

RIVM rapport 260854009/2005

**Methodologie Lokale en Nationale Monitor
Volksgezondheid**
Gevolgen voor vergelijkbaarheid van gegevens

Brink CL van den, Viet AL, Boshuizen HC,
Ameijden EJC van¹, Droomers M

Contact:

Carolien van den Brink, Centrum voor Preventie- en Zorgonderzoek, RIVM
carolien.van.den.brink@rivm.nl

¹ GG&GD Utrecht, Utrecht

Dit onderzoek werd verricht in opdracht en ten laste van het Ministerie van Volksgezondheid, Welzijn en Sport in het kader van project V/260854/01/AA, Nationale en Lokale Monitor Volksgezondheid.

Rapport in het kort

Methodologie Lokale en Nationale Monitor Volksgezondheid

Verschillen in methoden van gegevens verzamelen blijken te leiden tot verschillen in uitkomstmaten. Deze bevinding, beschreven in dit rapport, vormt de basis voor een advies voor uniforme gegevensverzameling door GGD'en. In het project "Lokale en Nationale Monitor Volksgezondheid" worden lokaal verzamelde gegevens samengevoegd om landelijke referentiecijfers over gezondheid en leefstijlfactoren te verkrijgen. GGD'en gebruiken echter verschillende methoden van gegevensverzameling en de vraag is in hoeverre de verzamelde gegevens vergelijkbaar zijn en samengevoegd mogen worden. Met behulp van literatuuronderzoek, data-analyses en overleg met deskundigen is onderzocht wat het effect is van verschillen in de steekproeftrekking, non-respons, seizoen waarin gemeten is en de manier van enquêteren op de uitkomsten. Tevens is bestudeerd hoe met deze verschillen omgegaan kan worden. Verschillen in de grootte van de steekproef hebben geen invloed op de vergelijkbaarheid van uitkomsten. Non-respons beïnvloedt de uitkomsten alleen als bepaalde groepen van de bevolking vaker meedoen dan andere. In dit geval kan de representativiteit van de respondenten voor de populatie vergroot worden door te wegen naar factoren die samenhangen met zowel respons als de uitkomstmaten, zoals leeftijd, geslacht, burgerlijke staat en urbanisatiegraad. Voor zowel seizoen als manier van enquêteren geldt dat gegevens niet vergelijkbaar zijn wanneer de methode verschilt. In de zomer werd namelijk een betere gezondheid en een gezondere leefstijl gerapporteerd dan in de winter. Verder rapporteerden mensen die een enquête via internet invulden een ongezondere leefstijl dan mensen die dat schriftelijk deden.

Trefwoorden: manier van enquêteren; methode; non-respons; representativiteit; seizoen; steekproef; weegmodel

Abstract

Methodology of the Public Health Monitor; comparability of data

In the project “Local and National Public Health Monitor” national reference data are generated by merging regional data, that are collected by the Community Health Services. The aim of this report is to investigate whether differences in the methods of data collection affect the comparability of regional data and the merging of regional data into national reference data. The effects of differences in the sampling methods, non-response, seasonal variation, and mode of data collection on outcome measures are investigated by a literature review, data analyses, and by consulting experts. It appears that differences in the size or fraction of the sample do not influence the comparability of collected data. When a stratified sample is drawn, weights can be assigned to the respondents in order to achieve representativeness of the population. Bias of outcome measures caused by non-response can be largely adjusted for by weighting the data, using a model consisting of variables like age, sex, and marital status. These variables are known to be associated with response as well as with outcome measures. Both seasonal variation and the mode of data collection (mailed questionnaire/interview/internet) are found to influence health indicators and life style factors. Based on these findings it is concluded that differences in the methods of data collection may result in differences in outcome measures.

Keywords: mode of data collection; methodology; non-response; representativeness; seasonal variation; weighting model

Voorwoord

Het project “Lokale en Nationale Monitor Volksgezondheid” (Monitor VGZ) is in 2002 van start gegaan. Het project is een samenwerking tussen GGD'en, het RIVM en GGD Nederland. Met dit project wordt een basis gelegd voor continue monitoring van de belangrijkste indicatoren van de gezondheid op lokaal en landelijk niveau. Het project is erop gericht de lokale gegevensverzamelingen die plaatsvinden bij GGD'en, op elkaar af te stemmen zodat na samenvoeging van de gegevens een nationaal beeld kan worden verkregen. Een eerste stap daarin is het ontwikkelen van standaardvraagstellingen geweest. Daarnaast speelt ook de methode van gegevensverzameling een rol bij de vergelijkbaarheid van gegevens. In dit rapport wordt beschreven in hoeverre gegevens met elkaar vergelijkbaar zijn en samengevoegd kunnen worden als er verschillen zijn tussen GGD'en in de methode waarmee die gegevens verzameld zijn. Aan de hand van de bevindingen zal een lijst opgesteld worden van methodologische informatie die GGD'en aan moeten leveren wanneer hun lokale gegevens ingevoerd worden in de landelijke database. Daarnaast vormt dit rapport de basis voor een aanbeveling voor een zo uniform mogelijke gegevensverzameling voor alle GGD'en.

Inhoud

Samenvatting	9
1. Monitor Volksgezondheid	11
2. Inleiding	13
3. Steekproefkader en non-respons	15
3.1 Inleiding	15
3.2 Steekproeftrekking	15
3.2.1 <i>Steekproefomvang</i>	15
3.2.2 <i>Steekproeffractie</i>	16
3.2.3 <i>Stratificatie</i>	16
3.2.4 <i>Leeftijdsgrenzen</i>	16
3.2.5 <i>Uitsluiten van groepen</i>	17
3.2.6 <i>Conclusie en aanbevelingen</i>	17
3.3 Non-respons	17
3.3.1 <i>Theorie</i>	17
3.3.2 <i>Contactpogingen</i>	17
3.3.3 <i>Verhoging respons</i>	18
3.3.4 <i>Conclusie en aanbevelingen</i>	18
3.4 Weging	18
3.4.1 <i>Achtergrond</i>	18
3.4.2 <i>Weegmodel</i>	19
3.4.3 <i>Koppeling gegevens</i>	20
3.4.4 <i>Conclusie en aanbevelingen</i>	20
3.5 Samenvoegen of vergelijken van gegevens	20
3.6 Conclusie	21
4. Seizoensvariatie	23
4.1 Inleiding	23
4.2 Situatieschets	23
4.3 Methode	23
4.3.1 <i>Literatuuronderzoek</i>	23
4.3.2 <i>Data-analyse op bestand GG&GD Utrecht</i>	24
4.4 Resultaten	27
4.4.1 <i>Uitkomsten literatuuronderzoek</i>	27
4.4.2 <i>Uitkomsten data-analyse GG&GD</i>	28
4.5 Gevolgen voor vergelijkbaarheid	33
4.6 Conclusie en aanbevelingen	34
5. Manier van enquêteren	35
5.1 Inleiding	35
5.2 Situatieschets	35
5.3 Methode	35
5.3.1 <i>Literatuuronderzoek</i>	36
5.3.2 <i>Data-analyse op bestand GGD'en regio Limburg</i>	36

