



Centraal Bureau
voor de Statistiek

Paper

Gezondheidsenquête 2014: een analyse van de methodebreuk

Mattijn Morren
Rob Willems
Mei 2015

De Gezondheidsenquête (GE) wordt vanaf 1981 elk jaar uitgevoerd. In 2014 vond er een herontwerp van de GE plaats. Het aantal modes werd beperkt tot twee en de vragenlijst werd ingekort. Hierbij werden o.a. de twee vragenlijstdelen van de GE tot dan toe samengevoegd. Ook werd het weegmodel uitgebreid. Dit rapport beschrijft een onderzoek naar de gevolgen van dit herontwerp op de cijfers van 2014. Hiertoe zijn methodebreukanalyses uitgevoerd op een selectie van 17 variabelen. Voor deze variabelen is nagegaan of het cijfer van 2014 afwijkt van het verwachte cijfer berekend op basis van de reeks t/m 2013. Dit is gedaan voor zowel het cijfer van de totale bevolking als voor de uitsplitsing naar geslacht en leeftijd (0–25 jaar, 25–65 jaar en 65 jaar of ouder). De resultaten tonen dat er voor een meerderheid van de variabelen sprake is van een methodebreuk. Bij het interpreteren van de GE cijfers van 2014, en met name ook bij een vergelijking met de cijfers van de jaren daarvoor, is het van belang hiermee rekening te houden.

Met dank aan Jan-Willem Bruggink, Harm Jan Boonstra en Bart Buelens.

Inhoud

1. Inleiding	4
1.1 Doelstelling	4
1.2 Opbouw van het rapport	5
2. Methode	5
2.1 Geselecteerde variabelen	5
2.2 Samenstelling van het analysebestand	6
2.3 Analyse van de methodebreuken	7
3. Het herontwerp van 2014	11
3.1 Wijziging van het onderzoeksdesign	11
3.2 Wijziging van de vragenlijst	13
3.3 Wijziging van vragen	14
3.4 Wijziging van het weegmodel	18
4. Resultaten	18
4.1 Voorwaarden statistisch model	19
4.2 Methodebreuk gehele groep	19
4.3 Methodebreuk naar geslacht	21
4.4 Methodebreuk naar leeftijd	22
4.5 Methodebreuk per variabele	23
5. Discussie	25
5.1 Algemene verklaringen	25
5.2 Verklaring per variabele	26
5.3 Beperkingen van het onderzoek	31
6. Conclusies	32
Literatuur	33
Bijlagen	35

1. Inleiding

De Gezondheidsenquête (GE) is een jaarlijks continu onderzoek dat wordt uitgevoerd vanaf 1981. Het onderzoek richt zich op de gezondheid, medische consumptie en leefstijl van de Nederlandse bevolking. De gegevens worden verzameld door middel van een vragenlijst. Tot 2010 bestond de GE uit twee delen, waarbij deel 1 via een persoonlijk interview (CAPI) en deel 2 via een schriftelijke vragenlijst (PAPI) werd afgenomen. Tussen 1981 en 2010 werd de vragenlijst af en toe aangepast of werd de leeftijdsgrens voor het stellen van vragen gewijzigd. De mode waarin de vragenlijst werd afgenomen is in deze periode echter hetzelfde gebleven. Tot 2010. In dat jaar werd een mixed-mode design ingevoerd. De vragenlijst bestond nog steeds uit twee delen, maar deelname kon naast de bestaande modes nu ook via een internetvragenlijst (CAWI) of een telefonisch interview (CATI). Dit mixed-mode design had echter enkele nadelen. Ten eerste was er hoge uitval na onderzoeksdeel 1, waardoor de respons voor onderzoeksdeel 2 daalde en de kans op selectiviteit toenam. Ten tweede was de logistiek ingewikkeld, omdat in deel 1 in drie modes kon worden gerespondeerd (CAWI, CATI of CAPI) en in deel 2 in twee (CAWI en PAPI).¹⁾

Mede om die reden vond in 2014 een herontwerp van de Gezondheidsenquête plaats. Bij dit herontwerp werden verschillende aspecten van de GE gewijzigd met als doel het vereenvoudigen van de gegevensverzameling en het verbeteren van de datakwaliteit. De twee onderzoeksdelen waaruit de GE tot dan toe bestond werden samengevoegd, zodat deelonderzoek 2 ophield te bestaan. Het aantal modes van onderzoeksdeel 1 (nu het enige deel) werd gereduceerd tot twee (CAWI en CAPI). Hierdoor werd niet alleen de logistiek van het benaderen van deelnemers eenvoudiger, maar ook de gegevensanalyse. Er hoefde niet langer rekening gehouden te worden met selectiviteit van uitval voor de tweede vragenlijst (Goris, 2013b). Ook de vragenlijst zelf werd fors gewijzigd. Vragen en vragenblokken verdwenen en werden gewijzigd of verplaatst, en nieuwe vragen werden toegevoegd. Dit had te maken met een samenwerkingsverband met de Leefstijlmonitor en verplichtingen aan het Europees statistisch bureau Eurostat ten aanzien van de Europese Gezondheidsenquête EHIS, de European Health Interview Survey. Al met al werd de vragenlijst korter, wat noodzakelijk was, omdat de invulduur van het enige resterende vragenlijstdeel niet langer dan 35 minuten mocht bedragen. Tot slot werd het weegmodel van 2014 uitgebreid ten opzichte van het weegmodel zoals vóór 2014 voor deel 1 werd toegepast. De variabelen inkomen en vermogen werden eraan toegevoegd om tot betere cijfers te komen.

1.1 Doelstelling

Het doel van dit onderzoek is om in kaart te brengen wat de gevolgen zijn van het GE herontwerp 2014 voor de cijfers van 2014. Indien blijkt dat er bij een variabele sprake is van een methodebreuk, wordt vervolgens gekeken wat de omvang is van deze breuk en welke factoren een rol kunnen spelen bij het ontstaan en/of het verklaren van de breuk. We hanteren hierbij de volgende vraagstellingen:

1. In hoeverre heeft de methodebreuk ten gevolge van het herontwerp van de Gezondheidsenquête in 2014 geleid tot een verandering van de cijfers van 2014 ten opzichte van de periode daarvoor?

¹⁾ CAPI = Computer Assisted Personal Interviewing (face-to-face interview); PAPI = Paper Assisted Personal Interviewing (papieren vragenlijst); CAWI = Computer Assisted Web Interviewing (internet vragenlijst); CATI = Computer Assisted Telephone Interviewing (telefonisch interview).

2. In hoeverre verschillen deze veranderingen tussen mannen en vrouwen en naar gelang leeftijd?
3. In hoeverre verklaart het nieuwe weegmodel, met inkomen en vermogen als additionele weegvariabelen, eventuele veranderingen in de cijfers van 2014?
4. Welke andere mogelijke verklaringen zijn er voor deze verschillen?

Om deze vraagstellingen te onderzoeken, zijn er 17 variabelen geselecteerd voor de methodebreukanalyse. Dit zijn variabelen uit elk van de drie onderzoeksgebieden van de GE, gezondheid, zorggebruik en leefstijl.

1.2 Opbouw van het rapport

Dit rapport is als volgt opgebouwd. In hoofdstuk 2 worden de variabelen beschreven die zijn geselecteerd voor de methodebreukanalyses. Tevens wordt beschreven hoe de analyses zijn uitgevoerd. Vervolgens worden in hoofdstuk 3 de wijzigingen van GE 2014 ten opzichte van 2013 op een rij gezet. Het betreft hier wijzigingen in het onderzoeksdesign, in de structuur van de vragenlijst, in de vragen zelf en in het weegmodel. Daarna worden in hoofdstuk 4 de resultaten van de methodebreukanalyses gepresenteerd. Er wordt daarbij een onderscheid gemaakt tussen algemene bevindingen en bevindingen die specifiek betrekking hebben op een variabele. Op basis van die resultaten worden dan in hoofdstuk 5 mogelijke verklaringen gegeven voor de resultaten uit hoofdstuk 4. Hoofdstuk 6 sluit af met de conclusies.

2. Methode

Het idee achter deze methodebreukanalyse is dat er een vergelijking wordt gemaakt tussen een schatting van het cijfer van 2014 op basis van het verleden enerzijds en het reële cijfer van 2014 anderzijds. Het verschil tussen die twee geeft een indicatie van de mate waarin het herontwerp invloed heeft op de cijfers van 2014 en daarmee of er sprake is van een methodebreuk.

In dit hoofdstuk wordt besproken hoe te werk is gegaan bij het analyseren van de methodebreuk. In paragraaf 2.1 wordt de selectie van variabelen voor de analyses beschreven en in paragraaf 2.2 de wijze waarop de gegevens zijn verkregen die de input vormen voor de analyses. In paragraaf 2.3 wordt uitgelegd hoe de analyses zijn uitgevoerd.

2.1 Geselecteerde variabelen

Het is niet mogelijk om voor alle variabelen waarover op StatLine wordt gepubliceerd en voor alle uitsplitsingen daarvan een methodebreukanalyse uit te voeren. Dit zijn er simpelweg te veel. Daarom is een aantal kernvariabelen geselecteerd, waarbij er wel voor gezorgd is dat er variabelen uit elk van de drie onderzoeksgebieden van de GE, gezondheid, zorggebruik en leefstijl, vertegenwoordigd zijn. Daarnaast zijn er variabelen geselecteerd uit zowel het eerste als het tweede onderzoeksdeel van de voorgaande jaren en zijn enkele variabelen die in de analyses van 2010 gevoelig bleken voor een methodebreuk in de selectie opgenomen (zie Wong, Roels & Van den Brakel, 2011). Voor de volgende variabelen zijn analyses uitgevoerd:

Gezondheid:

- Ervaren gezondheid
- Diabetes
- Verkoudheid
- Psychische ongezondheid
- Minstens 1 beperking (OESO-indicator)
- Beperking in bewegen (OESO-indicator)
- Lichamelijke beperking (ADL-indicator)

Zorggebruik:

- Contact met de huisarts
- Aantal contacten met de huisarts
- Contact met de specialist
- Ziekenhuisopname
- Zelfmedicatie

Leefstijl:

- Mammografie
- Roken
- Zware drinker (onder drinkers)
- Alcoholconsumptie drinkers (aantal glazen per dag)
- Ernstig overgewicht

Voor deze variabelen worden naast de totalen de cijfers opgesplitst naar geslacht en naar leeftijd (in 3 groepen: tot 25 jaar, 25 tot 65 jaar, 65 jaar en ouder). Merk op dat de variabele mammografie alleen betrekking heeft op vrouwen van 30 jaar of ouder en de variabele lichamelijke beperking op personen van 55 jaar of ouder. Een uitgebreide beschrijving van deze variabelen is te vinden in Bijlage 1.

2.2 Samenstelling van het analysebestand

Voor een methodebreukanalyse is het wenselijk om een zo lang mogelijke reeks van cijfers beschikbaar te hebben. Dit heeft een positief effect op de kwaliteit van de schatting van de breuk. Tegelijkertijd wordt echter het effect van het verleden op het heden kleiner naarmate dit verleden verder weg ligt. De Gezondheidsenquête gaat terug tot 1981, maar door wijzigingen in de enquête zijn er niet voor alle geselecteerde variabelen cijfers vanaf dat jaar beschikbaar. In Bijlage 2 staat voor elke variabele vanaf welk jaar het cijfer in de methodebreukanalyses is opgenomen en hoe en waarom dit eventueel afwijkt van de jaren waarover op StatLine is gepubliceerd.

De input van de methodebreukanalyses bestaat uit de cijfers en steekproefomvang (het aantal mensen dat gerespondeerd heeft) van de geselecteerde variabelen voor de jaren waarvoor deze beschikbaar en bruikbaar zijn. Voor de meeste variabelen zijn de cijferreeksen op StatLine beschikbaar voor de hele populatie en voor mannen en vrouwen apart. Toch is ervoor gekozen de reeksen opnieuw samen te stellen op basis van de bronbestanden. Dit heeft de volgende redenen:

1. Voor de drie leeftijdsgroepen (0–25, 25–65 en 65 jaar of ouder) waarvoor breukanalyses worden uitgevoerd zijn op StatLine geen gegevens beschikbaar.
2. In de loop van de jaren zijn er voor verschillende variabelen wijzigingen doorgevoerd. Dit betrof bijvoorbeeld wijzigingen in de definitie, en daarmee de afleiding, van variabelen

- (bijv. zware drinker) of in de wijze van presenteren van de indicator (bijv. psychische ongezondheid). Zie Bijlage 2 voor een overzicht van de verschillen tussen de variabelen zoals ze in de analyse zijn opgenomen en hoe deze op StatLine zijn gepubliceerd.
3. Op StatLine zijn niet voor alle individuele jaren de exacte responsaantallen gepresenteerd. Deze aantallen worden gebruikt voor het uitvoeren van de methodebreukanalyses.
 4. Op StatLine worden de cijfers op één cijfer achter de komma gepresenteerd. In de herberekening zijn meer cijfers achter de komma beschikbaar, hetgeen de precisie van de analyses zal verhogen.

2.3 Analyse van de methodebreuken

In deze paragraaf wordt uiteengezet hoe de methodebreukanalyses zijn uitgevoerd. Eerst wordt een algemene beschrijving van de analyses gegeven, die vervolgens verder wordt gepreciseerd. De paragraaf sluit af met een toelichting van de procedure waarmee de analyses zijn uitgevoerd.

