande database waarin de declaraties worden verwerkt en het fraudedetectiepakket. De wijze waarop dat plaats kan vinden, zal per pakket en per bestaand systeem verschillen.

Indien daarentegen gekozen wordt voor een analyse *after-the-fact,* die ook nog eens plaatsvindt in een datacentrum van de leverancier van het softwarepakket, zal de implementatie aanzienlijk minder complex zijn. In dat geval zijn soft- en hardware aanpassingen binnen de organisatie zeer waarschijnlijk niet nodig en zal alleen een periodieke *dump* van de declaratiegegevens gemaakt dienen te worden en worden aangeleverd bij de provider van de service. Bij één van de in het onderzoek betrokken zorgverzekeraars heeft het feit dat een dergelijke oplossing een minimale impact op de organisatie heeft, een belangrijke rol gespeeld in het beslissingsproces om te kiezen voor een *hosted solution*.

Het beheer van het systeem zal een zelfde verschil van complexiteit tussen de verschillende oplossingen laten zien. Een oplossing die extern gehost wordt brengt voor de organisatie, anders dan het *up-to-date* houden van de systematiek van de *datadump*, nauwelijks tot geen onderhoud met zich mee. Is het softwarepakket echter geïnstalleerd binnen de organisatie dan kan dit aanzienlijk server- en databaseonderhoud met zich meebrengen.

Ook in het gebruik van de systemen zullen, afhankelijk van de gekozen configuratie, verschillen bestaan in het gemak waarmee men daarmee aan de slag kan. Van de besproken systemen zijn tijdens het onderzoek FICO en IBM Modeler elk bij een zorgverzekeraar in de praktijk gezien. IBM NAMS, het FRISS Fraudeplatform en het SAS *Fraud Framework* worden nog niet (operationeel) gebruikt door een Nederlandse zorgverzekeraar voor de detectie van fraude in declaratiegegevens.

Voor een fraudeonderzoeker bestaat de output van FICO uit een *ranking* van zorgaanbieders waar op de hoogste plaats de zorgaanbieder staat waarvan op basis van de declaraties en (in de mondzorgmodule) 800 indicatoren de grootste kans op potentieel afwijkend declaratiegedrag bestaat. De onderzoeker kan elke zorgaanbieder in de lijst selecteren en kan dan via hyperlinks inzoomen tot detailniveau waarbij de redenen voor de afwijking zichtbaar zijn en tevens de afwijking per reden in verhouding tot andere declaranten gevisualiseerd kan worden. De gegevens van het lopende jaar zijn benaderbaar alsook de gegevens voor de afgelopen 5 jaar (indien aanwezig in de dataset) en analyses kunnen dus ook in de tijd worden gezet.

Het gebruik van IBM Modeler is iets complexer dan FICO aangezien het pakket eigenlijk bedoeld is voor analisten die op basis van de rauwe data en hypothesen over mogelijke fraude zelf alle analyses kunnen

maken. IBM Modeler is op zich een gebruiksvriendelijk datamining en analyse systeem, maar de gebruiker dient wel zelf de productieomgeving te snappen (welke data is voorhanden) alsook kennis te hebben van de zorgsector om effectieve en efficiënte analyses te kunnen maken. Er is dus een echte analist nodig, niet iedere medewerker kan snel aan de slag met dit softwarepakket. Na een initiële keuze voor een bepaald type analyse, bijvoorbeeld de gemiddelde kosten per patiënt van een bepaald type zorgverlener, geeft het pakket een visuele weergave van de verdeling en kan op individuele zorgverleners worden ingezoomd en kunnen keuzes worden gemaakt voor verder onderzoek. Het is mogelijk vaste analyses, zogenoemde *streams* te definiëren op bijvoorbeeld een bepaalde sector (bv. fysiotherapeuten of mondzorg) die periodiek worden uitgevoerd.

Een softwarepakket, zoals FICO, waarin de detectiemodellen voor zorgfraude opgenomen zijn, vergt over het algemeen minder specifieke (statistische) kennis en levert output waar een fraudeonderzoeker direct mee aan de slag kan. Het nadeel van deze pakketten kan zijn gelegen in het feit dat de zorgverzekeraar afhankelijk is van de beschikbare *rules* en modellen van de leverancier en dat meer diepgaande analyse aan de hand van opgedane kennis niet direct kan worden benut. Een pakket als IBM Modeler vereist daarentegen meer specialistische kennis om het pakket en de beschikbare functies goed te implementeren en te configureren voor het specifieke doel. Daar staat tegenover dat die benodigde kennis dan ook in de organisatie beschikbaar is (en blijft) voor de fraudedetectie en -bestrijding. Tijdens de expertmeeting waren de fraudecoördinatoren vrijwel unaniem van mening dat voor een goede fraudedetectie een *full-time* analist belangrijker is dan het type softwarepakket dat in gebruik is. Ongeacht de gebruikte datamining / fraudedetectie software zal de fraudeonderzoeker wel voldoende domeinkennis moeten hebben om de output van de softwarepakketten ook inhoudelijk aan te wenden voor het onderzoek.

2.3.2 Detectie en verklaringen

Zoals eerder aangegeven detecteren de verschillende softwarepakketten potentiële fraude op diverse wijzen. Indien declaraties of een zorgaanbieder door middel van datamining als potentieel frauduleus worden bestempeld, is in de output veelal mogelijk in te zoomen op de gegevens – tot op regelniveau in een (aantal) declaratie(s) – die ten grondslag liggen aan de gevonden afwijking. Vaak wordt de output visueel gepresenteerd en kan men direct verder inzoomen. Echter een onomstotelijke verklaring voor eventueel ontdekte afwijkingen en verbanden kan niet worden gegeven. Dit is een algemeen kenmerk van datamining (zie Seifert 2004: 3). Hoewel datamining wel aanwijzingen voor potentiële fraude, zoals bijvoorbeeld *upcoding*[10], kan opleveren, zal de daadwerkelijke verklaring voor gedetecteerde afwijkingen en verbanden altijd door nader onderzoek aan het licht moeten komen. Bijvoorbeeld, een tandarts kan opvallen omdat deze altijd bij alle

[10] Het declareren van een duurdere variant van de werkelijk verleende zorg.

patiënten de UTP[11] code voor uitgebreide gebitsreiniging declareert in afwijking van tandartsen in vergelijkbare situaties. Hiervoor zijn verschillende verklaringen mogelijk, waaronder frauduleuze *upcoding* maar evengoed ook de situatie dat de tandarts vanuit een preventieve visie bij al zijn patiënten daadwerkelijk altijd uitgebreide gebitsreiniging uitvoert. Een dataminingsysteem kan de juiste verklaring voor dergelijke afwijkingen niet geven.