5.4	Resultaten	36
5.4.1	<i>Uitkomsten literatuuronderzoek</i>	36
5.4.2	<i>Uitkomsten analyses op bestand GGD'en regio Limburg</i>	38
	Gevolgen voor vergelijkbaarheid	41
5.5	Conclusie en aanbeveling	42
6.	Conclusie	43
	Dankwoord	44
	Literatuur	45
	Bijlage 1. Gebruikte vraagstellingen GG&GD Utrecht; 1996-1998	51

Samenvatting

Achtergrond en inleiding

GGD'en verzamelen periodiek (gezondheids)gegevens om de lokale gezondheidstoestand en daarmee samenhangende factoren in kaart te brengen. Het project "Lokale en Nationale Monitor Volksgezondheid" beoogt standaardisatie van de lokale gegevensverzamelingen zodat na aggregatie een landelijk representatief beeld wordt verkregen. De manier waarop de gegevens verzameld worden bepaalt of de gegevens vergelijkbaar zijn en dus samengevoegd kunnen worden. Te denken valt daarbij aan het steekproefkader, non-respons, het seizoen waarin gemeten wordt en de manier van enquêteren. Per onderwerp wordt besproken wat het effect is op uitkomstmaten, en dus op de vergelijkbaarheid, en hoe met verschillen omgegaan kan worden.

Steekproefkader en non-respons

Met behulp van literatuuronderzoek en gesprekken met deskundigen van het Centraal Bureau voor de Statistiek (CBS) is bekeken wat het effect is van verschillen in de steekproeftrekking en non-respons op de vergelijkbaarheid van resultaten. Aspecten van de steekproeftrekking die verschillen tussen GGD'en, zijn de omvang, de fractie, en stratificatie naar bepaalde groepen. De eerste twee hebben geen effect op de uitkomsten, hooguit op de nauwkeurigheid. Wanneer dankzij of ondanks het stratificeren de steekproef niet representatief is voor de totale populatie, kunnen de uitkomsten gewogen worden naar de factoren waarnaar gestratificeerd is.

Selectieve non-respons vermindert de representativiteit van de respondenten voor de totale populatie. Door de uitkomsten te wegen naar kenmerken waarin de respondenten verschillen van de non-respondenten, kan hiervoor (gedeeltelijk) gecorrigeerd worden. Voor het weegmodel worden in navolging van de POLS-enquête van het CBS leeftijd, geslacht, burgerlijke staat en urbanisatiegraad gebruikt. GGD'en leveren deze gegevens aan voor hun respondenten, evenals een weegfactor waarmee zij regionale representatieve cijfers berekenen.

Seizoensvariatie

Aan de hand van data van de GG&GD Utrecht en literatuuronderzoek is bekeken of er seizoensvariatie optreedt in verschillende gezondheidsindicatoren en leefstijlfactoren. In Utrecht zijn er maandelijks enquêtes afgenomen, die we gegroepeerd hebben in kwartalen. Volgens onze data-analyse leken ervaren gezondheid, psychische problemen en ontsteking van de herten samen te hangen met seizoen, evenals lichamelijke activiteit. Uit het literatuuronderzoek bleek dat verschillende gezondheidsindicatoren (depressie, body mass index, incidentie beroerte, sterfte) samenhangen met het seizoen waarin is gemeten. In de wintermaanden werd een slechtere gezondheid gerapporteerd. Verder was het niveau van lichamelijke activiteit en van de consumptie van fruit en groente hoger in de zomermaanden ten opzichte van de wintermaanden.

Manier van enquêteren

We hebben literatuuronderzoek uitgevoerd naar het effect van de manier van enquêteren op gezondheidsindicatoren en leefstijlfactoren. Bij het invullen van een schriftelijke vragenlijst bleken mensen minder geneigd tot sociaal wenselijke antwoorden dan tijdens een interview. Bij het gebruik van interviews werden een betere gezondheidstoestand en kwaliteit van leven, en minder depressie gerapporteerd. Ook werd er een hoger niveau van lichamelijke activiteit gerapporteerd. Voor roken en alcoholconsumptie waren de studies niet consistent. Hoewel er nog weinig bekend is in de literatuur over het invullen van vragenlijsten via internet, leek deze methode qua uitkomstmaten niet te verschillen van schriftelijke vragenlijsten. Aanvullende analyses van data van de GGD'en Limburg, waarbij een deel van de respondenten via internet reageerde, lieten echter wel verschillen zien voor rapportage van leefstijlfactoren. Via internet is niet de totale populatie te bereiken.

Conclusie

Verschillen in methoden van gegevensverzameling kunnen leiden tot verschillen in uitkomstmaten. GGD'en dienen te vermelden in welk seizoen en op welke manier (schriftelijk, mondeling, internet) geënuquêteerd is. Alleen uniforme gegevensverzamelingen kunnen samengevoegd of vergeleken worden.

Voor (verschillen in) selectieve non-respons kan gecorrigeerd worden door te wegen naar leeftijd, geslacht, burgerlijke staat en urbanisatiegraad. GGD'en leveren deze gegevens aan voor de respondenten, zodat landelijke representatieve cijfers bepaald kunnen worden. Voor de presentatie van regionale representatieve cijfers leveren de GGD'en hun eigen weegfactor aan.

1. Monitor Volksgezondheid

In 2002 is het project “Lokale en Nationale Monitor Volksgezondheid” (Monitor VGZ) van start gegaan. Deze monitor levert een bijdrage aan de afstemming van informatievoorziening over de gezondheidstoestand en daarmee samenhangende factoren, op lokaal en landelijk niveau. Om dit doel te bereiken wordt de bij de GGD'en, het RIVM en GGD-NL aanwezige kennis en expertise gebundeld.