Opzet analyse

De cijfers van 2014 zijn alleen volgens de onderzoeksmethode van na het herontwerp verzameld en niet ook volgens de methode van daarvoor. Er is dus niet dubbel gedraaid. Hierdoor kan er geen directe vergelijking worden gemaakt tussen deze beide methoden. Wel kunnen de verschillen tussen de methoden indirect gekwantificeerd worden. Dit kan door voor de geselecteerde GE-variabelen de verwachte waarden (geschat op basis van de enquêtes vóór 2014) en de werkelijke waarden van 2014 met elkaar te vergelijken. Een gangbare manier om dit soort verschillen te kwantificeren is via een interventieanalyse aan de hand van een structureel tijdreeksmodel. Deze aanpak wordt beschreven in Van den Brakel en Roels (2010) en Van den Brakel, Smith en Compton (2008) en is een directe toepassing van de interventieanalyse zoals voorgesteld door Harvey en Durbin (1986). Een tijdreeks is een serie waarnemingen of datapunten van een variabele gedurende een bepaalde tijd. Bekende voorbeelden zijn beurskoersen, temperatuurmetingen, werkloosheidscijfers, etc. Het analyseren van zulke reeksen om hieruit betekenisvolle informatie te kunnen afleiden, wordt tijdreeksanalyse genoemd. In tijdreeksanalyse worden daartoe wiskundige modellen geformuleerd, die deze reeksen zo goed mogelijk beschrijven. In een structureel tijdreeksmodel wordt de reeks beschouwd als zijnde opgebouwd uit één of meer van de volgende componenten:

1. Een trendcomponent: deze component beschrijft de werkelijke trend van een tijdreeks, ontdaan van andere effecten (onderstaande componenten 2 t/m 5) die een helder beeld op deze trend in de weg staan. Deze component maakt deel uit van elk structureel tijdreeksmodel.
2. Een seizoencomponent: deze component modelleert herhalende, kortdurende fluctuaties met een vaste en bekende periode, zoals fluctuaties in de seizoenen.
3. Een cyclische component: deze component modelleert fluctuaties die gedurende een meer langdurig tijdsbestek plaatsvinden en niet een vaststaande periode kennen, zoals de economische conjunctuur (met een periodiciteit van ± 10 jaar). Er kan sprake zijn van meerdere cyclische componenten. Vaak is de cyclische component in de trendcomponent opgenomen.
4. Een regressiecomponent: deze component wordt opgenomen als er één of meer additionele verklarende variabelen zijn.
5. Een storingsterm (ook wel irreguliere term genoemd): deze component bevat de tijdreeks ontdaan van de vorige componenten en representeert de steekproefruis.

Iedere component wordt beschreven door een stochastisch model: dit is een wiskundig model dat ruimte laat voor onzekerheden en ontwikkelingen in de tijd. Voor meer informatie over structurele tijdreeksmodellen wordt verwezen naar Durbin en Koopman (2001).

Soms wordt de continuïteit van het fenomeen dat door een tijdreeks wordt beschreven plotseling blijvend en voor langere tijd verstoord. Dit geldt ook voor het invoeren van een gewijzigde enquête en/of onderzoeksopzet en het effect daarvan op de cijfers. Om het effect van de nieuwe vragenlijst te kwantificeren, kan een tijdreeksmodel worden opgesteld dat met deze verandering rekening houdt. Bij structurele tijdreeksen kan dat middels een zogenaamde interventieanalyse.

Bij een interventieanalyse wordt in de regressiecomponent een variabele toegevoegd die de verstoring van de continuïteit markeert, een zogenaamde interventievariabele. Deze variabele is een soort stapfunctie, die op het tijdstip van de wijziging overgaat van de waarde 0 naar 1. In het onderhavige herontwerp van de GE kan de bijbehorende regressiecoëfficiënt worden geïnterpreteerd als de grootte van de methodebreuk door de overgang van het oude naar het nieuwe ontwerp. Dit kunnen we stellen onder de aanname dat de werkelijke tijdreeks goed wordt benaderd door de andere componenten in het tijdreeksmodel. Met andere woorden, er wordt verondersteld dat op het tijdstip dat wordt overgegaan op het nieuwe ontwerp, de werkelijke ontwikkeling niet afwijkt van het veronderstelde tijdreeksmodel.

De invoering van het mixed-mode design voor de GE in 2010 en alle andere grote en kleine wijzigingen in de vragen en in de vragenlijst kunnen worden beschouwd als methodebreuken. Er is echter besloten om alleen de breuk van 2010 mee te nemen in het tijdreeksmodel, omdat deze recent plaatsvond en ingrijpend was (zie voor een beschrijving van deze breuk Wong, Roels en Van den Brakel, 2011b). Eventuele oudere breuken zouden nauwelijks nog effect hebben op de breukanalyse voor 2014. De breuk van 2014 is derhalve bepaald als ware de tijdreeks gecorrigeerd voor de periode voorafgaand aan de breuk van 2010. Dat wil zeggen dat de periode vóór 2010 op het niveau is gebracht van de periode ná 2010. Er is voor gekozen om de breuk van 2010 altijd in het model op te nemen, ongeacht of deze significant is. De reden hiervoor is dat hierdoor de interpretatie van de uitkomsten voor alle variabelen en opsplitsingen daarvan hetzelfde is en daarmee eenvoudiger.

Statistisch tijdreeksmodel

In het algemene structurele tijdreeksmodel komen enkele componenten voor die bij de GE niet relevant zijn. Zo is er geen seizoencomponent, omdat de reeksen jaarcijfers zijn. Andere cyclische componenten worden eveneens niet apart in het model opgenomen, omdat de modevaluatie daartoe geen aanleiding geeft. Duidelijke effecten van bijvoorbeeld de economische conjunctuur zijn niet zichtbaar en voor zover ze wel aanwezig zijn, worden ze feitelijk al door het tijdsafhankelijke trendmodel beschreven.

De volgende componenten blijven dan over: de trendcomponent, de regressiecomponent en de storingsterm. Dit betekent dat voor ieder van de onderzochte tijdreeksen het volgende structurele tijdreeksmodel verondersteld wordt:

$$\hat{y}_{t,i} = L_{t,i} + \beta_i \delta_{t,i} + \beta_i^* \delta_{t,i}^* + \varepsilon_{t,i} \quad (1)$$

Hierbij is $L_{t,i}$ de trend op tijdstip t voor reeks i ; en $\delta_{t,i}$ is een interventievariabele voor reeks i op tijdstip t , die aangeeft of op het betreffende tijdstip sprake is van het oude of van het nieuwe ontwerp, ofwel:

$$\delta_{t,i} = \begin{cases} 1 & (t = 2014) \\ 0 & (t < 2014) \end{cases} \quad (2)$$

En $\delta_{t,i}^*$ is de interventie variabele die het effect van het herontwerp van 2010 weergeeft:

$$\delta_{t,i}^* = \begin{cases} 0 & (t < 2010) \\ 1 & (t \geq 2010) \end{cases} \quad (3)$$

Verder zijn β_i en β_i^* de regressiecoëfficiënten die respectievelijk behoren bij $\delta_{t,i}$ en $\delta_{t,i}^*$. Ten slotte is $\varepsilon_{t,i}$ de storingsterm van reeks i . We nemen aan dat de storingstermen onderling onafhankelijk zijn en normaal zijn verdeeld volgens

$$\varepsilon_{t,i} \equiv N\left(0, \frac{\sigma_{\varepsilon_i}^2}{n_t}\right) \quad (4)$$

met n_t de steekproefomvang in jaar t . De variantie in de meetvergelijking wordt dus verondersteld *alleen* van de grootte van de steekproef af te hangen en neemt natuurlijk toe als de steekproef kleiner wordt. De aanname is hier dat het verband omgekeerd evenredig is. De stochastische trend wordt gemodelleerd met

$$\begin{aligned} L_{t,i} &= L_{t-1,i} + R_{t-1,i} \\ R_{t,i} &= R_{t-1,i} + \eta_{t,i} \end{aligned} \quad (5)$$

Hierbij is $R_{t,i}$ de stochastische hellingsparameter en $\eta_{t,i}$ zijn de storingstermen waarvan verondersteld wordt dat ze onderling onafhankelijk zijn en normaal verdeeld volgens

$$\eta_{t,i} \equiv N(0, \sigma_{\eta_i}^2) \quad (6)$$

De methodebreuk van 2014, gegeven de methodebreuk van 2010, is dan β_i . Met andere woorden, β_i geeft het effect weer van de veranderingen in 2014 als ware de methodebreuk gecorrigeerd voor de breuk van 2010. Merk tot slot op dat is verondersteld dat de onderzochte tijdreeksen van de verschillende variabelen, en daarmee ook de berekende trendbreuken, van elkaar onafhankelijk zijn.

Uitvoering analyse

Het model in de vorige paragraaf wordt geïmplementeerd in de programmeertaal *Ox* in het pakket *OxMetrics*, daarbij gebruik makend van de subroutines uit *Ssfpack 3.0*. Zie Doornik (2009) en Koopman, Shepard en Doornik (2008) voor verdere details.

De uiteindelijke bedoeling van deze implementatie is om de grootte van β_i te schatten. Als eerste worden daartoe via de methode van *Maximum Likelihood* de onbekende modelparameters geschat. Dit zijn de σ_{ε_i} en σ_{η_i} van (4) respectievelijk (6), de zogenaamde hyperparameters. Daarna worden met het Kalmanfilter de waarden β_i van en β_i^* geschat. Alle waarnemingen van de gehele tijdreeks (1) worden hiervoor gebruikt. De schattingen worden dan *gesmoothde* schattingen genoemd.

Als het gekozen model de ontwikkeling van de doelvariabelen goed benadert, dan kunnen de berekende regressiecoëfficiënten voor de interventievariabelen (dus β_i en β_i^*) geïnterpreteerd worden als de omvang van de methodebreuk. In dit geval is het dus de $\hat{\beta}_i$ (de *gesmoothde* schatting van β_i) die de methodebreuk van 2014 kwantificeert.

Om het effect van de methodebreuk teniet te doen zou β_i dus afgetrokken moeten worden van de waarneming van 2014. De zo berekende waarde is dan de voor de trendbreuk gecorrigeerde waarde, of kortweg de gecorrigeerde waarde.

Om te bepalen of het model goed past bij de waarnemingen, zijn er een aantal standaard hulpmiddelen (zie m.n. Durbin en Koopman, 2001).

- Allereerst kunnen procesgegevens gegenereerd door de gebruikte software worden gebruikt. De onbekende modelparameters worden via een iteratief proces geschat. Deze schatting kan al dan niet leiden tot convergentie. Sterke convergentie is een eerste indicatie dat het gekozen model goed past.
- Normaliteit van de innovaties. Innovatie is de gangbare term voor het verschil tussen de voorspelde waarde en de werkelijke waarde in een gegeven jaar of tijdstip. Nauwkeuriger gezegd: een innovatie is het verschil tussen de voorspelde waarde voor tijdstip $t+1$ (uitgaande van alle waarnemingen tot en met tijdstip t) en de waargenomen waarde (op tijdstip $t+1$). Deze innovaties worden voor alle beschikbare tijdstippen in de analyse, behalve het eerste, berekend en mogen niet al teveel afwijken van een normale verdeling. Als bijvoorbeeld de innovaties veel vaker positief zijn dan negatief, dan is dat een teken dat het model niet goed past. Voor het bepalen van normaliteit bestaat een standaard statistische test.
- Variantie van de innovaties. De innovaties behoren gedurende de periode waarop de tijdreeks betrekking heeft geen al te grote veranderingen in variantie te vertonen. Om dit te bepalen worden de varianties van de innovaties uit het begin van de tijdreeks (het eerste derde deel) vergeleken met die uit het einde (het laatste derde deel). Als de grootte van de varianties substantieel verandert, is er sprake van heteroscedasticiteit. Er bestaat een standaard statistische test om te bepalen of hiervan sprake is.
- Seriële correlatie van de innovaties. De innovaties op tijdstip t en op tijdstip $t+\Delta t$ behoren niet gecorreleerd te zijn voor iedere Δt . Voor het bepalen van seriële correlatie bestaat een standaard statistische test.
- Visuele inspectie van de innovaties. In sommige gevallen kan formeel worden geconcludeerd dat er sprake is van heteroscedasticiteit, maar wordt na visuele inspectie toch besloten dat er geen sprake is van een duidelijke vergroting of verkleining van de varianties.

Een methodebreuk die op deze manier is berekend, is niet per definitie ook significant. De verandering kan immers ook door toeval tot stand zijn gekomen. Om significantie aan te tonen, wordt gebruik gemaakt van de varianties van de breuken. Voorts wordt onder de aanname van normaliteit de kans berekend dat een breuk van een bepaalde omvang door toeval tot stand is gekomen. Als deze kans kleiner is dan een vastgestelde grenswaarde, wordt besloten dat er sprake is van een significante breuk. Die grenswaarde wordt vaak vastgesteld op 5%, ofwel er is p waarde van maximaal 5%.

De methodebreuk (d.w.z. $\hat{\beta}_i$) kan vervolgens worden opgesplitst in twee delen:

1. Eén deel dat verklaard wordt door veranderingen in design, benaderingsstrategie en vragenlijst: hierna designbreuk.
2. Eén deel dat kan worden verklaard door de weging: hierna wegingsbreuk.

In praktijk is het echter niet mogelijk om deze twee delen volledig van elkaar te scheiden. Dit komt doordat het oude weegmodel is opgesteld voor de modes en de benaderingsstrategie van het onderzoek na het herontwerp van 2010. Hierdoor kan het zo zijn dat het oude weegmodel selectieve respons in de periode 2010 t/m 2013 goed ondervangt, maar selectieve respons van 2014 niet. Er kan daarom niet worden uitgesloten dat een deel van een wegingsbreuk feitelijk het gevolg is van het veranderen van het design, of omgekeerd. Het is wel mogelijk om een inschatting te maken van de grootte van de designbreuk en van de wegingsbreuk. Er moet dan worden aangenomen dat het weegmodel van 2013 een goede

benadering is van het weegmodel zoals dat in 2014 zou zijn gebruikt als op de oude manier was gewogen. Dit is methodologisch het meest waarschijnlijk, hoewel voor dit geval geen kwantitatieve onderbouwing berekend is.

Het is mogelijk om voor de totaalbreuk een exacte p waarde te berekenen op basis van de uitkomsten van het structureel tijdreeksmodel. Hiermee is dan ook exact aan te geven wanneer de breuk significant is. Maar omdat de designbreuk en de wegingsbreuk niet volkomen los van elkaar mogen worden beschouwd, kan voor deze breuken geen exacte p waarde berekend worden. En dus is ook niet exact aan te geven wanneer zo'n breuk significant zou zijn. Om toch enig houvast te bieden bij de interpretatie van de breuken, is er voor gekozen om waarden te markeren die significant zouden zijn indien beide delen van de breuk onafhankelijk zouden zijn.

3. Het herontwerp van 2014

In de inleiding werd al kort beschreven op welke punten de Gezondheidsenquête in 2014 als gevolg van het herontwerp is gewijzigd ten opzichte van 2013 en daarvoor. In dit hoofdstuk worden wijzigingen in vier aspecten van het onderzoek beschreven. Op het hoogste niveau zijn er wijzigingen die alle variabelen betreffen. Dit zijn wijzigingen in het onderzoeksdesign, zoals de mode en de waarneemstrategie (3.1). Daarnaast kan de vragenlijst zelf worden aangepast, waardoor de positie of de context van vragen verandert doordat vragen of vragenblokken worden verplaatst of aangepast. Dit soort wijzigingen, waarbij de vraag zelf gelijk blijft, worden beschreven in paragraaf 3.2. De vraag zelf of haar antwoorden kunnen uiteraard ook worden aangepast. Om dit soort wijzigingen in kaart te brengen worden in paragraaf 3.3 de vraagstellingen van 2014 vergeleken met die van t/m 2013.