Hetzelfde geldt met betrekking tot de detectie van zwakheden in de wetgeving. Zo zal bij een fraude, waarbij misbruik gemaakt wordt van zwakheden in de onderliggende wetgeving, mogelijk een bepaalde zorgaanbieder opvallen omdat deze een declaratiepatroon heeft dat afwijkt van vergelijkbare zorgaanbieders met vergelijkbare patiënten. Echter, dat de verklaring in een zwakheid van de onderliggende wetgeving moet worden gezocht, zal ook pas na inhoudelijk onderzoek van de reden(en) voor afwijking duidelijk worden. Uiteraard zal het bij een éénmaal geïdentificeerde zwakheid in de wetgeving door middel van datamining eenvoudig zijn alle zorgaanbieders te identificeren die van die kennelijke zwakheid gebruik maken maar datamining software is geen vervanging van analisten en fraudeonderzoekers.

2.3.3 Toepasbaarheid in de zorg

Uit de literatuur op het gebied van datamining ten behoeve van fraudedetectie in de zorg blijkt dat met name hoge verwachtingen bestaan van *Supervised classification systems* (Travialle *et al.* 2011: 8):

"*Supervised classification models are particularly appropriate for use in health care fraud, as they can be trained and adjusted to detect sophisticated and evolving fraud schemes.*"

Maar Travaille *et al.* merken tegelijkertijd op:

"*The drawback to these techniques is that new fraud schemes are not immediately detectable due to the lag of discovering and labeling new fraud in training data",* en, *"However, no one technique, supervised or unsupervised, is applicable to discover all fraud strategies and schemes. A fraud detection system consisting of multiple techniques, with a flexible, modular approach capable of adapting to the continuous changes in the fraud detection field, must be employed to effectively combat fraud and abuse.*"

Met uitzondering van IBM Modeler wordt in alle andere besproken softwarepakketen gebruik gemaakt van meerdere analysetechnieken die naast elkaar worden gebruikt. Daarmee zijn – als we Travaille *et al.* volgen - al deze pakketten makkelijk toepasbaar voor datamining ten behoeve van fraudedetectie in de zorg. Het feit dat deze pakketten ook al daadwerkelijk worden toegepast voor fraudedetectie in de zorg – zij het nog niet allemaal in Nederland – ondersteunt dit. De fraudecoördinatoren van de zorgverzekeraars waar FICO en IBM

[11] Uniforme Particuliere Tarieven

Modeler in gebruik zijn, bevestigen dat de functionaliteit van de pakketten voldoet aan hun verwachtingen en dat de pakketten daadwerkelijk potentiële fraudes blootleggen die op andere wijze mogelijk niet gevonden waren. In de *proof-of-concept* die FICO voor één van de zorgverzekeraars uitvoerde, werden alle reeds bekende fraudegevallen geïdentificeerd, alsmede een zelfde aantal dat nog niet door de zorgverzekeraar was ontdekt. Een vergelijking tussen de effectiviteit en efficiëntie van de pakketten is op basis van dit onderzoek niet te maken.

Wat in het algemeen ten aanzien van de bruikbaarheid van dataminingsoftware voor het detecteren van fraude geldt, is dat het van belang is dat de beschikbare data voldoende variëteit en kwaliteit heeft. In data met een hoog aggregatieniveau vallen onderliggende patronen en verbanden niet of nauwelijks te ontdekken. Een praktisch voorbeeld hiervan is het – nu nog – ontbreken van de verplichting voor tandartsen om de element-nummers in de declaraties te vermelden. De elementen kunnen daarom niet in een analyse betrokken worden. Bovendien wordt vaak gedeclareerd op één code van de tandartspraktijk waardoor de individuele zorgverlener niet aan de declaratieregel is gekoppeld en niet in de analyse kan worden betrokken. Daarnaast is zorg die verleend wordt in ziekenhuizen lastig te onderzoeken, omdat de variëteit in de gegevens, na de overgang van declaraties op basis van verrichtingen naar declaraties op basis van DBCs (en sinds 1 januari 2012, DOTs) onvoldoende is. In gedetailleerde specificaties van behandelingen kunnen door middel van datamining afwijkingen sneller worden gedetecteerd, maar bij declaratie op basis van (nu) DOTs wordt (vrijwel) niets meer gespecificeerd.. Tenslotte bevat ook AWBZ declaratiedata volgens geïnterviewden relatief weinig detailinformatie waardoor de effectiviteit van datamining beperkter is dan wanneer er meer detail beschikbaar zou zijn.

Het grote voordeel dat het Nederlandse zorgsysteem heeft, is dat meer dan 95% van de zorgdeclaraties centraal en digitaal via VECOZO bij de zorgverzekeraars worden ingediend. Om deze informatiestromen zo uniform mogelijk te laten verlopen, zijn vaste Externe Integratie (EI) standaarden ontwikkeld (Vektis 2011) op basis waarvan uniforme gegevens beschikbaar zijn. Om de kans op succesvol gebruik van datamining zo hoog mogelijk te maken, zouden de EI standaarden opname van zo veel mogelijk gegevens over de verleende zorg, de zorgverlener, de verzekerde, en andere relevante gegevens (b.v. verwijzingen) verplicht moeten stellen in de declaraties. Vanzelfsprekend zal wel een balans gezocht moeten worden met de administratieve last die door uitgebreide(re) EI standaarden eventueel bij de zorgverleners komt te liggen.

2.4 Beantwoording onderzoeksvraag 3

• Wat zijn de effecten van deze systemen? Welke voor- en nadelen hebben ze? Welke vaardigheden hebben de gebruikers nodig om ermee te kunnen werken? Wat zijn de kosten van deze systemen (aanschaf en licentie?) Hoe makkelijk / moeilijk zijn ze te implementeren en wat kost dat? Hoe makkelijk / moeilijk zijn ze uit te breiden? Hoe gebruiksvriendelijk zijn ze? Welke onderdelen van de zorg lenen zich het beste voor instrumenten als datamining?

2.4.1 Effecten en voor en nadelen

Het effect van het gebruik van fraudedetectiesystemen is niet eenduidig te benoemen. Om te beginnen geven de fraudecoördinatoren aan dat, op basis van de bestaande systemen van fraudedetectie, incidentmeldingen en onderling (via ZN) uitgewisselde informatie, al een grote hoeveelheid vermoedelijke fraude bekend wordt. De afgelopen jaren is bij de zorgverzekeraars geïnvesteerd in fraudebestrijding met als gevolg dat de meeste fraudecoördinatoren over een behoorlijke werkvoorraad beschikken, waardoor onderzoeken geprioriteerd moeten worden. Meer detectie van potentiële fraude zou daardoor niet per definitie tot meer bewezen fraude leiden, wel zou een meer systematische en gestructureerde detectie kunnen bijdragen aan een betere prioritering van de te onderzoeken fraudesignalen.