De opzet van het project is lokale gegevensverzamelingen, welke plaatsvinden bij GGD'en, op dusdanige wijze op elkaar af te stemmen dat na samenvoeging van gegevens een nationaal beeld wordt verkregen. Daarmee is het onderling vergelijken van gegevens uit diverse GGD-regio's en het vergelijken van lokale of regionale gegevens met landelijke referentiecijfers ook mogelijk. De gegevens worden samengevoegd in een landelijke database waarin alle data worden opgeslagen. De database is toegankelijk via de website <http://www.monitorgezondheid.nl>. Onderzoekers kunnen op deze website ook landelijke referentiecijfers bekijken en vergelijkingen maken met gegevens uit andere GGD-regio's.

2. Inleiding

Voor de vergelijkbaarheid van gegevens is allereerst de vergelijkbaarheid van de vraagstelling belangrijk. Hiertoe zijn (worden) standaardvraagstellingen ontwikkeld die door alle GGD'en gebruikt kunnen worden. Het al dan niet vergelijkbaar zijn van gegevens wordt echter niet alleen bepaald door de gebruikte vraagstelling, maar ook door de manier waarop de gegevens verzameld worden. Dit bleek o.a. uit studies waarin surveys uit verschillende landen vergeleken werden.^(1;2) Ook bij de vergelijking tussen gebieden of plaatsen binnen een bepaald land zullen deze methodologische aspecten een rol spelen.

Omdat de gegevens verzameld door de GGD'en in Nederland in de toekomst samengevoegd zullen worden om een landelijk referentiekader te creëren en om regionale verschillen te bekijken, wordt onderzocht welke aspecten van de gegevensverzameling van invloed zijn op prevalentieschattingen en dus op de vergelijkbaarheid van gegevens.

Mogelijke verschillen tussen GGD'en kunnen optreden door onder andere verschillen in de steekproeftrekking, de al dan niet selectieve (non-) respons en de methode van gegevensverzameling (manier van enquêteren, seizoen). In dit rapport zal op de invloed van deze aspecten nader ingegaan worden. Allereerst wordt aangegeven wat het effect is van verschillen in methodologie op uitkomstvariabelen. Vervolgens wordt beschreven hoe met verschillen in de gegevensverzameling omgegaan kan worden. Daarbij richten we ons op het gebruik van de gegevens voor het berekenen van landelijke referentiecijfers en de vergelijking tussen verschillende GGD-regio's. Het accent van dit rapport ligt op het monitoren van de gezondheidssituatie van volwassenen, voor de jeugdmonitor wordt een aanvullend stuk geschreven.

Uiteindelijk zullen de uitkomsten van dit rapport leiden tot een lijst met gegevens (meta-informatie) over de methode van gegevensverzameling bij de verschillende GGD'en. Deze informatie wordt gebruikt bij het samenvoegen en vergelijken van resultaten.

Daarnaast zal dit rapport de basis vormen voor een aanbeveling voor een zo uniform mogelijke gegevensverzameling voor alle GGD'en.

In dit rapport zal per aspect van de gegevensverzameling allereerst de situatie geschetst worden van de stand van zaken bij de GGD'en in Nederland. Vervolgens wordt beschreven hoe het effect van deze aspecten onderzocht is en worden de uitkomsten, conclusies en aanbevelingen vermeld.

3. Steekproefkader en non-respons

3.1 Inleiding

Iedere gemeente in Nederland is verplicht om elke vier jaar de lokale gezondheidstoestand in kaart te brengen. Dit wordt over het algemeen uitgevoerd door GGD'en. Elke GGD in Nederland bestrijkt een regio of een grote stad. Voor het bepalen van de gezondheidstoestand worden gezondheidsenquêtes georganiseerd. Hiervoor wordt een random steekproef getrokken uit de Gemeentelijke Basis Administratie (GBA). De mensen uit die steekproef worden benaderd om een vragenlijst in te vullen en terug te sturen, of om naar de GGD te komen voor onderzoek. Dit gebeurt niet altijd elke keer voor dezelfde groep. De ene keer betreft het volwassenen, terwijl bijvoorbeeld 4 jaar later de jongeren benaderd worden voor het onderzoek. De gegevens worden anoniem verwerkt.

Twee aspecten van de gegevensverzameling zijn van belang voor de representativiteit van de respondenten voor de bevolking van de gemeente waartoe ze behoren. Dit zijn de manier waarop de steekproef getrokken wordt en non-respons. Belangrijke aspecten van de steekproeftrekking zijn de omvang, de fractie (percentage van totale populatie), en stratificatie (indelen van populatie in bepaalde groepen). Tussen GGD'en zijn er verschillen in deze aspecten van steekproeftrekking, evenals in de non-respons.

Dit hoofdstuk beschrijft hoe (verschillen in) deze aspecten van invloed kunnen zijn op de representativiteit van gegevens voor een GGD-regio of het hele land en op de vergelijkbaarheid van gegevens tussen regio's of met landelijke referentiecijfers. Tevens wordt bediscussieerd op welke manier met deze aspecten rekening dient te worden gehouden.

3.2 Steekproeftrekking

3.2.1 Steekproefomvang

Voor het bepalen van de steekproefomvang bestaan veel formules (powerberekening). Deze formules zijn echter vaak bedoeld voor experimentele interventiestudies. Onderdeel van deze formules zijn dan het te verwachten effect van de interventie, standaarddeviatie van de uitkomstvariabelen en gewenste power en significantie. Ten aanzien van deze parameters worden assumpties gemaakt.

Deze formules zijn echter niet goed te gebruiken voor de gezondheidsenquêtes van de GGD'en, aangezien GGD'en in eerste instantie geïnteresseerd zijn in het zo goed mogelijk in kaart brengen van de lokale gezondheidstoestand en niet in effecten van interventies. Verder hebben we bij de gezondheidsenquêtes te maken met allerlei verschillende uitkomstvariabelen, hetgeen het creëren van een sluitende formule bemoeilijkt. Als GGD'en specifiek iets willen weten, zouden ze zelf op basis van informatie over dat onderwerp een berekening kunnen uitvoeren. In de praktijk zullen echter statistische overwegingen een minder grote rol spelen dan de logistieke en financiële overwegingen.

Verschillen in grootte van steekproeven tussen de GGD'en zullen overigens niet snel leiden tot andere uitkomsten, hoewel de nauwkeurigheid van de uitkomst wel met de grootte van de steekproef samenhangt. Voor verschillen in grootte hoeft dus niet gecorrigeerd te worden.