Wijzigingen in het weegmodel betreffen eveneens alle variabelen in een onderzoek, maar dit beïnvloedt niet de wijze van gegevensverzameling. Wanneer het weegmodel wordt aangepast, is dit enkel van invloed op de manier waarop de gegevens worden geanalyseerd, niet op de gegevens zelf. Om die reden wordt de uitbreiding van het weegmodel onderscheiden van de andere wijzigingen in GE en in een aparte paragraaf beschreven (3.4).

3.1 Wijziging van het onderzoeksdesign

De veranderingen in het onderzoeksdesign van GE zijn eerder al kort beschreven. De belangrijkste wijziging betreft het samenvoegen van de twee onderzoeksdelen, met als gevolg dat GE vanaf 2014 nog maar uit één vragenlijst bestaat (zie ook 3.2). Daarnaast is het aantal modes waarin de vragenlijst kon worden ingevuld teruggebracht van vier (CAPI, CATI, CAWI en PAPI) tot twee (CAWI en CAPI). Door deze structurele veranderingen diende ook de waarneemstrategie, de wijze waarop potentiële deelnemers aan het onderzoek worden benaderd en gerappelleerd, te worden aangepast. In deze paragraaf wordt eerst het onderzoeksdesign van GE vóór 2014 beschreven, gevolgd door een beschrijving van het design in 2014.

De Gezondheidsenquête vóór 2014

Tot en met statistiekjaar 2013 was de GE opgebouwd uit twee onderzoeksdelen, waarbij respondenten van 12 jaar of ouder van onderzoeksdeel 1 werden herbenaderd voor

deelname aan onderzoeksdeel 2 (Goris, 2013ab). Vanaf 2010 was daarnaast sprake van een zogenaamd mixed-modes design, waarbij onderzoeksdeel 1 niet langer uitsluitend via persoonlijke interviews (CAPI), maar ook via telefonische interviews (CATI) en een internetvragenlijst (CAWI) werd waargenomen. Voor onderzoeksdeel 2 werd CAWI als extra mode aan PAPI toegevoegd. In de periode van 2010 t/m 2013 werden mensen als volgt benaderd om aan de Gezondheidsenquête mee te doen.

Mensen die in de steekproef zitten ontvangen een aanschrijfbrief met het verzoek om de vragenlijst van *onderzoeksdeel 1* via internet (CAWI) in te vullen. Wanneer men naar aanleiding hiervan niet respondeert, wordt na 1 week en eventueel na 2 weken een rappelbrief verstuurd. Mensen die er in deze periode geen blijk van geven deelname aan het onderzoek te willen weigeren, maar nog niet hebben gerespondeerd, worden vanaf een week na de tweede rappelbrief herbenaderd. Als het telefoonnummer bekend is, gebeurt dit via een telefonisch interview (CATI). Er worden dan 3 contactpogingen gedaan verspreid over de waarneemperiode. Elke contactpoging omvat 3 belpogingen die zijn verspreid over de dag: één in de ochtend, één in de middag en één in de avond. Als er geen telefoonnummer bekend is, worden mensen herbenaderd voor een persoonlijk interview bij hen thuis (CAPI). De interviewer doet voor elk adres maximaal 6 bezoekpogingen en eventueel een 7e als dit een concrete afspraak voor het interview betreft. Wanneer er geen contact kan worden gelegd of afspraak kan worden gepland, wordt deze persoon beschouwd als non-respondent. Personen van 12 jaar of ouder die in de vragenlijst van onderzoeksdeel 1 aangeven dat zij aan *onderzoeksdeel 2* willen deelnemen, worden hiervoor opnieuw benaderd. De wijze waarop dit gebeurt is gerelateerd aan de mode waarin hij of zij heeft gerespondeerd:

- CAWI-respondenten ontvangen uiterlijk 2 weken na responderen een brief met de inloggegevens voor de tweede vragenlijst. Eén en twee weken daarna wordt indien van toepassing een rappelbrief verstuurd. Bij het tweede rappel wordt een schriftelijke vragenlijst gevoegd.
- CATI-respondenten kunnen kiezen of zij de tweede vragenlijst via internet of schriftelijk willen invullen. Hiertoe krijgen zij uiterlijk 4 weken na het telefonische interview een gebruikersnaam en inlogcode of een schriftelijke vragenlijst toegestuurd. Na twee en vier weken wordt vervolgens een rappelbrief verstuurd met inlogcode dan wel een nieuwe schriftelijke vragenlijst. Respondenten die voor de internetvragenlijst hebben gekozen ontvangen bij de tweede rappelbrief tevens een schriftelijke vragenlijst. De optie om de tweede vragenlijst via internet in te vullen is in 2013 komen te vervallen, omdat via deze weg weinig respons werd verkregen.
- CAPI-respondenten krijgen na het interview nog tijdens het bezoek van de interviewer de tweede vragenlijst. De respondent vult deze vragenlijst bij voorkeur in terwijl de interviewer nog aanwezig is, zodat deze laatste de vragenlijst kan meenemen. Als dit niet mogelijk is, laat de interviewer de schriftelijke vragenlijst achter en wordt er net als bij CATI respondenten 2 keer schriftelijk gerappelleerd.

De belangrijkste reden om deze benaderingsstrategie in 2010 in te voeren was om het verzamelen van de gegevens goedkoper te maken ten opzichte van de waarneemmethodiek die tot en met 2009 werd gebruikt, waarbij tegelijkertijd de kwaliteit van de gegevens voldoende bleef.

De Gezondheidsenquête in 2014

Hoewel er in 2010 mixed-modes waarneming in de GE kon worden ingevoerd, had dit ook een aantal nadelen. Ten eerste daalde de respons op het tweede onderzoeksdeel sterk, waardoor een grotere steekproef nodig was en het risico op selectiviteit van de respons toenam. Ten tweede vereiste de benaderingsstrategie dat er met verschillende versies van

de vragenlijsten en van aanschrijf- en rappelbrieven moest worden gewerkt. Dit maakte de logistiek foutgevoelig. Ten derde noopte de uitval in het tweede onderzoeksdeel ten opzichte van het eerste tot het gebruik van een uitgebreider weegmodel met als gevolg dat de analyses ingewikkelder en foutgevoeliger werden.

Vanwege deze nadelen is besloten om vragenlijsten van de twee onderzoeksdelen te integreren tot één vragenlijst en om de benaderingsstrategie te vereenvoudigen tot 2 modes. Net als in de periode 2010 t/m 2013 wordt aan iedereen in de steekproef in 2014 een aanschrijfbrief verstuurd met gebruikersnaam en inlogcode om GE via een CAWI-vragenlijst in te vullen. Wanneer naar aanleiding van deze brief en de rappels 1 en 2 weken daarna niet wordt gerespondeerd, wordt men thuis benaderd door een interviewer. De periode waarin iemand wordt benaderd voor CAPI start uiterlijk twee maanden na het versturen van de aanschrijfbrief en duurt een maand. Het aantal bezoekpogingen en de wijze waarop mensen worden benaderd is niet veranderd ten opzichte van voorgaande jaren (Goris, 2013b).

3.2 Wijziging van de vragenlijst

Er zijn verschillende redenen waardoor de positie van een vraag in een vragenlijst kan veranderen. De 'positie' kan dan op twee manieren worden opgevat. Letterlijk, als de plaats in de vragenlijst waar een vraag zich structureel bevindt. Maar ook figuurlijk, als de context waarin een vraag wordt gesteld, wat er voorafgaat aan het stellen van de vraag. In beide interpretaties kan de positie van een vraag invloed hebben op het antwoord dat erop wordt gegeven.

De structurele positie van een vraag in de vragenlijst verandert wanneer een vraag bijvoorbeeld van het einde naar het begin van een vragenlijst wordt verplaatst, of andersom. In het geval van de GE was er een duidelijke structurele aanpassing, omdat de twee vragenlijstdelen zijn samengevoegd. Hierdoor, maar ook omdat vragen zijn toegevoegd en verwijderd, veranderde de plek van vragen in de vragenlijst. Gedurende het invullen van een vragenlijst treden er effecten op zoals vermoeidheid en een verminderde invulmotivatie. Wanneer de structurele positie van een vraag in een vragenlijst verandert, verandert ook de mate waarin deze effecten betrekking hebben op die vraag.

De relatieve positie van een vraag verandert automatisch doordat haar structurele positie verandert. Het antwoord van een respondent op een bepaalde vraag wordt namelijk bewust, maar meestal onbewust, beïnvloed door voorafgaande vragen. Dit kan de uitkomsten van een onderzoek beïnvloeden, sterker naarmate die voorafgaande vragen dichterbij de vraag in kwestie staan. Helaas is het vaak onduidelijk hoe vragen elkaar beïnvloeden en is het lastig en op zijn minst uitermate tijdrovend om dit voor alle vragen in een vragenlijst te doen. In specifieke gevallen is het echter wel relatief duidelijk dat een gewijzigde inbedding van een vraag van invloed zal zijn op de beantwoording ervan. Dit is het geval wanneer er een routevraag vooraf gaat, die bepaalt of de vraag in kwestie wordt gesteld. Wanneer deze routevraag wordt aangepast, verandert mogelijk ook de groep respondenten die de vraag in kwestie beantwoordt en de wijze waarop.

Het herontwerp van de Gezondheidsenquête in 2014 omvat een aantal wijzigingen in de vragenlijst die tot een verandering in de context van vragen hebben geleid. Niet alleen is de GE vragenlijst van twee delen naar één deel teruggebracht, ook de vragenblokken zijn ingekort, gewijzigd en binnen de vragenlijst verplaatst. Daarnaast is de routevraag aan het begin van het blok Alcoholconsumptie, op basis waarvan wordt bepaald of iemand drinker is of niet, gewijzigd. Een totaaloverzicht van de verschillende blokken waaruit de GE vragenlijsten van 2013 en 2014 zijn opgebouwd staat in Bijlage 3.

3.2.1 Positie van geselecteerde variabelen in de vragenlijst van 2013 en 2014

Variabele	Vragenlijstdeel (t/m 2013)	Bloknummer 2013 ¹⁾	Bloknummer 2014 ¹⁾	Aantal blokken verschil
Ervaren gezondheid	1	11	10	-1
Diabetes				
tot 12 jaar	1	27	17	-10
vanaf 12 jaar	2	4		13
Verkoudheid	2	7	19	12
Psychische ongezondheid	2	2	18	16
Minstens 1 beperking OESO	2	1	14	13
Beperking in bewegen OESO	2	1	14	13
Lichamelijke beperking ADL	1	25	12	-13
Contact met de huisarts	1	14	22	8
Aantal contacten met de huisarts	1	14	22	8
Contact met de specialist	1	17	24	7
Ziekenhuisopname	1	16	23	7
Zelfmedicatie	1	18	25	7
Mammografie	2	9	30	21
Roken	1	31	36	5
Zware drinker	2	5	37	32
Alcoholconsumptie	2	5	37	32
Ernstig overgewicht	1	29	32	3

¹⁾ Het rangnummer van het blok van het betreffende vragenlijstdeel of van de betreffende vragenlijst.

Tabel 3.2.1 is een samenvatting van Bijlage 3 en laat zien in welke blokken de vragen die bij de geselecteerde variabelen horen zich in de vragenlijst van 2013 en 2014 bevinden. Vervolgens is hieraan het rangnummer van het betreffende blok gekoppeld en het verschil tussen 2013 en 2014 berekend. De laatste kolom van de tabel laat zien dat de positie in de vragenlijst van sommige vragen weinig is veranderd (bijv. ervaren gezondheid en ernstig overgewicht), maar dat andere veel eerder of later in de vragenlijst worden gesteld (bijv. mammografie en zware drinker/alcoholconsumptie).

3.3 Wijziging van vragen

Door het samenvoegen van de twee vragenlijsten werd de vragenlijst langer. Tevens moesten er in 2014 gegevens voor de European Health Interview Survey (EHIS) worden verzameld en is er een samenwerkingsverband rond de Leefstijlmonitor (LSM). Voor de LSM is er een kernset van leefstijlvragen vastgesteld die onderdeel is van de Gezondheidsenquête. Dit waren deels vragen die eerder niet in de GE vragenlijst waren opgenomen.

Het gevolg van deze wijzigingen is een toename van de lengte van de vragenlijst. Dit is nadelig voor de respons. Daarom is de gemiddelde interviewduur voor CAPI bepaald op maximaal 35 minuten (in praktijk is die interviewduur dan iets korter). Om aan deze eis te kunnen voldoen, moesten er vragen worden geschrapt en variabelen efficiënter worden uitgevraagd door het aantal vragen waarop de afleiding van een variabele is gebaseerd terug te dringen. Dit kan door vragen aan te passen of te vernieuwen, zodat andere vragen overbodig worden. Een voorbeeld hiervan is de wijze waarop de vragen voor de variabele Contact met de huisarts is gewijzigd (zie onder).

In dit hoofdstuk wordt voor alle 17 variabelen waarvoor wordt onderzocht of er sprake is van een methodebreuk beschreven welke wijzigingen er in de vragenlijst van 2014 hebben plaatsgevonden ten opzichte van die van 2013. Zie voor een vergelijking van de vragen Bijlage 4.

Ervaren gezondheid

Geen wijzigingen.

Diabetes

Tot en met 2013 was er een aparte vraag over suikerziekte voor kinderen (in deel 1) en volwassenen (deel 2). De vraag zelf is voor kinderen gewijzigd, maar voor volwassenen niet. Wel is er voor volwassenen een introductietekst toegevoegd, voor kinderen is deze tekst gewijzigd.

Verkoudheid

Tot en met 2013 zat de vraag over verkoudheid in deel 2. De vraag zelf is in 2014 niet gewijzigd ten opzichte van 2013. Wel waren kleine er tekstuele wijzigingen in de introductietekst en is de periode van de afgelopen 2 maanden verder gespecificeerd door de betreffende datum toe te voegen: 'dus sinds ...'. In 2014 zijn er 6 vragen over acute aandoeningen, één minder dan in 2013. Ook zijn de vervolgvragen voor de betreffende aandoening komen te vervallen (d.i. raadplegen huisarts, verzuimen dagelijkse bezigheden van school, werk en thuis, frequentie waarmee geplaagd door de aandoening).

Psychische ongezondheid

Deze vragen voor de variabele psychische ongezondheid zaten tot en met 2013 in deel 2. De introductietekst, waarin wordt aangegeven dat de vragen over de afgelopen 4 weken beantwoord moeten worden, is licht gewijzigd doordat er een datumspecificatie voor die periode is toegevoegd. De vragen zelf zijn in 2014 niet gewijzigd. Wel is de in 2013 als vierde gestelde vraag 'Voelde u zich erg energiek' komen te vervallen.