Het is in dit verband van belang te onderkennen dat de definitie van fraude, die door de zorgverzekeraars wordt gehanteerd, de kenmerken heeft dat sprake moet zijn van opzettelijk handelen, waarbij een regel wordt overtreden en financieel voordeel wordt behaald. In de praktijk blijkt volgens de fraudecoördinatoren een groot aantal regels ten aanzien van het declareren voor meerdere uitleg vatbaar te zijn, met als gevolg dat per geval daadwerkelijk onderzoek noodzakelijk is om vast te stellen of inderdaad sprake is van het opzettelijk overtreden van een regel. Daarbij ontstaan volgens geïnterviewden vaak tijdrovende discussies met zorgaanbieders over de wijze van declareren. Veel vermoedens van fraude leiden derhalve niet tot een vaststelling van fraude, maar vergen wel de nodige onderzoekscapaciteit. De eventuele toegevoegde waarde van datamining voor het (semi)geautomatiseerd detecteren van fraude kan daarom vooral worden verwacht in een efficiëntere prioriteitstelling waarbij een eventuele afhankelijkheid van (toevallige) incidentmeldingen waarschijnlijk zal afnemen.

Het is overigens niet te verwachten dat het gebruik van meer en/of betere systemen voor fraudedetectie heel snel een zichtbaar effect in de fraudecijfers zullen hebben, tenzij in de gehele achterliggende keten eveneens wordt geïnvesteerd in het oppakken van de gegenereerde fraudesignalen. Dit betekent naar verwachting dat, los van de kosten-/batenafweging, meer geavanceerde detectie slechts in combinatie met investering in meer onderzoekscapaciteit zal leiden tot meer aangetoonde fraude. Het is daarnaast te verwachten dat als gevolg van het onderscheid bij de zorgverzekeraars tussen formele controle (zie ook paragraaf 2.5.1) en fraudebestrijding een deel van de toegevoegde waarde niet bij de fraudeomvang uitgedrukt in geld, maar bij de

besparingen op basis van formele controle zichtbaar worden. Immers, de kenmerken van éénmaal nieuw ontdekte fraudemethoden kunnen, daar waar mogelijk, worden toegevoegd aan de set regels waarlangs de formele controle plaatsvindt. Op die wijze wordt eventueel toekomstige soortgelijke fraude op efficiënte wijze geweerd. Echter declaraties die in de formele controle worden tegengehouden, worden door zorgverzekeraars niet als frauduleuze declaraties of potentieel frauduleuze declaraties bestempeld en dragen dus niet bij aan 'meer' ontdekte fraude.

Uit het bovenstaande volgt dat het grootste voordeel van (semi)geautomatiseerde fraudedetectie door middel van datamining zeer waarschijnlijk zal liggen in het systematisch(er) en efficiënt(er) kunnen stellen van prioriteiten voor de beschikbare onderzoekscapaciteit. Met behulp van een goede analist en een dataminingsysteem kan direct worden ingezoomd op de meest waarschijnlijke potentiële fraudes en de grootste financiële risico's. De fraudebestrijding wordt daarmee minder afhankelijk van incidentmeldingen en toevallig ontdekte fraudes. Een nadeel van fraudedetectie softwarepakketten zijn de kosten en inspanningen die met implementatie en onderhoud gemoeid zijn, uiteraard afhankelijk van het type pakket dat gekozen wordt. Het is vervolgens per zorgverzekeraar sterk afhankelijk van de daar bestaande situatie in welke mate een commercieel datamining / fraudedetectie softwarepakket toegevoegde waarde heeft boven de activiteiten van een goede analist en reeds bestaande methoden van (semi)geautomatiseerde fraudedetectie. Gelet op de verschillen in organisatie van de fraudedetectie en bestrijding tussen de zorgverzekeraar, is daar geen algemene conclusie over te trekken.

Wel kan generiek worden gesteld dat de grote zorgverzekeraars vanuit hun situatie makkelijker zelf succesvolle geautomatiseerde fraudedetectie kunnen laten plaatsvinden omdat zij beschikken over grotere datasets en over het algemeen een grote(re) geografische dekking. Hierdoor kunnen zij bijvoorbeeld cumulatieve fraudes (meer declareren dan een zorgverlener feitelijk kan uitvoeren) makkelijker detecteren dan de kleinere zorgverzekeraars. Het belang om gebruik te maken van bij andere partijen opgedane kennis is bij kleine zorgverzekeraars derhalve over het algemeen groter. In ZN verband worden momenteel de mogelijkheden onderzocht om op basis van de bij Vektis beschikbare gegevens (zie ook paragraaf 2.5.1) meer structureel gezamenlijk analyses ten behoeve van fraudedetectie te maken.

2.4.2 Vaardigheden en gebruikersvriendelijkheid

De vaardigheden die gebruikers nodig hebben, verschillen per systeem. Zoals eerder aangegeven is voor het gebruik van IBM Modeler, naast domeinkennis over de zorg, gedegen kennis van statistische technieken en vaardigheid in het maken van modellen aan de hand van hypothesen over potentiële frauderisico's nodig. Dat heeft er mede mee te maken dat deze oplossing volledig binnen de eigen organisatie draait en tevens geen

gebruik maakt van reeds bestaande modellen. Anderzijds zal de mondzorgmodule van FICO, die momenteel bij één zorgverzekeraar draait, minder vaardigheden van de eindgebruiker vergen aangezien de resultaten van de analyses kant en klaar worden gepresenteerd.

Anders gezegd, met uitzondering van een configuratie waarbij de gekozen softwareapplicatie buiten de organisatie draait op basis van bestaande modellen, zullen diverse niveaus in vaardigheid van de gebruikers van de software nodig zijn. Van de uiteindelijke fraudeonderzoekers die de door het systeem gegenereerde fraudesignalen krijgen aangeleverd, worden voor het gebruik geen specifieke vaardigheden verlangd anders dan een algemene vaardigheid met geautomatiseerde systemen om te gaan. Echter van de analisten en applicatiebeheerders die de (parameters van de) systemen moeten inregelen en onderhouden, zullen aanzienlijk meer vaardigheden worden verlangd. Analisten zullen voor de verschillende systemen zowel statistische kennis als domeinkennis van de zorg nodig hebben.

De gebruikersvriendelijkheid van de besproken softwarepakketten is niet exact meetbaar. Uit de beschikbare productinformatie lijkt de gebruikersvriendelijkheid hoog te zijn, waarbij veel met visuele output en interfaces wordt gewerkt. De demonstraties van enkele softwarepakketten die tijdens het onderzoek aan ons zijn gegeven, geven eveneens een beeld van een hoge gebruikersvriendelijkheid.