3.2.2 Steekproeffractie

Sommige GGD'en trekken een steekproef van een bepaald percentage (bijvoorbeeld 3%) van de bevolking uit iedere gemeente, terwijl andere GGD'en werken met een vaste omvang van de steekproef, bijvoorbeeld 500 personen per gemeente. Als de steekproeffractie in alle gemeenten even groot is, kunnen de gegevens van de verschillende gemeenten bij elkaar gevoegd worden voor het creëren van een representatief beeld van de regio. Indien de fracties tussen gemeenten verschillen, zullen voor een representatief beeld van de regio de kleinere fracties zwaarder moeten meewegen dan de grote.

3.2.3 Stratificatie

Steekproeven worden vaak random getrokken. Dit betekent dat elke inwoner van de GGD-regio een even grote kans heeft om benaderd te worden voor de gezondheidsenquête. Er wordt vooraf soms wel gestratificeerd. Dit betekent dat de populatie in bepaalde groepen wordt ingedeeld, bijvoorbeeld naar leeftijd, etniciteit of wijk. Reden hiervoor kan zijn dat een GGD specifiek geïnteresseerd is in de situatie van een bepaalde groep. Een andere reden kan zijn dat bepaalde groepen slechts een klein deel van de populatie uitmaken en dat een random steekproef daardoor geen nauwkeurige uitspraken oplevert over die groep of dat bekend is dat de verwachte respons in bepaalde groepen laag is. Om toch uitspraken over deze specifieke groepen te kunnen doen, wordt dan in die groepen de steekproef opgehoogd. We spreken dan van een disproportioneel gestratificeerde steekproef.

Voor het verkrijgen van een representatief beeld van de bevolking in een gemeente, een regio of het hele land, moet rekening gehouden worden met zulk een ophoging. De groepen die relatief vaker zijn benaderd, moeten dan minder zwaar meetellen.

3.2.4 Leeftijdsgrenzen

Er wordt onderscheid gemaakt tussen jongeren- en volwassenenenquêtes. De leeftijdsgrenzen voor deze enquêtes variëren per GGD. Voor de volwassenenenquête varieert de ondergrens tussen de 16 en 20 jaar en de bovengrens tussen 54 en 75 jaar. Gegevens van enquêtes die verschillende leeftijdsgrenzen hanteren kunnen niet direct met elkaar vergeleken worden. Bepaalde leeftijdsgroepen moeten uitgesloten worden voor de vergelijking, en over andere leeftijdsgroepen zal informatie ontbreken. De resultaten van steekproeven zullen het best vergelijkbaar zijn als GGD'en gelijke leeftijdsgrenzen hanteren. De richtlijn voor de volwassenenenquête is 19 t/m 64 jaar, en voor de jongeren 0 t/m 18.

3.2.5 Uitsluiten van groepen

Bij sommige GGD'en worden mensen die in verpleeghuizen of inrichtingen wonen uitgesloten van deelname. Redenen hiervoor zijn vaak praktische. Het gaat vaak om kleine aantallen, hetgeen de uitkomsten niet veel zal beïnvloeden.

3.2.6 Conclusie en aanbevelingen

Samenvattend kunnen we zeggen dat verschillen in de steekproefomvang of –fractie de vergelijkbaarheid van gegevens niet zullen beïnvloeden. Voor het creëren van een representatief beeld van bijvoorbeeld een regio moet wel gecorrigeerd worden voor de grootte van de steekproef in de gemeenten. Indien er bij de trekking van een steekproef gestratificeerd is naar bepaalde factoren, moeten die factoren meegenomen worden in het weegmodel.

Om de verzamelde gegevens zo effectief mogelijk te kunnen gebruiken, is het raadzaam dat alle GGD'en dezelfde leeftijdsgrenzen hanteren.

3.3 Non-respons

3.3.1 Theorie

De verzamelde gegevens van de GGD-enquêtes moeten een representatief beeld geven van de populatie. Uit de praktijk blijkt dat de respons op gezondheidsenquêtes onder volwassenen tussen de 45 en 75% ligt. Dat is met name een probleem omdat de respons vaak selectief is, hetgeen betekent dat bepaalde groepen uit de populatie minder geneigd zijn om deel te nemen aan het onderzoek. Factoren als geslacht, leeftijd, sociaal-economische status (SES), burgerlijke staat en etniciteit hangen samen met deze (non-)respons.⁽³⁻⁷⁾ Een oplossing voor deze vertekening door selectieve non-respons is te wegen naar die bepaalde kenmerken van de populatie. Een aanname die hierbij gedaan wordt is dat mensen uit een bepaalde groep die wel deelnemen niet verschillen van mensen uit die groep die niet deelnemen. Als bijvoorbeeld het responspercentage onder mannen lager is dan bij vrouwen wordt er voor het vergroten van de representativiteit aan de mannen een zwaardere weegfactor toegekend. Dit impliceert dat de mannelijke respondenten dezelfde uitkomsten op de onderzoeksvariabelen hebben als de mannelijke non-respondenten zouden hebben. Omdat we verwachten dat deze aanname niet geheel juist is, zijn we ons ervan bewust dat weging van de resultaten vertekening door selectieve non-respons niet geheel uitsluit, maar vermindert.

3.3.2 Contactpogingen

Analyses uitgevoerd door het Centraal Bureau voor de Statistiek (CBS) op gegevens van hun doorlopende enquête (POLS = Permanent Onderzoek Leefsituatie) lieten zien dat niet alleen factoren als geslacht en leeftijd in belangrijke mate de non-respons bepalen, maar ook de contactpogingen (bijvoorbeeld aantal benaderingen, responstijd).⁽³⁾ Zeer waarschijnlijk lijkt de groep die pas in laatste instantie na een bepaald aantal herinneringen deelneemt het meest

op de non-respondenten. Bestuderen van de kenmerken van de groep die in laatste instantie pas deelneemt, kan meer inzicht geven in de groep non-respondenten.

De registratie van de verschillende aspecten van contactpogingen is een soort kwaliteitsindicator voor de schattingen; het geeft de gebruiker inzicht in wat er is gebeurd en hoe de uitkomsten geïnterpreteerd kunnen worden. Ook de reden van non-respons zou meer inzicht kunnen geven in vertekening van de resultaten. Het is echter niet mogelijk om te corrigeren voor deze aspecten van contactpogingen.