Minstens 1 beperking (OESO-indicator)

De OESO-indicator is gebaseerd op 7 enquêtevragen die tot en met 2013 in deel 2 waren opgenomen. De introductie van het vragenblok met deze vragen is op een aantal punten gewijzigd:

- Aan de zin dat het erom gaat wat iemand normaal kan doen is toegevoegd dat dit is 'zo nodig met hulpmiddelen zoals bril of hoorapparaat die u gewoonlijk gebruikt'.
- In de tweede zin dat het niet gaat om tijdelijke problemen is het woordje 'niet' onderstreept.
- Aan de tweede zin is tussen haakjes toegevoegd 'van voorbijgaande aard'.
- Na de introductie werden tot en met 2013 9 vragen gesteld, vanaf 2014 zijn dit alleen de 7 op basis waarvan de indicator wordt berekend en zijn de 5e en 9e vraag geschrapt. Dit zijn:
 - Kunt u hard voedsel bijten en kauwen zoals bijvoorbeeld een harde appel?
 - Kunt u normaal verstaanbaar praten?

Beperking bewegen (OESO-indicator)

De indicator Beperking in bewegen is gebaseerd op 3 van de 7 OESO vragen (zie vorige paragraaf), namelijk vraag 5 t/m 7. Zoals in de vorige paragraaf al werd aangegeven, is de introductievraag van de OESO-vragen gewijzigd. Daarnaast ging de in 2014 geschrapte vraag over kauwen in 2013 vooraf aan deze drie vragen. In 2014 komt de vraag 'Kunt u op een afstand van 4 meter het gezicht van iemand herkennen?' vóór deze vragen.

Lichamelijke beperking (ADL-indicator)

De ADL-indicator is in 2014 gebaseerd op 11 enquêtevragen. Dit is er één meer dan in 2013, omdat de vraag 'Van het toilet gebruik maken?' als achtste is toegevoegd. Ook is de zevende vraag gewijzigd van 'Zich volledig wassen?' in 'In bad gaan of douchen?'. Daarnaast is in de introductietekst voor deze vragen 'aantal handelingen' veranderd in 'aantal dagelijkse handelingen'.

Contact met de huisarts en Aantal contacten met de huisarts

Het blok Huisarts bevat de vragen over het hebben van contact met de huisarts in het afgelopen jaar en het aantal contacten met de huisarts in de afgelopen vier werken. Dit blok is in 2014 substantieel gewijzigd ten opzichte van 2013 (zie voor een directe vergelijking van de vragen Bijlage 4):

- Het aantal voor deze variabelen relevante vragen is teruggebracht van vier naar twee, waarbij 3 vragen zijn geschrapt en 1 nieuwe vraag is toegevoegd. In 2013 staan deze vragen aan het begin van het blok, in 2014 zijn dit de enige vragen in het blok.
- De algemene introductie van het blok is gehandhaafd, met tekstuele wijzigingen:
 - De eerste twee zinnen zijn herschreven. Dit was tot 2014:
'De volgende vragen gaan over contacten met de *huisarts* in Nederland. Het gaat hierbij om bezoeken op het spreekuur, visites van de huisarts en telefonische consulten.'
Vanaf 2014 is dit ingekort tot:
'Nu iets over contacten met de huisarts. Denkt u hierbij aan bezoeken aan de huisartspraktijk, huisbezoeken en telefonische consulten.'
 - In de laatste zin is het woordje 'niet' in 2014 onderstreept om te benadrukken dat de praktijkondersteuner en -verpleegkundige niet moeten worden meegeteld.
- In 2014 is de eerste vraag na de algemene introductie (HACont) nieuw en vraagt wanneer de respondent voor het laatst voor zichzelf contact heeft gehad met de huisarts (minder of meer dan 12 maanden geleden, of nooit). Deze vraag wordt in 2014 gebruikt om te bepalen of iemand in het afgelopen jaar de huisarts heeft bezocht.
- De tweede vraag in 2014 was de eerste vraag in 2013 en daarvoor (BezKHA). De tekst is hetzelfde gebleven, afgezien van het onderstrepen van de tijdsbepaling 'in de afgelopen 4 weken' en het verlagen van het maximum aantal contacten dat mag worden ingevuld van 97 naar 28. Deze vraag wordt in 2013 gebruikt om te bepalen of iemand in het afgelopen jaar contact heeft gehad met de huisarts.
- Indien respondenten in 2013 de vraag (BezKHA) niet willen/kunnen beantwoorden of aangeven dat zij in de afgelopen 4 weken geen contact hadden met de huisarts, beantwoorden zij ook de vraag of zij in de afgelopen 12 maanden dan wél contact hadden (LastHA). Deze laatste vraag is in 2014 geschrapt en werd niet gebruikt om het aantal contacten met de huisarts in het afgelopen jaar te bepalen.
- Dat gold wel voor de laatste twee vragen in het blok Huisarts in 2013 (DagMndHA1 en DagMndHa2). Deze vragen zijn in 2014 geschrapt en vroegen naar de datum van het laatste en, indien van toepassing, het voorlaatste contact met de huisarts.

Contact met de specialist

Om te bepalen of een respondent in het afgelopen jaar contact had met de specialist, zijn de twee vragen die hiertoe werden gesteld t/m 2013 in 2014 vervangen door één nieuwe vraag. De introductie tot deze vragen is gehandhaafd met enkele tekstuele wijzigingen:

- De eerste zin was in 2013 'De volgende vragen gaan over contacten met specialisten in Nederland' en is vanaf 2014 'Nu iets over contacten met specialisten'.

- Tekst die in 2013 betrekking had op contacten met de psychiater, die in 2013 niet mochten worden meegeteld, is in 2014 geschrapt.

Ziekenhuisopname

De vraag met introductie naar ziekenhuisopnames in de afgelopen 12 maanden is in 2014 ten opzichte van daarvoor alleen tekstueel aangepast:

- In 2014 is de introductie anders geformuleerd. Daarbij wordt niet meer expliciet vermeld dat psychiatrische ziekenhuizen niet mogen worden meegeteld. De zin dat opnames gerelateerd aan bevalling en geboorte niet mogen worden meegeteld stond in 2013 in de vraag, maar is in 2014 verplaatst naar de introductie. De volledige tekst luidt in 2013: 'Nu volgen enkele vragen over ziekenhuisopnamen in Nederland. Dit kunnen allerlei soorten ziekenhuizen zijn, behalve psychiatrische ziekenhuizen.'
- en in 2014: 'Nu iets over *ziekenhuisopnamen*. Dit kunnen allerlei soorten ziekenhuizen zijn. (§3: Het gaat hierbij niet om opname in verband met bevalling of geboorte.)'
- In de vraagtekst is in 2014 onderstreping toegevoegd aan 'in de afgelopen 12 maanden' en de zin over opname bij bevalling of geboorte is vanaf 2014 verplaatst naar de introductie.

Zelfmedicatie

In het blok met de vraag over het gebruik van medicijnen die niet door een arts zijn voorgeschreven is de introductie geschrapt:

'Nu volgen enkele vragen over medicijngebruik en het gebruik van voedingssupplementen, zoals vitamines, mineralen en kruidenpreparaten.'

Verder is de vraagtekst veranderd, waarbij 'voedingssupplementen' is vervangen door 'kruidengeneesmiddelen of vitamines en 'in de afgelopen dagen' is onderstreept.

Mammografie

Om te achterhalen of vrouwen in de afgelopen 2 jaar een mammografie hebben laten maken, werd er in 2013 één vraag gesteld die werd ingeleid door een introductie. In deze introductie werd verwezen naar het bevolkingsonderzoek borstkanker. Deze verwijzing is in 2014 verwijderd, en vraag en introductie zijn samengevoegd tot één vraag met uitleg, waarbij de vraagstelling is aangepast. In de vraagstelling van 2014 is de term mammografie toegevoegd. Alleen de specificatie van het soort röntgenfoto's dat niet in het beantwoorden van de vraag moet worden meegenomen is gehandhaafd. De aard van de vraag is eveneens veranderd: vóór 2013 werd gevraagd naar de frequentie waarmee een borstfoto is gemaakt, in 2014 naar de laatste keer dat dit het geval was. De antwoorden zijn daarmee uiteraard ook gewijzigd en in 2014 is er één antwoordcategorie meer dan in 2013. Deze vraag zat tot en met 2013 in deel 2.

Roker

Geen wijzigingen.

Zware drinker onder drinkers en alcoholconsumptie van drinkers

Het blok Alcohol zat tot en met 2013 in deel 2. Het wordt zowel gebruikt voor het bepalen van het percentage zware drinkers als voor het aantal glazen alcohol dat wordt geconsumeerd. In 2013 bevat dit blok 10 vragen (waarvan 2 niet voor de indicatoren werden gebruikt) en in 2014 een introductie gevolgd door 8 vragen.

- In 2013 betrof de eerste vraag in het blok welke soorten alcoholhoudende drank de respondent in de afgelopen 12 maanden heeft gedronken. Als antwoordopties worden zes

- verschillende soorten drank genoemd en kan de respondent daarnaast aangeven dat hij of zij in die periode geen alcohol heeft gedronken of nooit alcohol dronk. Deze eerste vraag is in 2014 vervangen door een introductie ('De volgende vragen gaan over het gebruik van alcohol.') en twee nieuwe vragen over of men in de afgelopen 12 maanden alcohol heeft gedronken en zo nee, of men ooit alcohol heeft gedronken. Als men in deze periode geen alcohol dronk of nooit alcohol heeft gedronken, worden de daarop volgende vragen over het volume van de alcoholconsumptie niet gesteld. Door deze veranderingen is de wijze waarop wordt bepaald of iemand in de afgelopen 12 maanden alcohol dronk gewijzigd.
- De volgende vier vragen worden gebruikt om bij drinkers het gemiddeld aantal glazen alcohol dat men per dag drinkt te bepalen. In deze vragen wordt gevraagd hoeveel glazen alcoholhoudende drank men doorgaans door de week en in het weekend drinkt. Naast enkele kleine tekstuele wijzigingen (komma's i.p.v. haakjes), is in de vragen naar het aantal glazen alcohol dat men drinkt de antwoordoptie '11 of meer glazen' vervangen door de opties '11–15 glazen' en '16 of meer glazen'.
 - Vervolgens zijn 2 vragen uit 2013 geschrapt, over wanneer men voor het laatst alcohol dronk en hoe vaak men dit dan deed.
 - Daarna volgen de twee vragen die worden gebruikt om te bepalen of drinkers ook zware drinkers zijn. In beide vragen is 'in de afgelopen 6 maanden' in 2014 onderstreept (in 2013 niet) en is hieraan in de vraag of men meer dan 4 glazen op één dag heeft gedronken een datumspecificatie toegevoegd.

Ernstig overgewicht

Geen wijziging.

3.4 Wijziging van het weegmodel

Het weegmodel van de GE is in 2014 op twee punten aangepast:

1. Het weegmodel voor 2010 t/m 2013 hield rekening met mogelijke selectiviteit van uitval na het eerste onderzoeksdeel. Om dit op te vangen werd voor onderzoeksdeel 2 een aparte weegfactor berekend (Bruggink en Van Herten, 2014). Met het verdwijnen van het tweede deel is deze extra weegfactor komen te vervallen en is het weeg model eenvoudiger geworden.
2. Het weegmodel van 2014 is uitgebreid met de variabelen inkomen en vermogen, om zo een meer representatief cijfer te kunnen opleveren (Boonstra, 2015).

4. Resultaten

In dit hoofdstuk worden de uitkomsten van de analyses van de methodebreuk van de 17 GE-variabelen beschreven. Allereerst wordt gekeken of de resultaten mogen worden geïnterpreteerd (4.1). Dit blijkt het geval en daarom worden in de vervolgpargrafen de methodebreuken gepresenteerd voor de hele groep (paragraaf 4.2), voor mannen en vrouwen apart (4.3) en voor de drie leeftijdsgroepen 0 tot 25 jaar, 25 tot 65 jaar en 65 jaar en ouder (4.4). In paragraaf 4.5 worden de bevindingen voor elke variabele apart beschreven.

4.1 Voorwaarden statistisch model

Om de methodebreuk van de geselecteerde variabelen te berekenen, zijn de tijdreeksen ervan geanalyseerd volgens het statistisch model uit paragraaf 0. Allereerst werd onderzocht of het model de tijdreeksen goed beschrijft. Dit is een voorwaarde om de regressiecoëfficiënten die de analyse oplevert te mogen interpreteren als methodebreuk. Het bleek dat dit het geval was: (1) het iteratieve proces dat de modelparameters schat, bleek sterk convergent te zijn, (2) vrijwel alle innovaties kwamen door de normaliteitstest (en zo niet bleek de afwijking van normaliteit klein), (3) de seriële correlaties van de innovaties waren niet substantieel en (4) hoewel de formele test in een aantal gevallen heteroscedasticiteit aantoonde, bleek dit niet uit visuele inspectie (en ook inhoudelijk is er geen reden om dit aan te nemen) en bleek een enkele uitbijter meestal de oorzaak. Op basis hiervan mag worden geconcludeerd dat de berekende breuken correct zijn. Vervolgens is getest of methodebreuken die uit de analyses volgden significant waren en is de bijbehorende p waarde berekend. In de volgende paragrafen worden de methodebreuken beschreven die uit de analyses naar voren kwamen.

4.2 Methodebreuk gehele groep

Tabel 4.2.1 toont de methodebreuken van de totale groep en ook de cijfers op basis waarvan deze breuken zijn berekend (zie voor de cijfers ook Bijlage 5). Bij de breuken wordt er onderscheid gemaakt tussen een wegingsbreuk en een designbreuk. De wegingsbreuk is de breuk die optreedt als gevolg van de weging. De designbreuk is de breuk die optreedt

4.2.1 Methodebreuken GE 2014

Variabele ¹⁾	Cijfers		schatting tijdreeks- model	Breuken ²⁾		
	cijfer 2014 (met nieuwe weging)	cijfer met oude weging		totaal	design	weging
Ervaren gezondheid	79,46	80,73	80,38	-0,92	0,35	-1,26 ³
Diabetes	4,65	4,48	4,74	-0,09	-0,26	0,18 ³
Verkoudheid	36,69	36,23	44,03	-7,34***	-7,80 ³	0,46 ³
Psychische ongezondheid	11,32	10,46	10,45	0,86	0,01	0,85 ³
Minstens 1 beperking OESO	13,60	12,48	13,00	0,60	-0,52	1,12 ³
Beperking in bewegen OESO	10,62	9,63	7,41	3,22***	2,22 ³	1,00 ³
Lichamelijke beperking ADL	15,61	14,41	11,12	4,49***	3,29 ³	1,20 ³
Contact met de huisarts	68,23	67,80	71,22	-2,99***	-3,42 ³	0,43
Aantal contacten met de huisarts	4,96	4,85	4,13	0,83***	0,73	0,11 ³
Contact met de specialist	39,69	39,33	38,16	1,52	1,16	0,36 ³
Ziekenhuisopname	7,63	7,45	6,83	0,80	0,62	0,18 ³
Zelfmedicatie	38,55	38,68	32,78	5,76***	5,90 ³	-0,14
Mammografie	45,69	45,75	48,51	-2,82*	-2,76 ³	-0,06
Roken	24,89	23,89	22,52	2,37***	1,37 ³	1,00 ³
Zware drinker	8,42	8,17	11,71	-3,29***	-3,54 ³	0,26
Alcoholconsumptie	1,36	1,35	1,18	0,18***	0,17 ³	0,01
Ernstig overgewicht	12,20	11,53	10,37	1,84***	1,16 ³	0,67 ³

¹⁾ Alle cijfers en breuken zijn percentages, behalve bij Aantal contacten met de huisarts en Alcoholconsumptie (glazen alcohol).