2.4.3 Kosten

De kosten van softwarepakketten voor fraudedetectie kunnen worden onderverdeeld in aanschaf/licentie-, en/of implementatie en onderhoudskosten. De daadwerkelijke kosten van de aanschaf, implementatie, gebruik en onderhoud van de verschillende systemen is sterk afhankelijk van de gekozen oplossing en de bestaande geautomatiseerde omgeving van de zorgverzekeraars.

In de eerste plaats speelt mee of een geheel platform wordt afgenomen of alleen een enkele module. Ook is van belang of gekozen wordt voor een *real time* analyse of een analyse *after-the-fact,* en of de oplossing extern gehost wordt of binnen de organisatie geïmplementeerd. *Real time* oplossingen zullen, met name vanwege het maken van de noodzakelijke koppelingen met de bestaande systemen meer implementatiekosten met zich mee brengen dan wanneer periodiek een dump voor analyse *after-the-fact* wordt aangeleverd. Oplossingen die extern worden gehost brengen veelal kosten met zich mee die gebaseerd zijn op het verwerkte (data)volume maar relatief weinig hardware- en implementatiekosten, terwijl oplossingen die binnen de organisatie worden geïnstalleerd juist wel (aanzienlijke) hardware- en implementatiekosten met zich mee brengen. Een andere belangrijke factor in de kosten is de bestaande geautomatiseerde omgeving en de eventuele aanpassingen die daarin gemaakt moeten worden om de connectie met het softwarepakket voor de fraudedetectie te maken.

Tijdens het onderzoek is getracht informatie over de aanschaf- en licentiekosten van de besproken softwarepakketten te verzamelen. Niet alle vertegenwoordigers van de softwareleveranciers waren bereid inzicht te geven in de aanschaf en licentiekosten die zij in rekeningen brengen. Zij gaven aan dat de kosten sterk afhangen van de exacte keuzes die bij aanschaf en implementatie door de klant worden gemaakt; vrijwel altijd is sprake van een maatwerkoplossing in plaats van een standaardpakket dat tegen een standaardprijs aangeboden kan worden. Vanuit concurrentieoverwegingen zijn de leveranciers daarbij terughoudend in het verstrekken van gedetailleerd inzicht in hun prijsmodel.

Uit de informatie die verzameld kon worden, lijkt dat IBM Modeler qua directe licentiekosten een relatief goedkope oplossing is waarbij moet worden opgemerkt dat het gebruik van dit pakket een goed opgeleide analist vergt, en ook een minder uitgebreide oplossing betreft dan enkele andere besproken pakketten. De implementatie brengt ongeveer € 100k aan directe kosten met zich mee, waarna de terugkerende licentiekosten ongeveer € 20k per jaar bedragen. De kosten van de oplossingen van FRISS en FICO zijn afhankelijk van respectievelijk het premievolume en declaratievolume. Volgens opgave van FRISS zijn de kosten, afhankelijk van het premievolume en het aantal actieve zorgsoorten tussen de € 3k en € 20k totaal per maand. Volgens opgave van FICO zijn de kosten voor de service tussen de 10 en 15 cent per declaratieregel in een degressief tariefstelsel.

Samenvattend, het is in ons onderzoek niet mogelijk gebleken een reële schatting te maken van (implementatie)kosten aangezien deze te afhankelijk zijn van de te maken keuzes en van de bestaande geautomatiseerde omgeving waarin deze moet worden geïntegreerd. Analyse van gegevens die maandelijks in één batch uit de systemen wordt gehaald en bijvoorbeeld in IBM Modeler of door FICO worden geanalyseerd, zijn eenvoudig te implementeren en brengen de minste implementatiekosten met zich mee. Een oplossing die binnen de organisatie draait waarbij koppelingen moeten worden gemaakt met de database van het declaratiesysteem, en de declaraties *real-time* op potentiële fraude worden geanalyseerd zal aanzienlijke implementatiekosten met zich meebrengen. Het is reëel om aan te nemen dat een pakket dat *real-time* gegevens analyseert bij een grote verzekeraar in ieder geval snel enige tonnen (euro) aan implementatiekosten zal meebrengen.

2.4.4 Bruikbaarheid datamining

Zoals in paragraaf 2.3.3. al aangegeven, bleek dat de besproken softwarepakketten voor datamining ten behoeve van fraudedetectie zonder meer toepasbaar zijn in de zorg. De mate waarin een onderdeel van de zorg zich meer of minder goed leent voor het gebruik van datamining technieken om potentiële fraude te detecteren is – net als elke andere situatie waar men datamining zou willen toepassen - met name afhankelijk

van de hoeveelheid gegevens die beschikbaar is, de variëteit in de gegevens en de kwaliteit van de gegevens. Simpel gezegd, hoe minder detail, hoe kleiner de potentiële toegevoegde waarde van datamining bij fraudedetectie. Het is bijvoorbeeld lastiger een frauderende specialist in een ziekenhuis te identificeren als alle declaraties op de AGB[12] code van een maatschap worden ingediend. De AGB code is dan geen onderscheidend kenmerk meer in de gegevens.

Een andere relevante factor voor de bruikbaarheid van datamining in de zorg, is de beschikbaarheid van historische gegevens met name als in de analyses (sterk) gesteund wordt op *Supervised Classification Techniques*. Bij wijzigingen in de declaratiesystematiek zijn historische gegevens niet altijd meer bruikbaar om als referentiekader te dienen waartegen afwijkende patronen herkend kunnen worden. Dit is in hogere mate het geval bij systemen die gebruik maken van modellen die mede gebaseerd zijn op referentiedata uit andere landen. Bijvoorbeeld, vanwege de wijzigingen in de UTP codes voor mondzorg in Nederland per 1 januari 2012, zal het voor de mondzorgmodule die door FICO wordt aangeboden gemiddeld enige maanden duren voordat voldoende gegevens met de nieuwe codes beschikbaar is om met voldoende betrouwbaarheid relevante afwijkende patronen te ontdekken. Omdat Nederland met de nieuwe UTP codes gaat afwijken van de internationale standaard zal referentiedata uit het buitenland dan eventueel handmatig aan de modellen moeten worden toegevoegd.

Samenvattend, in feite is datamining bruikbaar voor fraudedetectie binnen alle onderdelen van de zorg maar de meeste (betrouwbare) resultaten zijn te verwachten in de onderdelen van de zorg waar de meest gedetailleerde declaratiegegevens voorhanden zijn. Gelet op de huidige declaratiestructuur is te verwachten dat datamining op bijvoorbeeld ziekenhuiszorg- en AWBZ-gegevens minder snel en betrouwbaar afwijkingen en verbanden zal detecteren dan binnen bijvoorbeeld de mondzorg.

[12] Algemeen Gegevens Beheer

2.5 Beantwoording onderzoeksvraag 4

• Hoe komt het dat deze systemen nog niet op grote schaal worden gebruikt door zorgverzekeraars? Wat zijn hier kansen en belemmeringen? Met welke argumenten zijn verzekeraars over te halen op grote schaal gebruik te maken van deze systemen?