3.3.3 Verhoging respons

Voor kwalitatief goede uitkomsten is het wenselijk dat de respons zo hoog mogelijk is. Edwards et al. hebben een review studie uitgevoerd naar manieren die de respons op postenquêtes verhogen.⁽⁸⁾ Hun resultaten worden hierna beschreven. De respons was hoger indien een financiële vergoeding werd gegeven, er een postzegel op de retourenvelop zat, en de naam van de geadresseerde op de brief stond. Een vragenlijst die kort was, met een mooie opzet, en met gekleurde inkt, had de voorkeur. Verder werd er een hogere respons gevonden als er vooraf contact werd opgenomen met de mensen uit de steekproef en als er ook nog herinneringen verstuurd werden of mensen gebeld werden. Als de vragenlijst zeer persoonlijke vragen bevatte die gevoelig kunnen liggen was de respons lager. Het bleek dat de afzender ook van belang was; een vragenlijst vanuit universiteiten werd vaker ingevuld en teruggestuurd dan die uit commerciële instellingen. Aan de begeleidingsbrief moet veel aandacht besteed worden.

In Amsterdam, waar mensen worden uitgenodigd om zelf naar de GGD te komen voor onderzoek, is onderzocht of het voor de hoogte van de respons uitmaakt of mensen op een vast tijdstip werden uitgenodigd of tussen twee tijdstippen in (bijvoorbeeld een dagdeel). Er bleek geen verschil te zijn in de respons.⁽⁹⁾

3.3.4 Conclusie en aanbevelingen

Voor een representatief beeld van de populatie moet gewogen worden naar kenmerken van de populatie die samenhangen met de respons en met de uitkomstvariabelen. Aspecten rondom de benadering, zoals het aantal benaderingen en de responstijd, geven inzicht in wat er gebeurd is, maar er hoeft niet voor te worden gecorrigeerd wanneer gegevens worden vergeleken of samengevoegd.

3.4 Weging

3.4.1 Achtergrond

Om zo goed mogelijk voor vertekening door selectieve non-respons te corrigeren is informatie nodig over de variabelen (bijvoorbeeld leeftijd en geslacht) die samenhangen met non-respons. Van deze variabelen moet de verdeling bekend zijn in zowel de respondentengroep als de lokale populatie. Door deze verdelingen te vergelijken wordt

inzicht verkregen in de mate waarin de respondenten afwijken van de doelpopulatie, en dus in de grootte van de selectieve non-respons en de effecten van stratificatie. De respondenten worden vervolgens in een aantal groepen verdeeld gebaseerd op deze variabelen (strata). Aan deze groepen worden verschillende gewichten toegekend, die bepaald worden aan de hand van verschillen in de samenstelling van de groep respondenten en de lokale populatie. Hierdoor wordt de steekproef representatief met betrekking tot die variabelen waarnaar gewogen is. Als deze variabelen tevens een sterke relatie hebben met de uitkomsten, zullen ook de resultaten van de gewogen steekproef ongeveer representatief zijn voor de populatie.⁽¹⁰⁾ In de praktijk zullen de verschillende uitkomsten echter gerelateerd zijn aan verschillende variabelen, waardoor elke uitkomstmaat eigenlijk een ander weegmodel nodig zou hebben. Omdat dit lastig is, worden verschillende modellen doorgaans gecombineerd tot één model.

3.4.2 Weegmodel

De vraag is welke variabelen het weegmodel zou moeten bevatten. Het huidige weegmodel dat door het CBS bij de POLS-enquête gebruikt wordt, bevat traditionele demografische variabelen die bekend zijn via de GBA (geslacht, leeftijd, burgerlijke staat, regio, urbanisatie, grootte huishouden). Sinds kort bestaat er ook de mogelijkheid om gebruik te maken van het Sociaal Statistisch Bestand (SSB). Dit is een virtueel bestand waarbij verschillende registraties aan elkaar gekoppeld worden. Daardoor is informatie over een groter aantal variabelen dan bijvoorbeeld opgenomen in de GBA beschikbaar voor alle personen in Nederland. Naast demografische variabelen zijn er gegevens over sociaal-economische variabelen, zoals beroep en ontvangen uitkeringen. Overigens is niet voor elke burger informatie beschikbaar over deze verschillende variabelen. Uit onderzoek van het CBS bleek dat ook met weging op basis van deze nieuwe niet-demografische variabelen niet per definitie voor elke vertekening volledig gecorrigeerd werd.⁽¹¹⁾

Uit ander onderzoek bleek dat sociaal-economische status (SES) een belangrijke factor is, die samenhangt met zowel de respons als met uitkomsten.^(5;6) Hoewel opleiding een goede indicator voor SES zou zijn, is deze vaak niet geregistreerd. De GBA bevat geen indicator voor SES en ook in andere registraties is het moeilijk om een goede indicator voor SES te vinden die samenhangt met gezondheid. Deskundigen adviseren om voorlopig geen indicator voor SES mee te nemen in het weegmodel.

Een andere factor die zowel samenhangt met respons als met uitkomsten is etniciteit.⁽¹²⁾

Volgens deskundigen van het CBS moeten we echter voorzichtig zijn met het wegen naar etniciteit, omdat de respondenten binnen die groep te zeer verschillen van de non-respondenten. Hierdoor levert weging naar etniciteit een vertekend beeld op.

Op basis van het CBS-model wordt een weegmodel voor de landelijke cijfers aanbevolen dat de variabelen geslacht, leeftijd, burgerlijke staat en urbanisatiegraad bevat. Dit betekent dat alle lokale onderzoeken deze gegevens moeten meten. Om de representativiteit voor de regionale cijfers te vergroten zouden GGD'en zelf variabelen aan hun model toe kunnen voegen.

3.4.3 Koppeling gegevens

Voor het bepalen van de verdeling van de verschillende variabelen uit het weegmodel in de totale populatie kan informatie gehaald worden uit de registers die de basis vormen voor de steekproeftrekking, zoals de GBA. Voor het bepalen van de verdeling van deze variabelen onder respondenten zijn twee mogelijkheden. De eerste is dat de gegevens uit de vragenlijst gehaald worden. De tweede mogelijkheid is dat de teruggestuurde vragenlijsten gekoppeld worden aan de registers die de basis vormen voor de steekproeftrekking. Bij het gebruik van de eerste methode is het belangrijk dat de informatie verkregen uit de vragenlijst identiek is aan de gegevens die bekend zijn voor de totale populatie. Bij de tweede mogelijkheid moeten respondenten toestemming geven voor de koppeling.