²⁾ De breuken worden als volgt berekend: Totaalbreuk = cijfer 2014 (met nieuwe weging) min cijfer 2014 geschat met tijdreeksmodel; Designbreuk = cijfer 2014 met oude weging min cijfer 2014 geschat met tijdreeksmodel; Wegingsbreuk = cijfer 2014 met oude weging min cijfer 2014 (met nieuwe weging).

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$.

³⁾ Significant als designbreuk en wegingsbreuk onafhankelijk zouden zijn.

als gevolg van andere wijzigingen in het onderzoek van de GE, zoals wijzigingen in het onderzoeksdesign, de benaderingsstrategie en de vragenlijst. Wijzigingen in de cijfers door het optreden van een designbreuk kunnen overigens niet worden onderscheiden van reële wijzigingen in het cijfer van 2014.

Voordat we deze resultaten bespreken, is het belangrijk om te vermelden dat de cijfers en de breuken van de verschillende variabelen in alle gevallen percentages betreffen, behalve bij twee. De eerste van die twee is contact met de huisarts, die is uitgedrukt in het gemiddeld aantal contacten per persoon per jaar. De tweede variabele is alcoholconsumptie, die het gemiddeld aantal glazen alcohol inhoudt dat per dag per drinker wordt genuttigd (zie verder Bijlage 1).

Het algemene beeld dat Tabel 4.2.1 laat zien, is dat de cijfers door de methodebreuk in 2014 een negatiever beeld schetsen van de gezondheid, de medische consumptie en de leefstijl dan voorheen. Dit geldt voor 12 van de 17 variabelen; alleen voor diabetes, verkoudheid, contact met de huisarts, mammografie en zware drinker is het cijfer indicatief voor een betere gezondheid of een gezondere leefstijl.

De gevonden breuk is voor 11 van de 17 onderzochte variabelen significant. Voor deze 11 variabelen is het zeer waarschijnlijk is dat de verandering in het cijfer van 2014 ten opzichte van voorgaande jaren ten minste deels kan worden toegeschreven aan het herontwerp. Deze mate van waarschijnlijkheid is hoger naarmate de p -waarde van een verschil kleiner is; het significantieniveau is dan hoger. Zie ook Bijlage 6, waarin voor elke variabele in een grafiek de cijferreeks en de breuk wordt weergegeven.

De 11 variabelen waarvan het cijfer door het herontwerp significant veranderd zijn: verkoudheid, beperking in bewegen (OESO), lichamelijke beperking (ADL), contact met de huisarts, aantal contacten met de huisarts, zelfmedicatie, mammografie, roken, zware drinker, alcoholconsumptie en ernstig overgewicht. De grootte van de breuk varieert daarbij van 1,8 tot 7,3%. Het aantal contacten met de huisarts verandert met 0,8 en de alcoholconsumptie (van drinkers) met 0,2 consumpties. Breuken die er qua omvang uitschieten zijn die bij verkoudheid (-7,3%), zelfmedicatie (5,7%), lichamelijke beperking (4,5%), zware drinker (3,3%) en beperking in bewegen (3,2%).

De totaalbreuk kan worden opgesplitst in een deel dat wordt verklaard door het design en een deel dat wordt verklaard door veranderingen in het weegmodel. De *design*breuken zijn goed vergelijkbaar met de totaalbreuken, zowel qua richting als qua significantie. De richting is hetzelfde, behalve bij ervaren gezondheid en minstens 1 beperking. In deze twee gevallen omvat de totaalbreuk een *verslechtering* van de gezondheid en de designbreuk een *verbetering*.

De resultaten voor de breuken als gevolg van de aangepaste *weging* geven een heel ander beeld. De nieuwe weging leidt tot een ongezonder beeld van de gezondheid dan de oude weging. Uitzondering hierop zijn de variabelen zelfmedicatie en mammografie, al zijn deze effecten klein. De designbreuk is doorgaans groter dan de wegingsbreuk, maar niet voor ervaren gezondheid, diabetes en psychische ongezondheid. De breuk bij deze variabelen wordt (vrijwel) volledig verklaard door de weging. Het omgekeerde, waarbij de totaalbreuk (vrijwel) volledig wordt verklaard door de designbreuk, komt vaker voor: bij verkoudheid, contact met de huisarts, aantal contacten met de huisarts, zelfmedicatie, mammografie, zware drinker en alcoholconsumptie. Tot slot is het effect van design en weging bij enkele variabelen tegengesteld; bij ervaren gezondheid en minstens 1 beperking lijkt dit te verklaren waarom de totaalbreuk niet significant is.

4.3 Methodebreuk naar geslacht

Tabel 4.3.1 toont de methodebreuken voor mannen en vrouwen apart; de cijfers op basis waarvan deze breuken zijn berekend staan in Bijlage 5. De totaalbreuken zijn over het algemeen groter voor vrouwen dan voor mannen (11 van de 16²⁾ breuken). Voor de wegingsbreuk geldt dit in nog hogere mate (14 van de 16 groter voor vrouwen), maar voor de designbreuk geldt dit niet (7 van de 16 groter). Voor drie variabelen is het verschil voor de totaalbreuk tussen mannen en vrouwen groter dan 2%, namelijk bij verkoudheid, zware drinker en contact met de specialist. Bij die laatste is de breuk overigens groter bij mannen.

4.3.1 Methodebreuken GE 2014 voor mannen en vrouwen

Variabele ¹⁾	Mannen			Vrouwen		
	totaal	design	weging	totaal	design	weging
Ervaren gezondheid	-0,70	0,41	-1,11 [‡]	-0,89	0,53	-1,41 [‡]
Diabetes	-0,03	-0,19	0,16	0,11	-0,09	0,20 [‡]
Verkoudheid	-8,42***	-8,83 [‡]	0,40	-6,18***	-6,70 [‡]	0,52 [‡]
Psychische ongezondheid	0,02	-0,53	0,55 [‡]	1,66*	0,51	1,14 [‡]
Minstens 1 beperking OESO	-0,16	-1,12	0,96 [‡]	1,35	0,08	1,27 [‡]
Beperking in bewegen OESO	2,20***	1,42 [‡]	0,78 [‡]	4,09***	2,89 [‡]	1,20 [‡]
Lichamelijke beperking ADL	4,40***	3,26 [‡]	1,14 [‡]	3,56***	2,31 [‡]	1,25 [‡]
Contact met de huisarts	-2,72**	-3,26 [‡]	0,54	-3,45**	-3,78 [‡]	0,33
Aantal contacten met de huisarts	0,77***	0,71 [‡]	0,06	0,81**	0,67 [‡]	0,15 [‡]
Contact met de specialist	2,62*	2,16	0,46 [‡]	0,23	-0,03	0,26
Ziekenhuisopname	0,57	0,41	0,16	0,97*	0,78	0,19
Zelfmedicatie	4,86***	4,92 [‡]	-0,06	6,52***	6,73 [‡]	-0,21
Mammografie	.	.	.	-2,82*	-2,76 [‡]	-0,06
Roken	2,84*	1,91	0,94 [‡]	2,42***	1,35 [‡]	1,07 [‡]
Zware drinkerz	-4,33***	-4,53 [‡]	0,20	-2,19***	-2,45 [‡]	0,26 [‡]
Alcoholconsumptie	0,18***	0,18 [‡]	0,00	0,22***	0,21 [‡]	0,01
Ernstig overgewicht	0,79	0,37	0,42 [‡]	2,75***	1,82 [‡]	0,93 [‡]

¹⁾ Alle breuken zijn percentages, behalve bij Aantal contacten met de huisarts en Alcoholconsumptie (glazen alcohol).

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$.

[‡]Significant als designbreuk en wegingsbreuk onafhankelijk zouden zijn.

Voor de designbreuken geldt eveneens dat deze voor de meeste variabelen groter is voor vrouwen dan voor mannen. Het differentiële effect van de weging op geslacht is aanzienlijk kleiner dan dat van het design: maximaal 0,6% ten opzichte van 2,2% bij de designbreuk. De weging heeft vooral een differentieel effect op de variabelen psychische ongezondheid, ernstig overgewicht en minstens 1 beperking.

Wanneer we de individuele variabelen bekijken, valt een aantal zaken op:

1. Er is alleen een breuk bij vrouwen voor psychische ongezondheid, ziekenhuisopname en ernstig overgewicht. Bij psychische ongezondheid en ziekenhuisopname treedt die breuk niet op bij de gehele groep, bij ernstig overgewicht wel.
2. Er is alleen een breuk bij mannen voor contact met de specialist. Deze breuk is niet aanwezig in de gehele groep.
3. De breuken bij beperking in bewegen, zelfmedicatie en ernstig overgewicht zijn groter bij vrouwen dan bij mannen.
4. De breuken bij verkoudheid, contact met de specialist en zware drinker zijn groter bij mannen dan bij vrouwen.

²⁾ Voor mammografie is geen cijfer beschikbaar voor mannen.

4.4 Methodebreuk naar leeftijd

In Tabel 4.4.1 staan de breuken voor de drie leeftijdsgroepen. Het valt op dat in de laagste leeftijdsgroep van 0 tot 25 jaar over het algemeen minder (significante) en kleinere breuken worden gevonden dan in de andere twee leeftijdsgroepen. De breuken die wel worden gevonden, zijn echter meestal groter. Dit geldt voor de negatieve breuken bij contact met de huisarts en zware drinker en voor de positieve breuken bij roken en alcoholconsumptie, maar niet voor verkoudheid. Deze breuken kunnen worden over het algemeen voor een groter deel verklaard door het design.

4.4.1 Methodebreuken GE 2014 voor drie leeftijdsgroepen

Variabele ¹⁾	0–25 jaar			25–65 jaar			65 jaar en ouder		
	totaal	design	weging	totaal	design	weging	totaal	design	weging
Ervaren gezondheid	-0,70	-0,03	-0,66	-0,61	0,83	-1,45 [§]	-1,15	0,56	-1,71 [§]
Diabetes	-0,10	-0,08	-0,03	0,06	-0,12	0,19 [§]	-1,04	-1,54 [§]	0,50 [§]
Verkoudheid	-4,72*	-5,24 [§]	0,51	-8,34***	-8,95 [§]	0,61 [§]	-5,45**	-5,39 [§]	-0,05
Psychische ongezondheid	0,31	-0,16	0,47	1,31*	0,28	1,03 [§]	0,16	-0,50	0,66 [§]
Minstens 1 beperking OESO	-1,62	-1,89 [§]	0,27	-0,41	-1,55 [§]	1,14 [§]	6,92***	5,06 [§]	1,86 [§]
Beperking in bewegen OESO	-0,71	-0,88	0,16	2,51***	1,50 [§]	1,02 [§]	9,75***	8,03 [§]	1,72 [§]
Lichamelijke beperking ADL	.	.	.	3,58***	2,63 [§]	0,95 [§]	4,61**	3,22 [§]	1,39 [§]
Contact met de huisarts	-4,65***	-5,27 [§]	0,61	-3,25***	-3,66 [§]	0,41 [§]	-0,16	-0,36	0,21
Aantal contacten met de huisarts	0,45	0,37	0,08	0,77***	0,66 [§]	0,11 [§]	1,52***	1,38 [§]	0,13
Contact met de specialist	0,39	-0,08	0,47	1,97	1,61	0,36	0,67	0,50	0,17
Ziekenhuisopname	0,07	0,08	-0,02	0,98	0,72	0,26 [§]	1,20	0,96	0,24
Zelfmedicatie	2,17	2,18	-0,01	7,00***	7,10 [§]	-0,09	7,82***	8,32 [§]	-0,50
Mammografie	.	.	.	-3,35*	-3,48 [§]	0,13	-2,30	-1,73	-0,56
Roken	3,97*	3,12	0,85 [§]	2,57*	1,45	1,11 [§]	1,96	1,16	0,80 [§]
Zware drinker	-4,75	-5,42 [§]	0,67	-2,82***	-2,99 [§]	0,17	-1,76	-1,95	0,19
Alcoholconsumptie	0,27**	0,22 [§]	0,05	0,05	0,04	0,00	0,22***	0,25 [§]	-0,03 [§]
Ernstig overgewicht	0,89	0,50	0,38 [§]	2,35***	1,74 [§]	0,61 [§]	1,89*	0,63	1,26 [§]

¹⁾ Alle breuken zijn percentages, behalve bij Aantal contacten met de huisarts en Alcoholconsumptie (glazen alcohol).

* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$.

[§]Significant als designbreuk en wegingsbreuk onafhankelijk zouden zijn.

In de groep van 65 jaar of ouder zijn vooral de breuken gerelateerd aan beperkingen groot: minstens 1 beperking, beperking in bewegen en lichamelijke beperking. Dit geldt ook voor de breuk bij zelfmedicatie. De grootste breuk wordt gevonden in deze groep, namelijk van 9,8% bij beperking in bewegen. Opvallend is dat er in deze groep geen breuk wordt gevonden voor contact met de huisarts, maar dat het aantal contacten juist erg hoog ligt. Ook voor de variabele zware drinker is er bij 65-plussers geen breuk.

In de leeftijdsgroep van 25 tot 65 jaar zijn de breuken voor verkoudheid, psychische ongezondheid, mammografie en ernstig overgewicht groter dan in de andere leeftijdsgroepen. Die bij alcoholconsumptie is juist kleiner (niet significant).

4.5 Methodebreuk per variabele

In deze paragraaf wordt de methodebreuk voor elke variabele apart beschreven.