2.5.1 Huidige situatie

Om de vraag te beantwoorden (of, en zo ja,) waarom dataminingsystemen nog niet op grote schaal worden gebruikt door zorgverzekeraars, is het van belang eerst de bestaande situatie van fraudedetectie en fraudebestrijding door zorgverzekeraars te schetsen.

De eerste stappen met datamining / geautomatiseerde fraudedetectie door zorgverzekeraars zijn gezet in 2005/2006. Bij de landelijke inventarisatie van het fraudebeheersingsbeleid door ZN over 2010 gaf driekwart van de zorgverzekeraars aan geautomatiseerde systemen te gebruiken bij de detectie van fraude (ZN 2011). Uit dit onderzoek blijkt dat alle zorgverzekeraars die participeerden, gebruik maken van enig systeem van datamining / geautomatiseerde fraudedetectie alhoewel er grote verschillen bestaan in de mate waarin deze systemen gebruikt worden, hoelang men de systemen al in gebruik heeft, en of het commerciële software pakketten betreft of een eigen ontwikkeld systeem of methodiek.

Zo is bij de meeste zorgverzekeraars in de software die de declaraties verwerkt een *rule based* detectie ingebouwd die soms bestaat uit vele duizenden regels. Deze controles worden door de zorgverzekeraars als 'formele controle' gedefinieerd en zijn primair gericht op het vaststellen of de gedeclareerde zorg onder de verzekerde prestaties valt. Echter, in deze formele controles wordt in de praktijk – waarbij uiteraard verschillen bestaan tussen de zorgverzekeraars - ook potentiële fraude gedetecteerd en tegengehouden. Voorbeelden hiervan zijn de controle op dubbel ingediende declaraties en de eerder genoemde controle op 'onmogelijke' behandelingen, zoals een wortelkanaalbehandeling bij een driejarige. Door middel van deze systemen hebben de zorgverzekeraars in 2010 naar schatting € 1.1 miljard aan onterecht gedeclareerde kosten tegengehouden (ZN 2011).[13]

De volgende stap in het controle proces is de 'materiële controle' waarbij het gaat om het vaststellen of de geleverde zorg in overeenstemming is met de declaratie. Het vaststellen daarvan kan niet geautomatiseerd plaatsvinden, hoewel aanwijzingen voor het feit dat een declaratie een ander beeld dan de werkelijkheid geeft wel uit de declaratiegegevens gehaald kunnen worden. Een voorbeeld daarvan is *upcoding* waarbij een zorgverlener een duurdere variant van de zorg declareert dan hij/zij daadwerkelijk heeft verleend.

[13] Volgens de toelichting van ZN werd in eerste instantie € 2.3 miljard aan onjuiste declaraties tegengehouden waarvan 52% later alsnog op juiste wijze werd ingediend. Hieruit volgt dat 48% oftewel € 1.1 miljard onrechtmatig was ingediend.

Bij de meeste zorgverzekeraars vindt naast de (semi)automatische formele controle in het declaratieproces ook achteraf (semi)geautomatiseerde detectie van afwijkingen plaats ten behoeve van materiele controle en fraudedetectie. Hiervoor wordt niet altijd een specialistisch pakket voor datamining ten behoeve van fraudedetectie gebruikt. Vaak wordt gebruik gemaakt van andere (*business analytics*) software die in de organisatie beschikbaar is; ook zonder de specialistische fraudepakketten kunnen aanwijzingen voor fraude worden ontdekt (zie bijvoorbeeld Hasaart 2011). De twee in het onderzoek betrokken verzekeraars waar momenteel nog geen (structurele) analyses achteraf op de declaratiedata plaatsvindt, overwegen wel de aanschaf van een commercieel softwarepakket en zijn in het stadium van een *proof-of-concept*.

Tenslotte vindt incidenteel analyse van zorgdeclaratiegegevens ten behoeve van fraudebestrijding plaats door Vektis. Vektis beheert namens de zorgverzekeraars de declaratiegegevens en maakt op basis daarvan analyses over de kosten en de kwaliteit van de gezondheidszorg in Nederland. Vektis krijgt daartoe van de zorgverzekeraars en de zorgkantoren achteraf vrijwel alle gegevens van verzekerde zorg in Nederland aangeleverd en heeft in 2011 ook ondersteuning van fraudebeheersing toegevoegd aan het dienstenpakket (Vektis 2011). Analyses met een focus op fraudedetectie en onderzoek worden door Vektis gemaakt op speciaal verzoek vanuit de fraudewerkgroep van ZN. Op basis van, onder meer, afspraken tussen de zorgverzekeraars uit 2009 over samenwerking ten aanzien van fraudebestrijding, wordt momenteel ook onderzocht of binnen Vektis op landelijk niveau meer proactief datamining op de zorgdeclaratiegegevens kan plaatsvinden. Bij Vektis wordt voor de business analyses gebruik gemaakt van SAS *business analytics* software en van IBM software.

2.5.2 Gebruik systemen door zorgverzekeraars

Uit de bevindingen van het onderzoek blijkt dat commerciële dataminingsystemen nog niet op grote schaal (door ons geïnterpreteerd als door een meerderheid) worden gebruikt door zorgverzekeraars. Zonder verdere nuance geeft deze constatering echter geen getrouw beeld van de werkelijkheid van de (semi)geautomatiseerde fraudedetectie door zorgverzekeraars. Hiervoor zijn drie redenen.

In de eerste plaats is in het onderzoek gebleken, zoals hiervoor in 2.5.1 weergegeven, dat bij een groot aantal verzekeraars op andere wijze dan met commerciële softwarepakketten, wel degelijk semi(geautomatiseerde) fraudedetectie plaatsvindt. Dat vindt dan zowel plaats door middel van de controleregels die in de declaratiesystemen zijn ingebouwd, als door middel van analyse achteraf. Hoewel de controleregels in de declaratiesystemen primair zijn bedoeld voor de formele controle, worden deze in de praktijk ook gebruikt voor fraudedetectie. Daarnaast vindt bij een aantal zorgverzekeraars fraudedetectie plaats door middel van analyse achteraf waarbij analisten andersoortige software gebruiken.