Wij bevelen de tweede methode aan; het 1-op-1 koppelen van de vragenlijsten aan registraties die de basis vormen voor de steekproeftrekking, omdat dan zowel voor de respondenten als voor de totale populatie exact dezelfde variabelen gebruikt kunnen worden. De vragenlijsten dienen daarvoor echter gekoppeld te kunnen worden aan bijvoorbeeld GBA-gegevens. Het probleem hierbij is dat GGD'en dit aan deelnemers moeten melden of zelfs toestemming moeten vragen voor de koppeling van gegevens. Hierdoor zal de non-respons waarschijnlijk verhoogd worden.

3.4.4 Conclusie en aanbevelingen

Voor de berekening van landelijke representatieve cijfers wordt een weegmodel gebruikt. Het weegmodel bevat de variabelen leeftijd, geslacht, burgerlijke staat en urbanisatiegraad. Dit betekent dat deze variabelen bekend moeten zijn voor de respondenten.

GGD'en leveren zelf weegfactoren aan die gebruikt zijn om de representativiteit van de resultaten voor de regio te vergroten.

We blijven zoeken naar variabelen die het weegmodel verder kunnen optimaliseren.

3.5 Samenvoegen of vergelijken van gegevens

De door de GGD'en verzamelde gegevens worden samengevoegd om landelijk representatieve cijfers te kunnen rapporteren. Hierbij is het van belang dat de gegevens zo representatief mogelijk zijn voor de totale landelijke populatie. De verzamelde gegevens kunnen gewogen worden naar de verdeling van bepaalde kenmerken in de populatie. Zoals eerder vermeld gaat het om leeftijd, geslacht, burgerlijke staat en urbanisatiegraad. Indien er van alle regio's gegevens beschikbaar zijn, kunnen de landelijke referentiecijfers bepaald worden door het samenvoegen van alle regio's, waarbij gewogen wordt voor de grootte van de regio's.

Om een beeld te krijgen van een regio, kunnen de weegmodellen in principe dezelfde variabelen bevatten en kan de weegfactor op dezelfde manier worden toegekend. Voor sommige regio's kan het echter interessant zijn om naar een extra variabele te wegen, bijvoorbeeld etniciteit. GGD'en mogen zelf een weegmodel maken voor de regionale cijfers.

De weegfactoren zijn voor de bepaling van landelijke referentiecijfers niet voor alle uitkomstmaten hetzelfde. GGD'en zijn namelijk vrij in het kiezen van de indicatoren die nagevraagd worden. Dit kan tot gevolg hebben dat niet elke GGD alle indicatoren navraagt, waardoor er voor verschillende uitkomstmaten telkens andere regio's in het landelijke weegmodel terechtkomen. De gebruikte variabelen in het weegmodel zijn hetzelfde, maar de waarde zal anders zijn.

3.6 Conclusie

Bij het samenvoegen of vergelijken van gegevens hoeft niet gecorrigeerd te worden voor verschillen in de steekproefomvang of -fractie. Bij het trekken van een gestratificeerde steekproef moet gewogen worden naar de variabelen waarnaar gestratificeerd is om een representatief beeld van de populatie te krijgen.

Voor selectieve non-respons kan eveneens gecorrigeerd worden door de uitkomsten te wegen naar variabelen die zowel samenhangen met de respons als met de uitkomstmaten. In eerste instantie zullen leeftijd, geslacht, burgerlijke staat en urbanisatiegraad in het weegmodel voor landelijke cijfers opgenomen worden. Deze variabelen moeten bekend zijn voor alle respondenten. Het heeft onze voorkeur om die gegevens te halen uit de registraties die de basis vormden voor de steekproeftrekking.

Om regionale representatieve cijfers te kunnen bepalen leveren de GGD'en zelf een weegfactor aan. Daarbij moet worden aangegeven welke variabelen in die weegfactor meegenomen zijn.

4. Seizoensvariatie

4.1 Inleiding

Variatie in de seizoenen waarin gemeten wordt, beïnvloedt mogelijk de vergelijkbaarheid van de verzamelde gegevens. Resultaten kunnen beïnvloed worden door seizoenen, die onder andere variëren in weersgesteldheid (klimaat) en vakantieperioden. In dit hoofdstuk wordt beschreven welke uitkomsten te maken hebben met seizoensvariatie. Daarbij wordt onderscheid gemaakt tussen indicatoren voor gezondheid en leefstijlfactoren.

Allereerst wordt er een literatuuronderzoek beschreven. Resultaten van buitenlandse studies zijn mogelijk niet generaliseerbaar naar Nederland door verschillen in bijvoorbeeld klimaat en cultuur. Daarom worden aanvullende analyses gedaan op een Nederlandse dataset. Naast de invloed die seizoenen kan hebben op de onderzoeksvariabelen, kan ook de respons verschillen, afhankelijk van het seizoen waarin het onderzoek plaatsvindt. Daarbij richten we ons allereerst op het responspercentage. Vervolgens wordt ook de responstijd (tijd tussen versturen van enquête en ontvangst) in ogenschouw genomen.

4.2 Situatieschets

Hoewel de GGD'en verplicht zijn om periodiek onderzoek uit te voeren naar de gezondheidssituatie van de bevolking, staat er niet vast op welk tijdstip dat moet gebeuren. Zo meten de verschillende GGD'en in verschillende jaren en in verschillende maanden of seizoenen.

De referentieperiode verschilt per indicator. Sommige vragen gaan over de afgelopen maand, andere vragen betreffen het hele jaar. Er zijn ook vragen waarin gerefereerd wordt naar "doorgaans" of "gewoonlijk". Maar ook al wordt er naar een grotere tijdsperiode gevraagd, de referentieperiode voor deelnemers is mogelijk gewoon het moment waarop ze de vragenlijst invullen. Daarom zouden ook vragen met een langere referentieperiode onder invloed kunnen staan van seizoensvariatie.

4.3 Methode

Er is een literatuuronderzoek uitgevoerd om de seizoensvariatie in gezondheidsindicatoren en leefstijlfactoren te bestuderen. Tevens hadden we de mogelijkheid om analyses uit te voeren op een bestand van de GG&GD Utrecht, die maandelijks gezondheidsenquêtes verstuurt. De methoden en resultaten van beide onderzoeken zullen apart besproken worden.