Ervaren gezondheid

Voor ervaren gezondheid wordt geen totaalbreuk gevonden. Opvallend is dat het effect van het design positief is en dat van de weging negatief. Alleen het effect ten gevolge van de weging is significant en leidt tot een lager percentage mensen dat zijn of haar gezondheid als goed of zeer goed ervaart. Dit effect is iets sterker bij vrouwen dan bij mannen en bij 25–65-jarigen en vooral 65-plussers dan bij de jongste leeftijdsgroep.

Diabetes

Voor diabetes wordt eveneens geen totaalbreuk gevonden. Wel is er een klein effect van de weging dat het duidelijkst is in de groep van 65 jaar of ouder, dat een hoger percentage diabetici tot gevolg heeft. Bij die laatste groep is een iets groter tegengesteld effect van het design.

Verkoudheid

De totaalbreuk betreft een daling van 7,3% voor de variabele verkoudheid. Dit effect is groter bij mannen dan bij vrouwen en groter in de leeftijdsgroep van 25–65 jaar in vergelijking tot de andere twee. Deze totaalbreuk is een resultante van een nog iets groter designeffect (7,8%), dat enigszins wordt getemperd door de weging. Dit geldt voor alle subgroepen, behalve 65-plussers.

Psychische ongezondheid

Bij psychische ongezondheid is de totaalbreuk alleen significant bij vrouwen en in de leeftijdsgroep van 25–65 jaar. De psychische gezondheid wordt hierdoor minder positief ingeschat. Dit wordt verklaard doordat de breuk als gevolg van de weging hier significant is (dit geldt overigens voor alle subgroepen behalve 0–25-jarigen). De designbreuk is in geen enkele subgroep significant, maar is tegengesteld aan de wegingsbreuk, behalve bij vrouwen en 25–65-jarigen. Dit verklaart dan ook de significante totaalbreuk in deze subgroepen.

Minstens 1 beperking (OESO-indicator)

Er is geen breuk voor de variabele minstens 1 beperking voor de totale groep, maar wel in de groep van 65 jaar en ouder. Het percentage mensen met minstens 1 beperking neemt hierdoor met bijna 7% toe. Dit wordt grotendeels verklaard door het design (bijna 5%), maar in mindere mate ook door de weging. Opvallend is verder dat het designeffect tegenovergesteld is in de andere twee leeftijdsgroepen en hier dus juist leidt tot een afname van het percentage mensen met minstens 1 beperking.

Beperking in bewegen (OESO-indicator)

De totaalbreuk voor beperking in bewegen wordt eveneens verklaard door 65-plussers; hier wordt de grootste breuk van bijna 10 procent geobserveerd. Anders dan bij minstens 1 beperking wordt hier echter ook door de groep van 25–65 jaar bijgedragen aan het verklaren van de totaalbreuk. Dit komt voor het grootste deel door het design en in mindere mate door de weging (behalve bij de groep van 0–25 jaar, waar geen breuken waren voor deze variabele). Opvallend is dat de totaalbreuk bij vrouwen (ruim 4%) bijna twee keer zo groot is als die van mannen (ruim 2%).

Lichamelijke beperking (ADL-indicator)

Het patroon van bevindingen bij de variabele lichamelijke beperking is vergelijkbaar met dat bij de variabele beperking in bewegen. Wel is de breuk groter, wellicht omdat deze variabele vanaf 55 jaar wordt gemeten. Hierdoor is het verschil met de breuk bij 25–65-jarigen wellicht minder groot. Het verschil tussen mannen en vrouwen is niet groot.

Contact met de huisarts

De totaalbreuk voor contact met de huisarts, die tot een lager percentage mensen dat de huisarts bezoekt leidt, wordt volledig verklaard door het design. Het effect van de weging is klein, maar in tegenovergestelde richting. Deze effecten zijn groter bij vrouwen dan bij mannen, maar worden vooral verklaard door de leeftijdsgroepen van 0–25 en 25–65 jaar.

Aantal contacten met de huisarts

De interpretatie van het aantal contacten met de huisarts is anders dan die van de overige variabelen, omdat de breuk geen percentage betreft. De waargenomen totaalbreuk van 0,8 lijkt klein, maar omvat uitgaande van de nieuwe weging een toename in het aantal contacten van ruim 19 procent ten opzichte van de voorspelde 4,1 contacten gemiddeld. Deze toename komt vooral op het conto van 65-plussers (deze is hier zo'n 40%) en wordt meestendeels verklaard door het design.

Contact met de specialist

Voor de gehele groep wordt bij contact met de specialist geen totaalbreuk waargenomen. Dit geldt eveneens voor de drie leeftijdsgroepen en vrouwen. Bij mannen is er echter wel sprake van een breuk, hetgeen een optelsom is van het effect van het design en van de weging.

Ziekenhuisopname

Voor ziekenhuisopname geldt hetzelfde als voor contact met de specialist, alleen is er hier wel een effect bij vrouwen, maar niet bij mannen.

Zelfmedicatie

Bij zelfmedicatie is er een aanzienlijke totaalbreuk, die geheel te verklaren is door het design en op het conto komt van 25–65-jarigen en 65-plussers. Bij vrouwen is de breuk iets groter dan bij mannen.

Mammografie

De totaalbreuk bij mammografie wordt eveneens verklaard door het design en dan met namen doordat er een breuk wordt gevonden bij vrouwen van 25 tot 65 jaar.

Roken

Zowel een breuk door het design als door de weging verklaren bij roken de totaalbreuk. Bij deze variabele is de designbreuk groter. Er is geen verschil naar geslacht, maar bij de 65-plussers is er geen totaalbreuk (wel een kleine wegingsbreuk).

Zware drinker

Door de totaalbreuk bij zware drinker wordt dit percentage bijgesteld naar beneden. Dit kan worden toegeschreven aan het design en geldt meer voor mannen dan voor vrouwen. De sterkste breuk zit in de groep van 25 tot 65 jaar en wordt verklaard door het design. In de jongste leeftijdsgroep is alleen de designbreuk significant.

Alcoholconsumptie

De interpretatie van alcoholconsumptie wijkt, net als die van het aantal contacten met de huisarts, af van de andere variabelen. De alcoholconsumptie van drinkers wordt uitgedrukt in glazen alcohol. Voor de totaalbreuk betekent dit dat de consumptie met bijna een vijfde glas naar boven wordt bijgesteld, een toename van zo'n 15% ten opzichte van de op basis van het model geschatte consumptie van 1,2 glazen. De breuk bij alcoholconsumptie wordt volledig verklaard door het design, verschilt weinig qua geslacht en wordt verklaard door de oudste en jongste leeftijdsgroepen.

Ernstig overgewicht

Er is een totaalbreuk voor ernstig overgewicht, maar die wordt vooral verklaard door vrouwen en personen van 25 jaar of ouder. Deze breuk kan aan zowel het design als de weging worden toegeschreven. De bijdrage van de weging aan de totaalbreuk is relatief groter bij 65-plussers dan bij 0–25- en vooral 25–65-jarigen.

5. Discussie

In dit hoofdstuk wordt getracht een verklaring te vinden voor de gevonden methodebreuken, in het algemeen (5.1) en per variabele (5.2). Vervolgens worden in hoofdstuk 6 de conclusies gepresenteerd die op basis hiervan kunnen worden getrokken.

5.1 Algemene verklaringen

Er is een significante methodebreuk bij een meerderheid van de variabelen (11 van de 17). Deze breuk varieert tussen 1,8 en 7,3 procent. Daarnaast zijn er twee variabelen met een significante breuk die niet in procenten wordt uitgedrukt: het aantal contacten met de huisarts met een breuk van 0,8 contacten en alcoholconsumptie met een breuk van 0,2 glazen. De breuk resulteert voor de meeste variabelen in een minder goede gezondheid, meer zorgconsumptie en een minder gezonde leefstijl van de Nederlandse bevolking. Uitzonderingen zijn verkoudheid, contact met de huisarts, mammografie en zware drinker, waar juist sprake is van 'positieve' breuken.

De variabelen die een significante totaalbreuk hebben, vertonen allen een duidelijke designbreuk in dezelfde richting. Wanneer er daarnaast een aanzienlijke wegingsbreuk wordt gevonden, is deze voor alle variabelen behalve verkoudheid in dezelfde richting als de designbreuk, maar kleiner. Dit laat zien dat de invloed van het wijzigen van de benaderingsstrategie en van de structuur en inhoud van de vragenlijst voor de meeste variabelen meer invloed heeft dan het wijzigen van het weegmodel. Het individuele effect van de verschillende wijzigingen die de designbreuk omvat op de cijfers, en dus op het ontstaan van een methodebreuk, kunnen echter niet worden onderscheiden. Zo is voor alle variabelen de benaderingsstrategie veranderd, maar is de breuk niet voor alle variabelen even groot of in dezelfde richting. En wanneer een variabele voor 2014 in deel 2 van de vragenlijst zat, betekent dit niet dat er een grotere breuk gevonden wordt dan voor variabelen die in deel 1 zaten. Idem dito voor het al dan niet veranderen van de vraagstelling. Over het effect van het vernieuwen van de weging kunnen wel conclusies worden getrokken, want de wegingsbreuk leidt bij vrijwel alle variabelen tot een negatiever beeld van de

gezondheid. Dit is conform de verwachting, aangezien het weegmodel van 2014 is uitgebreid met de variabelen inkomen en vermogen. Deze variabelen hangen sterk samen met sociaal-economische status (SES), waarbij lage SES wordt geassocieerd met een minder goede gezondheid. Omdat de respons bij groepen met een lage SES gemiddeld lager ligt, is het plausibel dat de uitbreiding van het weegmodel met inkomen en vermogen een realistische bijstelling van de cijfers van 2014 naar beneden betekent. Relatief ondervertegenwoordigde groepen in de respons, die tevens gemiddeld een minder goede gezondheid hebben, te weten laag inkomen en weinig vermogen³⁾, krijgen door de nieuwe weging namelijk een zwaarder gewicht.

Er zijn enkele systematische verschillen tussen de breuken van vrouwen en mannen en tussen de breuken van de drie leeftijdsgroepen. Zo zijn de breuken gemiddeld kleiner bij mannen en in de jongste leeftijdsgroep (0 tot 25 jaar). Dit lijkt te komen door een onderliggend 'prevalentie-effect', waarbij de breuken groter zijn wanneer het gezondheidsaspect vaker voorkomt. Zo is ruim 13% van de mannen zware drinker tegenover 3% van de vrouwen. De bijbehorende breuk van 4,3% bij mannen is ongeveer 2 keer zo groot als die bij vrouwen. Hetzelfde effect is waarneembaar in de leeftijdsgroep van 65-plussers voor de beperkingen variabelen. Zo is de breuk voor het hebben van minstens één beperking voor die groep met bijna 7% aanzienlijk hoger dan de niet-significante breuk bij 0–25- en 25–65-jarigen (0,4 en 1,6%). Dit geldt ook voor de prevalentie ervan, die in oplopende leeftijd respectievelijk 3%, 11% en 33% is. Het lijkt er dus op dat breuken over het algemeen groter zijn als de prevalentie van een verschijnsel hoog is.

5.2 Verklaring per variabele

In deze paragraaf wordt voor elk van de zeventien variabelen apart gezocht naar verklaringen voor de geobserveerde methodebreuken. Hierbij wordt vooral ingegaan op het designeffect. Dit is met name omdat het weegeffect makkelijker te duiden is, namelijk aan de aanpassing van het weegmodel. Voor de designeffecten geldt dat verschillende factoren een rol kunnen spelen, waaronder de verandering van de opzet van de vragenlijst, van de benaderingsstrategie en de mode, van de positie van de vraag in de vragenlijst en van verandering van de vraagstelling. Het weegeffect is vrij stabiel in de richting van een minder goede gezondheid, een hogere zorgconsumptie en een ongezonder leefstijl.

Ervaren gezondheid

Voor ervaren gezondheid is er geen breuk ten gevolge van de veranderingen in de mode, het design en de vragenlijst. Dit is in lijn met de bevinding van Wong et al. (2011), die eveneens geen methodebreuk vonden voor ervaren gezondheid als gevolg van wijzigingen in de mode of in de vragenlijst van de Gezondheidsenquête in 2010. Dit suggereert dat ervaren gezondheid relatief resistent is tegen veranderingen in de mode en de benaderingsstrategie. Er was overigens geen aanleiding om te verwachten dat de positie van de vraag aan het begin in de vragenlijst en de vraagstelling een methodebreuk tot gevolg zouden hebben, aangezien beide niet zijn veranderd.

Er is wel een duidelijk weegeffect, dat wordt verklaard door de twee hoogste leeftijdsgroepen. Het percentage mensen dat zijn of haar gezondheid als (zeer) goed ervaart is hierdoor gedaald, vooral bij vrouwen en bij oudere mensen. Wellicht komt dit doordat

³⁾ Het betreft hier personen uit het eerste inkomenskwaantiel en uit het tweede vermogenskwaantiel. Het eerste vermogenskwaantiel bestaat vooral uit huiseigenaren die door hun hypotheekschuld een negatief vermogen hebben. Het tweede vermogenskwaantiel betreft vooral mensen met een huurhuis die weinig eigen vermogen hebben.

kinderen onder de 12 jaar de vragenlijst niet zelf invullen, dit doen hun ouders. Mogelijk verklaart deze proxy-invulling bij een in hoge mate subjectieve variabele als ervaren gezondheid de afwezigheid van een duidelijke wegingsbreuk bij 0–25-jarigen.

Diabetes

Voor diabetes is er, naast een klein wegeffect, alleen een duidelijk designeffect in de hoogste leeftijdscategorie. Er waren verschillende wijzigingen in de vragenlijst, zowel in de vraag voor respondenten tot 12 jaar als in de vraag voor respondenten van 12 jaar of ouder (zie Bijlage 4). Deze veranderingen bieden echter geen verklaring voor het gevonden verschil tussen de leeftijdsgroepen, omdat het designeffect van de middelste leeftijdsgroep vergelijkbaar is met dat van de jongste. Het in paragraaf 5.1 beschreven algemene prevalentie-effect biedt hier wel een verklaring.

Verkoudheid

Door de breuk voor de variabele verkoudheid (deze omvat verkoudheid, griep, keelontsteking en voorhoofdsholteontsteking) ligt het percentage bijna 8 lager dan op basis van het verleden mag worden verwacht. Deze afname kan mogelijk worden toegeschreven aan verschillende veranderingen in het GE onderzoek. Aan de vraagstelling zelf is niets veranderd, al is de introductie wel op enkele punten gewijzigd (zie bijbehorende tekst in paragraaf 3.2). Daarnaast is de variabele verplaatst van het tweede vragenlijstdeel naar het eerste en enige vragenlijstdeel in 2014 en is daarmee in de vragenlijst naar achteren verplaatst. De aandoeningen die binnen de variabele verkoudheid vallen zijn acute aandoeningen met een relatief kortdurend beloop. Dit betekent dat deze variabele relatief veranderlijk is en onderhevig aan omgevingsinvloeden. Er kan dan ook niet worden uitgesloten dat een reële daling voor deze variabele de designbreuk geheel of deels verklaard. Er zijn aanwijzingen dat dit het geval is, aangezien er in 2014 sprake was van een relatief milde griepgolf (Hooiveld, Donker & Schellevis, 2015). Om aan te tonen dat het cijfer van 2014 inderdaad een uitbijter is, zouden de methodebreukanalyses over enkele jaren moeten worden herhaald, zodat ook de cijfers van de jaren na 2014 in de analyses kunnen worden meegenomen.