In de tweede plaats is een toename van de aandacht voor fraudedetectie en –bestrijding te zien, zowel in het aantal zorgverzekeraars dat commerciële softwarepakketten gebruikt of dat momenteel overweegt, als in de algemene aandacht die fraudedetectie en –bestrijding bij de zorgverzekeraars krijgt. De geïnterviewde bestuurders van de zorgverzekeraars geven in dat verband aan dat in het verleden mogelijk minder expliciete aandacht voor fraudebestrijding bestond maar dat dit in de afgelopen jaren aanzienlijk is ingehaald. De fraudedetectie en –bestrijding is mogelijk nog niet overal optimaal maar krijgt nu wel gestructureerde aandacht bij alle zorgverzekeraars. Recent is structurele aandacht voor fraudebestrijding ook gestandaardiseerd in het 'protocol betreffende bewustwording, preventie, detectie en afhandeling van verzekeringsfraude en criminaliteit' (Verbond van Verzekeraars en Zorgverzekeraars Nederland, 2011). Een ander aspect dat daarbij van belang is, is de risico gestuurde aanpak die onder andere door de Solvency II richtlijnen wordt voorgeschreven, waardoor het frauderisico voor een zorgverzekeraar niet meer te negeren is.

Het is mogelijk dat de jaarlijks door ZN gepubliceerde fraudecijfers in vergelijking met zorgfraudecijfers uit het buitenland het beeld opwekken dat door de zorgverzekeraars weinig aan fraudedetectie wordt gedaan. In 2010 werd voor ´slechts´ € 6,2 miljoen aan fraude ontdekt. Dit lage cijfer (gelet op de € 60 miljard die in de zorg omgaat) is mede het gevolg van de relatief restrictieve definitie van fraude die door de zorgverzekeraars wordt gehanteerd. In Nederland worden alleen die gevallen waarin opzettelijk onrechtmatig wordt gedeclareerd en waarbij financieel gewin ontstaat als fraude bestempeld. Dit moet dan ook nog aantoonbaar zijn. De fraudecoördinatoren van de verschillende zorgverzekeraars geven bovendien aan dat het primaire doel is om onrechtmatig ingediende declaraties niet uit te betalen en dus niet zozeer om zoveel mogelijk fraude op te sporen maar om deze vroegtijdig tegen te houden. Veel van reeds genoemde €1.1 miljard aan niet terechte declaraties die in het Nederlandse systeem in de formele controle worden tegengehouden, en de €106 miljoen (zie ZN 2011) die na een materiële controle als onterecht gedeclareerd worden teruggevorderd, zouden bij gebruik van een minder restrictieve definitie waarschijnlijk onder de noemer fraude kunnen worden geschaard. Gelet op de grote verschillen in systemen en definities is daarom het simpelweg vergelijken met fraudecijfers uit het buitenland ook niet mogelijk.

De derde en laatste nuance is dat uit de bevindingen van dit onderzoek blijkt dat datamining niet als panacee voor de fraudedetectie en -bestrijding in de zorg mag worden gezien. De zorgverzekeraars wijzen er met nadruk op dat, naast de verantwoordelijkheid die zij zelf hebben en nemen voor fraudebestrijding, ook een aantal factoren buiten hun invloedsfeer ligt. Als voorbeelden worden genoemd de soms beperkte fraudebestendigheid van de regelgeving, de betrekkelijke eenvoud waarmee erkenningen als zorgaanbieder verkregen kunnen worden, en de wijze waarop indicaties worden gesteld. De grootschalige fraudes met het persoonsgebonden budget leveren schrijnende voorbeelden op en enkele fraudecoördinatoren van de zorgverzekeraars geven aan het gevoel te hebben dat hun signalen daarover aan de overheid niet serieus genomen zijn, en worden. Die signalen betreffen niet alleen de bovengenoemde factoren voorin de keten,

maar ook de achterzijde van de keten waar met betrekking tot de bij justitie en de NZa aangedragen zaken in ieder geval vaak de perceptie bestaat dat daar onvoldoende gevolg aan wordt gegeven. Overigens wordt bij de NZa - op één voorbeeld na - het beeld dat concrete zaken zijn blijven liggen, niet herkend.

Een ander belangrijk aspect dat in deze discussie nadrukkelijk naar voren is gebracht, is de zorgplicht die de verzekeraars hebben ten opzichte van de verzekerden. Het niet meer accepteren van declaraties van een frauderende zorgaanbieder kan ertoe leiden dat de verzekerden de factuur rechtstreeks krijgen en die op basis van restitutie aan de zorgverzekeraar sturen, die vervolgens verplicht is de factuur te voldoen. Tijdens het onderzoek bleek dat over de mogelijkheid van het weigeren van betalen van facturen van zorgaanbieders door middel van uitsluiting in de polisvoorwaarden een verschil van perceptie bestaat tussen het ministerie en de zorgverzekeraars. Het alternatief voor de zorgverzekeraars om facturen van zorgaanbieders met een akte van cessie van de verzekerde over te nemen en vervolgens alsnog de strijd met de zorgaanbieder aan te gaan, levert een significante administratieve last op. Die last had volgens de fraudecoördinatoren voorkomen kunnen worden als betreffende zorgaanbieder na duidelijke signalen van fraude zijn erkenning had verloren, of beter, deze door deugdelijke toetsing vooraf nooit had verkregen.

Het belang van goede controle en fraudedetectie – waar datamining een belangrijke bijdrage aan kan leveren – wordt door de zorgverzekeraars zeker onderkend maar heeft volgens de geïnterviewde fraudecoördinatoren alleen zin indien deze in balans is met andere inspanningen in de gehele keten van fraudepreventie, detectie en bestrijding, zowel door de zorgverzekeraars, als door andere betrokken partijen.

2.5.3 Kansen en belemmeringen

Ondanks de nuance zoals weergegeven in punt 3 van de vorige paragraaf, bestaat een positieve houding bij de zorgverzekeraars ten opzichte van datamining voor fraudedetectie die een duidelijke kans biedt voor een groter gebruik daarvan. De zorgverzekeraars geven bijvoorbeeld aan ook niet-kwantificeerbare effecten van fraudedetectie – zoals het preventieve effect – mee te nemen in investeringsbeslissingen. Hoewel er wel een logische business case voor de investeringen moet zijn, gaat het volgens hen bij de fraudebestrijding niet alleen om de harde resultaten op korte termijn. Een dergelijk uitgangspunt wordt ook geuit door de zorgverzekeraars met betrekking tot de aanpak van AWBZ fraude waar, ondanks dat voor de bestrijding hiervan geen financiële prikkel bij de verzekeraars bestaat, hierin toch een aanzienlijke hoeveelheid onderzoekscapaciteit wordt gestoken.

Daarnaast hebben enkele recente incidenten bij zorgverzekeraars geleid tot een concrete waardering van de voordelen die (semi)geautomatiseerde fraudedetectie door middel van datamining kan bieden; men beseft dat

de aanzienlijke fraudes, die bij toeval ontdekt zijn veel eerder ontdekt en voorkomen hadden kunnen worden met behulp van datamining. Het benutten van leermomenten uit incidenten biedt een duidelijke kans voor het vergroten van de bewustwording van het potentieel van datamining.