4.3.1 Literatuuronderzoek

Via zoekopdrachten in Pubmed zijn studies achterhaald die seizoensvariatie hebben onderzocht in indicatoren voor de gezondheid en leefstijlfactoren. Er zijn de volgende trefwoorden gebruikt: "season", "seasonal variation", "physical activity", "fruit", "smoking",

“depression”, “self-rated health”, “disease”, “mortality”. De keuze voor de verschillende indicatoren is gebaseerd op de indicatoren waarvoor standaardvraagstellingen beschikbaar zijn. Deze standaardvraagstellingen zijn door een werkgroep vastgesteld in het kader van de Monitor VGZ. Alleen studies van de afgelopen tien jaar zijn gebruikt. Indien interessante artikelen werden gevonden, is daarbij ook naar de genoemde referenties gekeken en naar gerelateerde artikelen.

4.3.2 Data-analyse op bestand GG&GD Utrecht

Populatie

De GG&GD Utrecht voert een continue gezondheidspeiling uit. Iedere maand ontvangt een a-selecte groep van 100 Utrechters van 1 jaar en ouder een gezondheidsenquête.

Voor dit onderzoek is gebruik gemaakt van de gegevens uit de periode 1996-1998 omdat voor deze jaren een dataset over het hele jaar en alle seizoenen beschikbaar was. Binnen de totale populatie is een selectie gemaakt van personen in de leeftijd van 19 t/m 64 jaar, aangezien dit de doelgroep is voor de volwassen enquêtes van de GGD'en. Deze selectie resulteerde in een onderzoekspopulatie van 1531 personen.

De vragenlijsten werden per post verstuurd. Mensen die na twee weken de vragenlijst niet hadden teruggestuurd, ontvingen een schriftelijk rappel met het verzoek de vragenlijst alsnog in te vullen. Als de deelnemer nog niet reageerde, belde een interviewer voor een afspraak voor een mondelinge afname bij de mensen thuis. Deelnemers die niet telefonisch bereikbaar waren, ontvingen een tweede rappel met de aankondiging dat een interviewer langskwam. Het percentage mensen dat via een mondeling interview deelnam was gemiddeld slechts 2,5% in de periode 1996-1998. De totale respons over de 3 jaar was gemiddeld 69,4%.

Onderzoeksvariabelen

Ook hier maken we onderscheid tussen indicatoren voor gezondheid en leefstijlfactoren. De dataset bevat de volgende variabelen:

- gezondheidsindicatoren – beperkingen, ervaren gezondheid, psychiatrische problemen en ontsteking van holten;
- leefstijlfactoren – bewegen, fruitconsumptie, groenteconsumptie en ontbijtgebruik.

De exacte vraagstellingen die in de enquête gebruikt zijn voor deze variabelen staan in bijlage 1. In de betreffende onderzoeksjaren waren er nog geen standaardvraagstellingen van de Monitor VGZ beschikbaar.

Voor de indicator beperkingen is het percentage mensen bepaald dat één of meerdere lichamelijke beperkingen rapporteerde. Een goede ervaren gezondheid is gedefinieerd als het rapporteren van een “zeer goede” of een “goede” ervaren gezondheid. Psychische problemen zijn bepaald aan de hand van de General Health Questionnaire (GHQ) die bestaat uit 12 items. Een totaal score is berekend volgens de standaard van Koeter en Ormel⁽¹³⁾ waarbij een score van 2 of meer een psychisch probleem indiceert. Uit de lijst van ziekten/aandoeningen is gekozen voor ontsteking van de holten in de afgelopen 12 maanden, vastgesteld door een arts. Andere ziekten uit de lijst waren meer chronisch en daardoor minder interessant voor de seizoensanalyse. Voor sommige andere ziekten was de prevalentie

te laag. De aanwezigheid van door de arts vastgestelde ontsteking van de holten is gedichotomiseerd als ja versus nee.

Een gunstig beweegpatroon is gedefinieerd als minstens 30 minuten fietsen en/of wandelen per dag. Fruitconsumptie is gedefinieerd als gunstig indien minstens acht stuks per week gegeten werden en voor groenteconsumptie een frequentie van 7 dagen per week. Er was geen informatie over de hoeveelheid groente die geconsumeerd werd.

Analyses

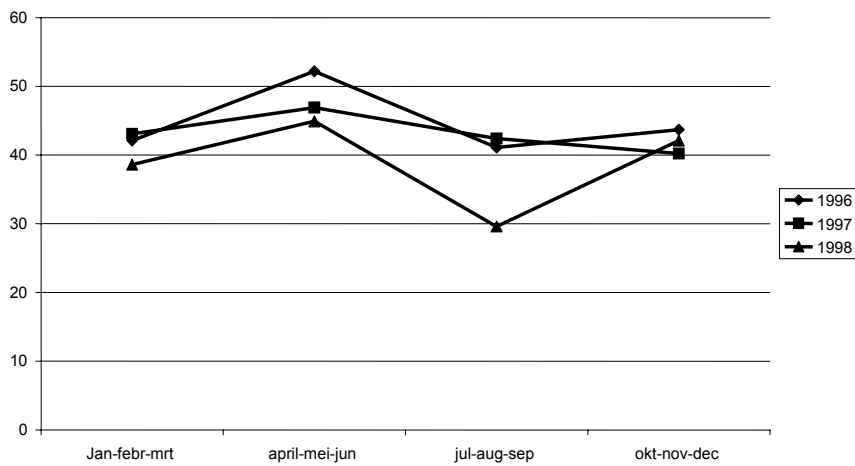
Voor de bepaling van seizoensvariatie is het jaar ingedeeld in kwartalen. Het eerste kwartaal (winterseizoen) bestaat uit de maanden januari, februari en maart. Het tweede uit april, mei, juni (voorjaar), enzovoort. Voor de indeling is gebruik gemaakt van de datum waarop de vragenlijst naar de deelnemer is verstuurd. Er is voor deze datum gekozen, aangezien de datum van ontvangst bij de GG&GD niet voor alle deelnemers bekend is (eerste helft 1996 ontbreekt). Overigens bleek 62% van de deelnemers waarvoor de ontvangstdatum wel bekend was de vragenlijst binnen een maand terug te sturen, waardoor we verwachten dat het verschil tussen de maand van verzenden en de maand van ontvangst bij de GG&GD de resultaten niet sterk zullen beïnvloeden.

Om de seizoensvariatie te bepalen is voor de verschillende gezondheidsindicatoren en leefstijlfactoren het percentage gunstig of ongunstig bepaald per seizoen. Dit is voor elk jaar apart gedaan, om te bekijken of de variatie tussen seizoenen consistent is over de jaren. Indien er over de verschillende jaren gelijke trends te zien waren, zijn de jaren samengevoegd om te kijken of het seizoen een statistisch significant effect had op de betreffende variabele. Conclusies worden echter niet gebaseerd op statistische significantie alleen, aangezien de power laag is.