Psychische ongezondheid

Voor psychische ongezondheid is er alleen een totaalbreuk bij vrouwen en in de leeftijd van 25 tot 65 jaar. De breuken bij deze groepen zijn een optelsom van design- en wegingsbreuk, waarbij de weging het grootste deel verklaart. De designbreuken waren relatief klein en dit valt te rijmen met het feit dat er geen grote veranderingen waren in de vragenlijst van 2014 ten opzichte van 2013. Wel was het zo dat de vragen in 2013 in het tweede blok van deel 2 zaten en in 2014 in het achttiende blok van de vragenlijst.

Minstens 1 beperking (OESO-indicator)

Afgezien van een algemene wegingsbreuk, is er bij 65-plussers een substantiële designbreuk die een verhoging van 5% van het aantal personen met minstens 1 beperking omvat. Tevens is er een kleinere tegengestelde designbreuk in de beide andere leeftijdsgroepen die een vermindering van 1,5 à 2% betekent.

De positie van deze vragen in de vragenlijst biedt hiervoor mogelijk een verklaring. Deze vragen zijn namelijk verplaatst van het eerste blok van deel 2 (2013) naar blok 14 (2014). Interessant daarbij is dat de twee voorgaande blokken in 2014 gaan over beperkingen. Blok 12 gaat over beperkingen in algemene dagelijkse levensverrichtingen (ADL); de meeste vragen worden alleen aan 55-plussers gesteld (alleen de eerste 4 vragen vanaf 12 jaar). Blok 13 gaat over beperkingen bij het uitvoeren van huishoudelijke activiteiten en wordt alleen aan 55-plussers voorgelegd. Dat 55-plussers al een flink aantal vragen over beperkingen

hebben beantwoord, heeft mogelijk de drempel om bij de OESO-vragen aan te geven dat er sprake is van een beperking verlaagd. Het resultaat hiervan zou dan zijn dat meer mensen aangeven dat zij 1 of meer beperkingen hebben.

Beperking in bewegen (OESO-indicator)

De designbreuk bij beperking in bewegen is eveneens het grootst in de leeftijdsgroep van 65 jaar of ouder. Dezelfde verklaring als bij minstens 1 beperking zou hier kunnen gelden. Voor beperking in bewegen is er echter ook sprake van een designbreuk voor de totale bevolking die hiermee niet is te verklaren. De vragen zelf zijn ongewijzigd gebleven, dus dit is eveneens geen verklaring. Een andere mogelijke verklaring heeft te maken met het verdwijnen van de CATI mode. Oudere personen responderen gemiddeld vaker via CATI. Via de telefoon is het makkelijker om de eigen beperkingen rooskleuriger te presenteren dan in een persoonlijk interview aan huis. Mogelijk heeft dit te maken met schaamte, maar wanneer er een interviewer aanwezig is heeft die ook de mogelijkheid de respondent te corrigeren op basis van hetgeen hij of zij bij die persoon ziet. Deze verklaring is echter niet heel waarschijnlijk, omdat bij het herontwerp in 2010 geen methodebreuk als gevolg van de wijziging van de mode van CAPI naar CAWI/CATI/CAPI werd gevonden. Het toevoegen van de CATI mode heeft destijds namelijk niet tot een afname geleid van het percentage personen met een beperking in bewegen.

Lichamelijke beperking (ADL-indicator)

Aangezien de bevindingen bij lichamelijke beperking hetzelfde patroon volgen als bij beperking in bewegen, kunnen dezelfde verklaringen worden gegeven. Wel is één vraag van de indicator, over het zich kunnen wassen (VolWas), anders geformuleerd (zie Bijlage 4). De frequentietabellen van deze variabele voor 2013 en 2014 laten zien dat er inderdaad een toename is van 2,5% (gewogen). Hiermee wordt de breuk van 4,5% niet volledig verklaard. Er is bij de 6 vragen die voorafgaan aan VolWas echter eveneens een toename in het percentage mensen dat 'Met grote moeite', 'Met hulp van anderen' of 'Zelfs niet met hulp van anderen' aankruist. Dit percentage varieert van 0,3 tot 1,6%. De laatste drie vragen van de indicator over het zich kunnen voortbewegen vertonen echter een daling van 0,8 tot 2,1 procent. Wanneer er meer mensen scoren op de ADL vragen, is het goed mogelijk dat er een breuk in de orde van grootte van de geobserveerde breuk ontstaat. Het is overigens opvallend dat de drie bewegingsitems die als enige een daling vertonen, direct na de nieuwe vraag over toiletbezoek worden gesteld. Het is moeilijk te zeggen of dit ook werkelijk aan de nieuwe vraag is toe te schrijven. De daling bij deze bewegingsitems is eveneens opvallend, omdat de breuk bij beperking in bewegen juist tot een toename leidt.

Contact met de huisarts en Aantal contacten met de huisarts

Een opvallende bevinding is dat het aandeel mensen dat contact heeft met de huisarts afneemt, terwijl het aantal contacten per persoon in de bevolking toeneemt. Het aandeel mensen met contact neemt vooral af bij mensen onder de 65 jaar, terwijl het aantal contacten vooral toeneemt boven de 25 jaar. Het is goed mogelijk dat deze breuken het resultaat zijn van de revisie van de vragen in dit blok (zie Bijlage 4). Om vast te stellen of iemand contact met de huisarts had, werden in 2013 twee vragen gesteld over het aantal contacten: eerst over de afgelopen 4 weken en bij geen contact over de afgelopen 12 maanden. Respondenten hadden dus twee kansen om aan te geven dat ze contact hadden met de huisarts en werden gedwongen om goed na te denken over hoe vaak dit dan was. In 2014 werd één gesloten vraag gesteld met drie antwoordopties over wanneer het laatste huisartscontact plaatsvond. Het is zeer aannemelijk dat deze wijziging heeft geleid tot een afname in het aantal personen dat contact heeft met de huisarts. Daarbij komt dat in 2014

niet langer expliciet wordt vermeld dat het gaat om contact met de huisarts in Nederland. Het is echter ook mogelijk dat er sprake is van een reële daling, doordat de milde griepgolf van 2014 minder mensen naar de huisarts heeft bewogen.

Dat het aantal contacten vervolgens toeneemt kan een artefact zijn van het gegeven dat mensen die minder vaak de huisarts bezoeken dit sneller over het hoofd zien. Het gevolg zou dan zijn dat mensen die relatief vaak de huisarts bezoeken worden gebruikt om het aantal contacten te berekenen, zodat een toename in het aantal contacten wordt geobserveerd. Daarbij komt dat het aantal contacten in 2013 werd afgeleid van de datum van de laatste twee contacten met de huisarts. Het vergeten van een contact leidt dan al snel tot een onderschatting van het aantal contacten. Overigens leidde de wijzigingen in 2010 niet tot een methodebreuk ten gevolge van de vragenlijst; de veranderingen waren destijds dan ook minder ingrijpend (Wong et al., 2011).

Contact met de specialist

De uitvraag voor het percentage personen dat contact had met de specialist is op dezelfde manier gewijzigd als die voor contact met de huisarts. Het is dan ook verrassend dat er voor deze variabele overall geen significante totaalbreuk is. Dit verschil kan niet worden verklaard door effecten van de vragenlijst, behalve dat door het inkorten van de voorafgaande blokken over de huisarts en ziekenhuis- en dagopname, dit blok kort na de vragen over contact met de huisarts volgt en er andere vragen aan voorafgaan. Een andere verklaring is dat mensen, doordat het hebben van contact met de specialist doorgaans een meer ingrijpend karakter heeft, deze contacten minder snel vergeten, zodat het percentage gelijk blijft.

Ziekenhuisopname

Voor ziekenhuisopname is dezelfde verklaring van toepassing als voor contact met de specialist. Een verschil is wel dat de introductietekst en de vraagstelling veel minder ingrijpend zijn gewijzigd.

Zelfmedicatie

Mogelijk kan de breuk bij zelfmedicatie (vooral bij vrouwen en 25-plussers) worden verklaard door wijzigingen in de vragenlijst. De introductie is in 2014 verwijderd en de uitleg dat medicijngebruik betrekking kan hebben op kruidenpreparaten (in 2014: kruidengeneesmiddelen) en vitamines is verplaatst naar de vraagtekst. Vóór de vraag naar zelfmedicatie zat in beide jaren een vraag over het gebruik van medicijnen op recept. Doordat de uitleg dat ook kruidenpreparaten en vitamines meegeteld mogen worden in 2013 niet direct vóór de zelfmedicatievraag stond, is het mogelijk dat mensen deze destijds niet hebben meegerekend bij het beantwoorden van de vraag. Dit kan hebben geleid tot de breuk van een toegenomen zelfmedicatie. Ook de mode heeft wellicht bijgedragen aan deze breuk; in 2010 bleek dat de verandering van de mode een significante methodebreuk tot gevolg had (Wong et al., 2011).

Mammografie

De significante totaalbreuk bij mammografie wordt verklaard door de designbreuk die optreedt bij vrouwen in de leeftijdsgroep van 25 tot 65 jaar. De daling van het aantal mammografieën als gevolg van deze breuk kan mogelijk worden verklaard door wijzigingen in de vragenlijst. Ten eerste is de verwijzing naar het bevolkingsonderzoek uit de vraag- en introductieteksten geschrapt, waardoor vrouwen wellicht zijn vergeten om het bevolkingsonderzoek mee te tellen. Daarnaast wordt niet langer gevraagd naar de frequentie van mammografieën in de afgelopen twee jaar, maar naar de laatste keer dat er een borstfoto werd gemaakt. Ook wordt de term mammografie in 2013 nog niet genoemd; wanneer

vrouwen de term niet kennen, vullen zij mogelijk in dat zij nooit een borstfoto hebben laten maken.

De gewijzigde mode kan eveneens een verklaring bieden voor de methodebreuk. Mogelijk is het voor een interviewer in een face-to-face situatie (CAPI) makkelijker de respondent erop te wijzen dat zij gezien haar leeftijd waarschijnlijk heeft deelgenomen aan het bevolkingsonderzoek dan aan de telefoon (CATI).

Roken

De vraagstelling van roken is niet gewijzigd, maar desondanks is er een methodebreuk als gevolg van de verandering van het design (vooral bij vrouwen en tot 65 jaar). Deze breuk leidt tot een toename van het cijfer, hetgeen haaks staat op de dalende trend in het percentage rokers die al jaren wordt waargenomen. Hoewel de vraagstelling niet is gewijzigd en de positie aan het einde van de vragenlijst ook niet, is wel het blok voorafgaand aan het blok Roken veranderd (zie Bijlage 3). In 2014 was dit het blok Ongevallen met vragen over het hebben gehad van een ongeluk in de afgelopen 12 maanden. Vóór 2013 stond vóór het blok Roken echter het blok Eigen bijdrage met vragen over de aanvullende zorgverzekering en de eigen bijdrage daarbij. Omdat verzekeringen wellicht wat beladen zijn in verband met roken, kan er sprake zijn geweest van enige onderrapportage. Een andere mogelijke verklaring is dat mensen via CATI makkelijker sociaal wenselijk antwoorden dat zij niet roken. Door het verdwijnen van de CATI mode hebben zij die mogelijkheid in 2014 niet meer, met als gevolg een toename van het percentage rokers. Deze verklaring is echter strijdig met de bevinding voor 2010 van Wong et al. (2011), die eveneens een significante mode-gerelateerde breuk vonden waarbij het percentage rokers toenam. Er was daarnaast destijds wel sprake van een niet-significant vragenlijsteffect, dat een daling inhield die niet kan worden verklaard door volgorde-effecten.

Zware drinker en Alcoholconsumptie

De breuk in het aantal zware drinkers onder de drinkers is een afname, terwijl het aantal glazen alcoholische drank dat drinkers nuttigen toeneemt. Dit kan worden verklaard doordat de vraag op basis waarvan wordt bepaald of iemand een drinker is, volledig is gewijzigd (zie Bijlage 4). Hierdoor worden licht alcoholische dranken, zoals alcoholarm bier, in 2014 niet langer als alcoholische consumptie werden beschouwd en niet meer worden meegenomen om te bepalen of iemand een drinker is. Dat betekent dat personen die in 2013 alleen licht alcoholische dranken nuttigden werden aangemerkt als drinker, maar in 2014 niet. Dit percentage lag in 2013 op 0,5% en kan dus slechts een deel van de breuk verklaren. Dit was echter niet de enige wijziging van de vraagstelling. Om vast te stellen of iemand een drinker is, werd er vóór 2014 voor diverse soorten alcoholische consumptie gevraagd deze in de afgelopen 12 maanden wel eens was genuttigd. Deze vrij gedetailleerde vraag is in 2014 vervangen door de vraag of men in de laatste 12 maanden wel eens alcoholhoudende drank heeft gedronken, met daarbij enkele voorbeelden. Een meer specifieke vraag zal in het algemeen meer positieve antwoorden opleveren, omdat mensen die weinig drinken eerder geneigd zijn te vergeten dat ze alcohol hebben geconsumeerd. Er worden dan minder drinkers geïdentificeerd, met als gevolg een hogere gemiddelde alcoholconsumptie per drinker. Daar komt bij dat er aan de vragen over week- en weekendconsumptie in de vraag naar het aantal glazen een antwoordoptie toegevoegd. Hierdoor is het maximaal aantal glazen dat kan worden aangekruist verhoogd van '11 of meer glazen' naar '16 of meer glazen'. Ook dit kan tot een hogere gemiddelde alcoholconsumptie per drinker hebben geleid. De breuk van de afname van het percentage zware drinkers is moeilijker te verklaren. Mogelijk speelt sociaal wenselijk antwoorden bij CATI minder een rol bij het rapporteren van het drankgebruik dan bij CAPI. Merk op dat dit effect tegenovergesteld is als bij roken; dit zou dan

komen doordat het voor iemand minder makkelijk is om te verbloemen dat hij of zij rookt. In het onderzoek van 2010 werd overigens geen methodebreuk gevonden voor het percentage zware drinkers. Dit suggereert dat deze vragen relatief bestendig zijn tegen wijzigingen in de mode en dat daarom de wijzigingen in de vragenlijst en de positie van de vragen in de vragenlijst verantwoordelijk kunnen zijn voor de breuk.