Anderzijds bestaan ook vragen ten aanzien van de te verwachten meeropbrengsten van de veelal relatief kostbare commerciële datamining softwarepakketten ten behoeve van fraudedetectie ten opzichte van de al bestaande maatregelen. De besparing op de zorgkosten wordt immers pas gerealiseerd als de investering daadwerkelijk tot afwijzing (vooraf) of terugvordering (achteraf) leidt en niet door de detectie van potentiële fraude op zich. Bovendien kan gesteld worden dat een bovengrens bestaat waarbij meer investeringen in de fraudedetectie en –bestrijding niet tot meer besparingen leidt omdat op enig moment de kosten van maatregelen voor fraudedetectie en –bestrijding de besparingen te boven gaan. De fraudecoördinatoren geven zelf aan dat men nog geen zicht heeft of die bovengrens al bereikt is, maar ook dat de belemmering veelal zit in de beschikbare onderzoekscapaciteit, die nodig is om de fraudesignalen op te volgen.

De bevindingen geven, voor zover na te gaan[14], onvoldoende houvast te concluderen dat een direct verband bestaat tussen het gebruik van een (extern) datamining softwarepakket en de hoeveelheid fraude die door een zorgverzekeraar ontdekt wordt. Daarnaast is van één zorgverzekeraar bekend dat men het gebruik van een commercieel datamining softwarepakket heeft beëindigd en vervolgens door middel van analyse binnen de eigen bedrijfsprocessensystemen door kundige analisten naar eigen zeggen tot betere detectieresultaten is gekomen.

2.5.4 Bevordering gebruik van datamining

Op basis van alle bevindingen tezamen, is niet eenduidig antwoord te geven op de vraag met welke argumenten de individuele zorgverzekeraars over zijn te halen op grote schaal gebruik te maken van datamining systemen. Uit het onderzoek blijkt dat zowel de organisatorische inbedding van de fraudedetectie en –bestrijding, als het daarvoor gebruikte instrumentarium sterk wisselt per zorgverzekeraar. Een aantal zorgverzekeraars beschikt over goed werkende bestaande systemen, waardoor overstappen op een commercieel fraudedetectie/datamining softwarepakket niet snel toegevoegde waarde zal hebben. De zorgverzekeraars waar dit wel een toegevoegde waarde kan hebben, zijn zich daarvan al zelf bewust geworden, soms enigszins geholpen door incidenten waarvan men achteraf inziet dat die met detectiesoftware voorkomen hadden kunnen worden. Het is overigens niet mogelijk één van de besproken softwarepakketten

[14] Niet alle zorgverzekeraars waren bereid ten behoeve van dit onderzoek gedetailleerde cijfers over besparingen en vastgestelde fraude te verstrekken. Het meest gehanteerde argument daarvoor was dat de cijfers - gelet op de grote verschillen tussen de verzekeraars – onvoldoende onderling vergelijkbaar zijn.

aan te wijzen dat primair en bij uitstek het meest geschikt zou zijn voor fraudedetectie bij zorgverzekeraars; elk van de op de markt beschikbare pakketten heeft voor- en nadelen en het kan per zorgverzekeraar verschillen hoe die gewogen (moeten) worden.

3. Aanbevelingen

De centrale bevinding van dit onderzoek is dat de toegevoegde waarde van datamining voor het (semi)geautomatiseerd detecteren van fraude vooral kan worden verwacht in het systematisch(er) en efficiënt(er) kunnen stellen van prioriteiten voor de beschikbare onderzoekscapaciteit waarbij een eventuele afhankelijkheid van (toevallige) incidentmeldingen waarschijnlijk zal afnemen. Aansluitend hierop, en ter beantwoording van de vraag wat door de verschillende partijen nog meer gedaan zou kunnen worden op het gebied van preventie en –bestrijding van fraude in de zorg, kunnen wij op grond van onze bevindingen in dit onderzoek de volgende aanbevelingen geven.

1. Geautomatiseerde fraudedetectie centraal organiseren

Gelet op de belangrijkste kenmerken van, en voorwaarden voor, het gebruik van dataminingsoftware voor de detectie van fraude in de zorg, kan de aanbeveling worden gedaan om waar mogelijk geautomatiseerde fraudedetectie op centraal niveau te organiseren. Vanwege de grotere hoeveelheid beschikbare gegevens is de kans op betrouwbare en bruikbare resultaten van datamining groter indien dat centraal plaatsvindt. Ook cumulatieve fraudes die een enkele zorgverzekeraar, op basis van alleen de eigen gegevens, veelal niet kan ontdekken, komen hiermee vermoedelijk eerder aan het licht. Zoals eerder gemeld, wordt de mogelijkheid tot een meer structurele analyse van de declaratiegegevens bij Vektis ten behoeve van fraudedetectie momenteel onderzocht. In de gezamenlijke analyse ontdekte fraudepatronen en verbanden zouden door alle verzekeraars zowel kunnen worden gebruikt voor daadwerkelijk onderzoek als ook in het aanscherpen van de controleregels in hun declaratiesysteem. Een bijkomend voordeel van het centraal organiseren van datamining van zorgdeclaratiegegevens is dat de kosten voor de aanschaf, implementatie en onderhoud van de benodigde software gedeeld kunnen worden.

2. Analysefunctie inrichten

Ongeacht de tools die voor geautomatiseerde fraudedetectie worden ingezet, blijkt in de praktijk dat een goed ingerichte analysefunctie in de organisatie van de fraudepreventie en –detectie van groot belang is. Een goede analist kan al met de juiste zoekvragen in de database van de organisatie en met standaard software naar patronen en verbanden zoeken ter detectie van potentiële fraude. Uit de bevindingen komt daarom als aanbeveling voor de zorgverzekeraars naar voren om binnen de fraudepreventie en –detectie in ieder geval de analysefunctie goed in te richten met bij voorkeur een full time analist.

3. EI-standaarden optimaliseren voor gebruik datamining

Voor de bruikbaarheid van dataminingsoftware voor het detecteren van fraude is het van groot belang dat de beschikbare data voldoende variëteit en kwaliteit heeft. In het algemeen is het voor het succesvol kunnen toepassen van datamining voor de detectie van fraude in zorgdeclaraties van belang dat het detailniveau van de zorgdeclaraties zo hoog mogelijk is en de systematiek niet (te vaak) wordt gewijzigd.

Het grote voordeel dat het Nederlandse zorgsysteem heeft, is dat de meeste zorgdeclaraties centraal en digitaal via VECOZO worden verwerkt op basis van vaste standaarden. Bij het wijzigen van de zorgsystematiek, en meer in detail de declaratie en EI-standaarden, zou rekening met het belang van fraudedetectie en -bestrijding gehouden moeten worden door waar mogelijk meer details over behandelingen uit te wisselen. Vanzelfsprekend zal wel een balans gezocht moeten worden met de eventuele administratieve last die door uitgebreide(re) EI standaarden bij de zorgverleners zou kunnen ontstaan.

4. Begrippen en registratie standaardiseren

Tijdens het onderzoek is, zowel uit de expertmeeting als uit de jaarlijks door ZN verzamelde gegevens, gebleken dat verschillen bestaan in de mate en wijze waarop (meta)gegevens over onregelmatigheden en fraude met zorgdeclaraties bij de zorgverzekeraars worden verzameld en vastgelegd. Hoewel binnen ZN gezamenlijk afspraken zijn gemaakt over de gehanteerde definities, lijken in de praktijk interpretatieverschillen te bestaan tussen de zorgverzekeraars Een goede vergelijking van en inzicht in de fraudebestrijding en resultaten daarvan, zowel tussen de zorgverzekeraars als met vergelijkbare omringende landen, is hierdoor veelal niet mogelijk. Het is daarom aan te bevelen het verzamelen en de wijze van vastleggen van (meta)gegevens over het fraudebestrijdingsproces (verder) te standaardiseren en transparant te maken.

5. Gebruiken fraudesignalen ter preventie

Onder het motto voorkomen is beter dan genezen, is het van groot belang dat aan de voorzijde van de zorgketen gebruik wordt gemaakt van de signalen die door de zorgverzekeraars worden afgegeven met betrekking op zwakheden in wet- en regelgeving en in processen. De fraudecoördinatoren van een aantal zorgverzekeraars hebben aangegeven dat zij aan de hand van concrete voorbeelden de aandacht te hebben gevestigd op fraudegevoelige zwakheden in wet- en regelgeving en processen, bijvoorbeeld inzake het PGB. In hun perceptie hebben deze signalen niet tot bijstelling van het beleid en de processen heeft geleid. Met name een rechtstreekse *feedback* van concrete signalen uit de praktijk van de fraudedetectie en –bestrijding bij de zorgverzekeraars zouden de fraudebestendigheid van wet- en regelgeving en processen kunnen verhogen.

6. Opvolgen aangiften en fraudesignalen

In 2011 zijn in een Convenant Aanpak Verzekeringsfraude tussen verzekeraars en OM en de opsporingsdiensten afspraken gemaakt over de route waarlangs aangiften en fraudesignalen vanuit de (zorg)verzekeraars bij de overheid terecht kunnen komen. Het is aan te bevelen dat de zorgverzekeraars consequent die route volgen, zodat het OM en de opsporingsdiensten een eenduidige input krijgen, maar ook opdat een eenduidige beeld kan ontstaan met betrekking tot de prioriteit die door OM, politie, FIOD, SIOD en Nza gegeven wordt aan zorgfraude. Op basis daarvan kunnen eventueel verdere afspraken over een gezamenlijk aanpak van zorgfraude gebaseerd worden. Een inventarisatie van de liggende fraudezaken en fraudeaanwijzingen, die in de perceptie van de zorgverzekeraars niet of onvoldoende zijn opgepakt door de relevante overheidsdiensten, zou eveneens kunnen bijdragen in een (beter) inzicht in de knelpunten in de samenwerking.

4. Slotopmerking

Dit rapport is opgesteld ten behoeve van het Ministerie van Volksgezondheid Welzijn en Sport en is afgestemd op het gebruik ten behoeve van de vermelde doelstelling. Dit rapport is dan ook niet bedoeld, en mogelijk ook niet geschikt, voor enig ander gebruik.

Grant Thornton Forensic & Investigation Services B.V.

Mark Hoekstra

Partner

Bijlage 1 - Organisaties waaruit de voor dit onderzoek geïnterviewden afkomstig zijn.

Achmea

Agis

CZ

De Friesland

DSW

ENO-Salland

FICO

Friss fraude en risicobeheersing

IBM Nederland

Krammer Software

Menzis

Ministerie van Volksgezondheid, Welzijn en Sport

Nederlandse Zorgautoriteit

Ordina

SAS Nederland B.V.

The Beagle Armada

UVIT (Univé / VGZ)

VEKTIS

Verbond van Verzekeraars

Zorg & Zekerheid

Zorgverzekeraars Nederland

Bijlage 2 - Bestudeerde documentatie

Conz N (2009) Insurers Leverage Data Mining and Predictive Analytics to Mitigate Increasingly Complex Fraud Schemes. In *Insurance & Technology,* December 2009.

Hasaart F (2011) *Incentives in the Diagnosis Treatment Combination payment system for specialist medical care. A study about behavioral responses of medical specialists and hospitals in the Netherlands.* Dissertatie. Maastricht: Universiteit Maastricht.

Kelley R (2009) *Where can $700 Billion in Waste be cut annually from the U.S. Healthcare System?* Whitepaper. Ann Arbour: Thomson Reuters.

Kobielus J (2010) *Predictive Analytics and Data Mining Solutions, Q1 2010*. Cambridge (USA): Forrester Research.

Mikut R and Reischl M (2011) Data mining tools. *WIREs Interdiciplinary Reviews: Data Mining and Knowledge discovery*, Vol. 1(5): 431-443.

NZa (2011) *Beleidsregel TH/BR-001 Toezichtkader zorgplicht zorgverzekeraars (Zvw) De reikwijdte van de zorgplicht in begrippen, verantwoordelijkheden en normen*. Utrecht: Nederlandse Zorgautoriteit.

Rexer Analytics (2011) *4th Annual Data Miner Survey – 2010 Survey Summary Report.* Winchester: Rexer Analytics.

Seifert J (2004) *Data Mining: An Overview. CRS Report for Congress.* Washington D.C.: Congressional Research Service.

Travaille P, Muller R, Thornton D and Van Hilligersberg J (2011) *Electronic Fraud Detection in the U.S. Medicaid Healthcare Program: Lessons Learned from other Industries.* AMCIS Proceedings – All Submissions. Paper 328. http://aisel.aisnet.org/amcis_2011_submissions/328.

Vektis (2011) *Notitie Informatiesystemen Vektis.* Zeist: Vektis.

Verbond van Verzekeraars en Zorgverzekeraars Nederland (2011) *Verzekeraars & Criminaliteit. Protocol betreffende Bewustwording, Preventie, Detectie en Afhandeling van Verzekeringsfraude en Criminaliteit.* Den Haag: Verbond van Verzekeraars.

VWS (2010) *Plan van aanpak 2010 2012 Verbetering fraudebestrijding in de zorg. Werkprogramma van de Regiegroep 'Verbetering fraudebestrijding in de zorg' onder leiding van het Ministerie van VWS*. Den Haag: Ministerie van Volksgezondheid, Welzijn en Sport.

ZN (2011) *Landelijke inventarisatie inzet, ervaringen en resultaten van het fraudebeheersingsbeleid 2010 – 2009*. Zeist: Zorgverzekeraars Nederland.