Ernstig overgewicht

De variabele ernstig overgewicht is opnieuw berekend volgens de criteria van 2014. De vraagstelling van deze variabele is niet veranderd en ook de positie vrij aan het einde van de vragenlijst is ongewijzigd gebleven. Het is daarom opvallend dat hier een methodebreuk optreedt, waarbij er vooral sprake is van een toename van het percentage vrouwen met overgewicht in met name de leeftijd van 25–65 jaar. De richting van deze breuk is in overeenstemming met het verdwijnen van de CATI mode. Het is bij CAPI minder makkelijk om een lager gewicht op te geven, omdat er een interviewer tegenover je zit. Die ziet immers wel ongeveer wat je gewicht is. Daar tegenover staat overigens de bevinding in het 2010-onderzoek dat er geen mode-gerelateerde breuk optrad na de invoering van CATI.

5.3 Beperkingen van het onderzoek

Dit onderzoek heeft enkele beperkingen. Ten eerste is er niet voor alle variabelen, waarover het Centraal Bureau voor de Statistiek (CBS) op basis van de Gezondheidsenquête publiceert in 2014, een methodebreukanalyse uitgevoerd (zie paragraaf 2.1). De mate waarin er voor een bepaalde variabele een breuk wordt gevonden, zegt niet per definitie iets over de mate waarin er voor een andere variabele sprake is van een breuk.

Bij het herontwerp in 2014 zijn verschillende aspecten van de gezondheidsenquête gelijktijdig veranderd. De benaderingsstrategie werd aangepast waarbij de CATI mode verdween, het aantal onderzoeksdelen werd teruggebracht van twee tot één, het weegmodel werd uitgebreid met twee variabelen, de vragenlijst werd ingekort en aangepast, waarbij de positie van vragen in de vragenlijst veranderde en sommige vragen werden gewijzigd. De eerste drie veranderingen waren generiek voor alle variabelen, de laatste twee betroffen verschillende variabelen in verschillende mate. Doordat verschillende aspecten van een vraag tegelijkertijd veranderden, kan een waargenomen methodebreuk niet met zekerheid aan een bepaalde verandering worden toegeschreven. Maar ook wanneer er geen methodebreuk wordt gevonden, kan niet met zekerheid worden geconcludeerd dat het herontwerp geen invloed heeft gehad: een reële verandering in de cijfers kan zijn tenietgedaan door het herontwerp. Tot slot kan niet worden uitgesloten dat een breuk feitelijk een reële veranderingen van de cijfers betreft. Dit is ook de reden dat wordt aangeraden om methodebreukanalyses enkele jaren na een herontwerp te herhalen, zodat kan worden uitgesloten dat de breuken te wijten zijn aan toevallige fluctuaties (uitbijters) van de cijfers.

Door dubbel te draaien kan het effect van afzonderlijke wijzigingen, bijvoorbeeld zoals in 2010 van de mode of van de vragenlijst, uit elkaar getrokken kunnen worden. Bij dubbeldraaien wordt er een experimentele conditie toegevoegd naast het reguliere vernieuwde onderzoek, waarbij de oude vragenlijst volgens de nieuwe benaderingsstrategie wordt toegevoegd, of de nieuwe vragenlijst volgens de oude benaderingsstrategie. De individuele effecten kunnen dan onafhankelijk van elkaar worden geschat. Voor de GE is er in 2014 niet dubbel gedraaid. Het probleem was dat er veel verschillende wijzigingen gelijktijdig moesten worden doorgevoerd: de mode, de opzet van de vragenlijst, de positie van vragen in de vragenlijst, de vragen zelf en het weegmodel. Uit praktisch oogpunt is ervoor gekozen om al deze wijzigingen in één keer door te voeren, zodat er maar één keer sprake zou zijn van

een breuk. Bij één voor één doorvoeren van de veranderingen zou enkele jaren achtereenvolgend een breuk ontstaan en bovendien de stabiliteit van de reeks in het geding zijn. Het alternatief van dubbeldraaien met meerdere experimentele condities (één voor elke verandering) werd eveneens niet haalbaar geacht, vanwege budgettaire beperkingen en de praktische problemen die dit zou opleveren bij het verzamelen van de gegevens.

Ondanks deze beperkingen wordt het uitvoeren van een methodebreukanalyse zinvol geacht. Het verschaft namelijk meer inzicht in de cijfers. En ook wanneer het niet mogelijk is om de oorzaak van een breuk met zekerheid te duiden, kunnen er wel degelijk verklaringen worden geformuleerd waarvan het aannemelijk is dat ze hebben bijgedragen aan het ontstaan van de breuk. In dit rapport is een poging gedaan om dergelijke verklaringen te formuleren.

6. Conclusies

In deze nota is onderzocht in hoeverre het herontwerp van de Gezondheidsenquête in 2014 heeft geleid tot een methodebreuk in de cijfers van 2014. Op basis hiervan kunnen de volgende conclusies worden getrokken:

1. Er is sprake van een duidelijke methodebreuk voor een meerderheid van de geselecteerde variabelen (11 van de 17). Voor de volgende variabelen trad er een methodebreuk op: verkoudheid, beperking in bewegen, lichamelijke beperking, contact met de huisarts, aantal contacten met de huisarts, zelfmedicatie, mammografie, roken, zware drinker, alcoholconsumptie en ernstig overgewicht.
2. De meeste breuken hebben tot gevolg dat de cijfers een minder goede gezondheid, hogere medische consumptie en een ongezonder leefstijl weergeven. Dit geldt echter niet voor verkoudheid, contact met de huisarts en zware drinker. Hier leidden de breuken tot een betere gezondheid, lagere medische consumptie en een minder ongezonde leefstijl.
3. De totale methodebreuk die ontstaat, kan worden opgesplitst in een designbreuk en een wegingsbreuk, die kunnen helpen de totaalbreuk te duiden. De designbreuken zijn het gevolg van het wijzigen van de benaderingsstrategie, de mode, de structuur van de vragenlijst, de positie van vragen in de vragenlijst en de vraagstelling. De wegingsbreuken zijn het gevolg van de uitbreiding van het weegmodel. De geobserveerde designbreuken zijn voor de meeste variabelen groter dan de wegingsbreuken. Twee variabelen vormen een uitzondering, hier zijn de wegingsbreuken groter dan de designbreuken. Dit zijn ervaren gezondheid en minstens 1 beperking volgens de OESO-indicator.
4. Het uitbreiden van het weegmodel met de variabelen inkomen en vermogen leidt ertoe dat de cijfers een slechtere gezondheid, een hogere zorgconsumptie en een minder gezonde leefstijl laten zien. Dit effect is goed te verklaren, omdat vooral groepen met een slechte gezondheid in dit nieuwe weegmodel een zwaarder gewicht krijgen. Dit zwaardere gewicht is toegekend, omdat de respons in deze groepen relatief laag is.
5. Er zijn geen duidelijke aanwijzingen gevonden dat variabelen uit het voormalige deel 2 van de vragenlijst een grotere methodebreuk hebben. Voor sommige variabelen die in 2013 in vragenlijstdeel 2 zaten wordt een grote breuk gevonden (bijv. verkoudheid), maar voor andere variabelen is er juist geen breuk (bijv. psychische ongezondheid). Hetzelfde geldt voor variabelen uit vragenlijstdeel 1, die soms wel (bijv. zelfmedicatie) en soms geen breuk vertonen (bijv. ziekenhuisopname).

6. Hoewel het mogelijk is om te spreken van een designbreuk, kan de precieze oorzaak ervan niet worden geïdentificeerd. Dit komt doordat aan de designbreuk verschillende wijzigingen in de Gezondheidsenquête ten grondslag liggen, namelijk wijzigingen in de benaderingsstrategie, in de mode, in de structuur van de vragenlijst, in de positie van vragen in de vragenlijst, en in de vragen en antwoordopties. De effecten van deze veranderingen kunnen, nog los van de veranderingen die in werkelijkheid optreden in de gezondheidstoestand van de bevolking, niet van elkaar worden losgekoppeld.
7. Voor een aantal variabelen wordt geen methodebreuk gevonden. Dit is een aanwijzing dat deze variabelen relatief resistent zijn tegen veranderingen die in het onderzoek van de GE worden doorgevoerd. Variabelen waarvoor de cijfers stabiel blijven na veranderingen in het design en in de vragenlijst en dergelijke, zijn ervaren gezondheid, psychische gezondheid, contact met de specialist en ziekenhuisopname. Voor deze variabelen werd in geen van de (sub)groepen een duidelijke designbreuk gevonden. Variabelen waarvoor de cijfers stabiel bleven na uitbreiding van het weegmodel zijn zelfmedicatie, mammografie en alcoholconsumptie. Voor deze variabelen werd geen duidelijke wegingsbreuk gevonden.
8. Het ligt voor de hand om aan te nemen dat de methodebreuken die in dit onderzoek gevonden zijn, ook zullen optreden bij variabelen uit de Gezondheidsenquête die niet in methodebreukanalyses werden meegenomen. Het is echter niet mogelijk om te voorspellen voor welke van deze variabelen er sprake zal zijn van een methodebreuk en wat de richting en omvang ervan zal zijn. Er is dan ook voorzichtigheid geboden, wanneer cijfers van GE uit 2014 (of later) en van GE 2013 (of eerder) worden vergeleken of samen in één figuur of tabel worden gepresenteerd. Het is raadzaam om hierbij steeds te vermelden dat er sprake kan zijn van een methodebreuk en om met deze breuk rekening te houden bij het interpreteren van eventuele verschillen.
9. De in dit rapport beschreven methodebreukanalyses zijn uitgevoerd met slechts één meetpunt na de breuk. Dit brengt een onnauwkeurigheid met zich mee, omdat er sprake kan zijn van toevallige schommelingen in het geobserveerde fenomeen of zelfs van uitbijters. Hierbij valt te denken aan de breuk bij de variabele verkoudheid, die mogelijk zijn oorsprong heeft in de milde griepgolf van 2014. Het is daarom zinvol om deze methodebreukanalyses over een aantal jaren nog eens te herhalen, wanneer er meerdere meetpunten na de breuk beschikbaar zijn. Uitbijters worden dan uitgemiddeld en het effect van realistische veranderingen kan worden losgekoppeld van veranderingen als gevolg van het herontwerp. Er kan dan een betere duiding van de in dit rapport gevonden methodebreuken worden gerealiseerd.

Literatuur

Boonstra, H.J. (2015). *Weging Gezondheidsenquête 2014*. Heerlen: Centraal Bureau voor de Statistiek.

Brakel, J.A. van den & Roels, J. (2010). Intervention analysis with state-space models to estimate discontinuities due to a survey redesign. *Annals of Applied Statistics*, 4, 1105–1138.

Brakel, J.A. van den, Smith, P.A. & Compton, S. (2008). Quality procedures for survey transitions, experiments, time series and discontinuities. *Journal for Survey Research Methods*, 2, 123–141.

Bruggink, J.-W. & Van Herten, M. (2014). *Weging Gezondheidsenquête, vanaf 2010*. Heerlen: Centraal Bureau voor de Statistiek.

Doornik, J.A. (2009). *An object-oriented Matrix Programming Language Ox 6*. London: Timberlake Consultants Press.

Durbin, J. & Koopman, S.J. (2001). *Time Series Analysis by State Space Methods*. Oxford: Oxford University Press. MR1856951.

Goris, G. (2013a). *Onderzoeksdesign GEZO 2013 1.3*. Heerlen: Centraal Bureau voor de Statistiek.

Goris, G. (2013b). *Onderzoeksdesign Gezondheidsenquête 2014 1.0*. Heerlen: Centraal Bureau voor de Statistiek.

Harvey, A.C. & Durbin, J. (1986). The effect of seat belt legislation on British road casualties: A case study in structural time series modelling. *Journal of the Royal Statistical Society Series A*, 149, 187–227.

Hooiveld, M., Donker, G.A. & Schellevis, F.G. (2015). *Wekelijkse surveillance cijfers*. NIVEL Zorgregistraties eerste lijn. Utrecht: NIVEL (www.nivel.nl/surveillance, geraadpleegd op 2 april 2015).

Koopman, S.J., Shephard, N. & Doornik, J.A. (2008). *Structural Time Series Analyser, Modeller en Predictor – Stamp 8.2*. London: Timberlake Consultants Press.

Wong, F.Y., Roels, J. & Van den Brakel, J. (2011a). *Analyse methodebreuken Gezondheidsenquête bij het Herontwerp Persoonsenquête*. Heerlen: Centraal Bureau voor de Statistiek.

Wong, F.Y., Roels, J. & Van den Brakel, J. (2011b). *Schattingen voor de Gezondheidsenquête 2010 op basis van voorspellingen met behulp van structurele tijdreeksmodellen (versie 3.4)*. Heerlen: Centraal Bureau voor de Statistiek.

Bijlagen

Bijlage 1	Beschrijving variabelen
Bijlage 2	Verschillen met StatLine
Bijlage 3	Opbouw GE vragenlijst 2013 en 2014
Bijlage 4	Verschillen in de GE vragenlijst van 2014 ten opzichte van 2013
Bijlage 5	Cijfers GE 2014
Bijlage 6	Grafieken methodebreuken

Verklaring van tekens

.	Gegevens ontbreken
*	Voorlopig cijfer
**	Nader voorlopig cijfer
x	Geheim
–	Nihil
–	(Indien voorkomend tussen twee getallen) tot en met
0 (0,0)	Het getal is kleiner dan de helft van de gekozen eenheid
Niets (blank)	Een cijfer kan op logische gronden niet voorkomen
2014–2015	2014 tot en met 2015
2014/2015	Het gemiddelde over de jaren 2014 tot en met 2015
2014/'15	Oogstjaar, boekjaar, schooljaar enz., beginnend in 2014 en eindigend in 2015
2012/'13–2014/'15	Oogstjaar, boekjaar, enz., 2012/'13 tot en met 2014/'15

In geval van afronding kan het voorkomen dat het weergegeven totaal niet overeenstemt met de som van de getallen.

Colofon

Uitgever
Centraal Bureau voor de Statistiek
Henri Faasdreef 312, 2492 JP Den Haag
www.cbs.nl

Prepress
Studio BCO, Den Haag

Ontwerp
Edenspiekermann

Inlichtingen
Tel. 088 570 70 70, fax 070 337 59 94
Via contactformulier: www.cbs.nl/infoservice

© Centraal Bureau voor de Statistiek, Den Haag/Heerlen/Bonaire, 2015.
Verveelvoudigen is toegestaan, mits het CBS als bron wordt vermeld.