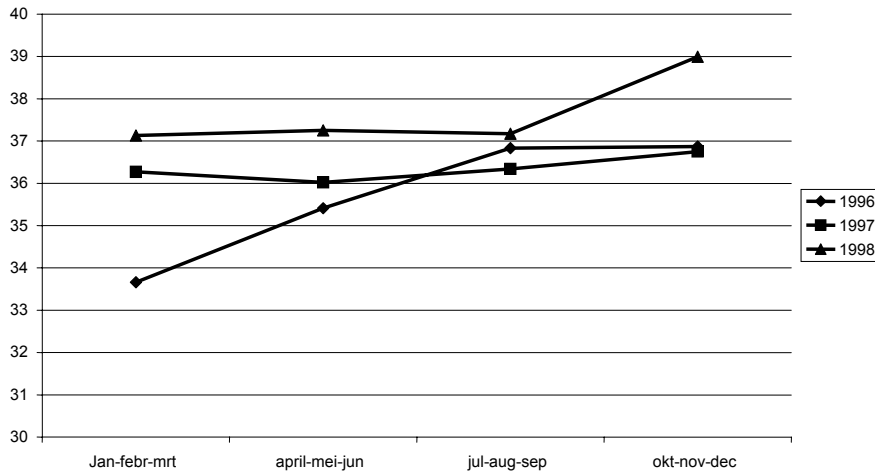
Tenslotte is het responspercentage berekend voor de verschillende seizoenen. Voor de jaren 1997 en 1998 is ook naar de responstijd gekeken. Deze is bepaald aan de hand van de datum waarop de vragenlijst naar de deelnemer is verstuurd en de datum van ontvangst van de vragenlijst op de GG&GD.

Verschillen in studiepopulatie tussen de seizoenen zouden de seizoenseffecten deels kunnen verklaren. Hierbij valt te denken aan geslacht en leeftijd. Als in het ene seizoen bijvoorbeeld meer mannen meedoen dan in het andere, kan dit de uitkomsten beïnvloeden. Figuur 4.1 en 4.2 laten de verdeling zien van respectievelijk het percentage respondenten van het mannelijk geslacht en de gemiddelde leeftijd van de respondenten per seizoen over de drie onderzoeksjaren. Omdat deze variabelen variëren over de seizoenen, is daar in de analyses voor gecorrigeerd.

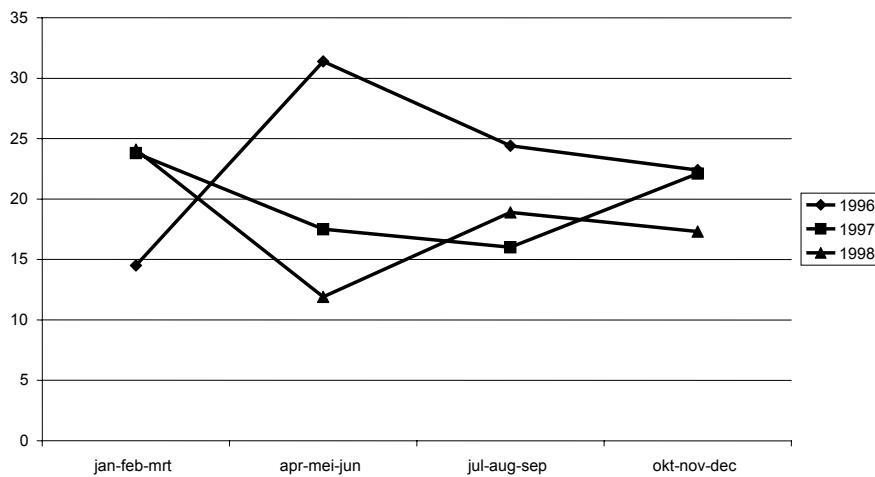
Er is ook gecorrigeerd voor sociaal-economische status (SES), aangezien deze ook varieerde over de seizoenen (Figuur 4.3). Als maat voor SES is het opleidingsniveau van de respondent genomen (laag = geen onderwijs, alleen lager onderwijs, vbo/lbo, mavo of leerlingwezen). Etniciteit van de respondenten varieerde ook over de seizoenen. Omdat deze variabele echter sterk gerelateerd was aan SES en zeer scheef verdeeld, hebben we besloten niet voor etniciteit te corrigeren.



Figuur 4.1 Percentage mannelijke respondenten per seizoen voor de drie onderzoeksjaren



Figuur 4.2 Gemiddelde leeftijd per seizoen voor de drie onderzoeksjaren



Figuur 4.3 Percentage deelnemers met een lage SES (= lage opleiding)

Zoals vermeld heeft een klein deel van de respondenten via een mondeling interview deelgenomen. Hoewel de manier van enquêteren invloed kan hebben op de uitkomsten, is hiervoor niet gecorrigeerd, aangezien de prevalentie van mondeling afgenomen interviews niet varieerde tussen de seizoenen (niet getoond in figuur).

4.4 Resultaten

4.4.1 Uitkomsten literatuuronderzoek

Gezondheidsindicatoren

De indicatoren voor gezondheid uit het literatuuronderzoek kunnen onderverdeeld worden in psychische gezondheid, endogene factoren, en registraties van bepaalde ziekten of sterfte. Zelfgerapporteerde psychische gezondheid, uitgedrukt in prevalentie van depressieve symptomen, angst, of boosheid, bleek seizoensafhankelijk. Het aantal symptomen was hoger in de herfst en winter dan in de zomermaanden⁽¹⁴⁻¹⁶⁾, hoewel dit niet altijd bevestigd werd.⁽¹⁷⁾ Verder kwam postnatale depressie vaker voor in de donkere maanden vergeleken met het voorjaar en de zomer.⁽¹⁸⁾

Endogene indicatoren van lichamelijke gezondheid hingen eveneens samen met seizoenen. Het cholesterolgehalte was in de winter hoger dan in de zomer⁽¹⁹⁾, evenals de body mass index (ook wel quetelet index genoemd) en de middelomtrek.⁽²⁰⁾ Tevens hadden in de winter meer mensen extreem overgewicht (obesitas).⁽²¹⁾

Uit verschillende registraties bleek ook dat lichamelijke gezondheid varieerde over de seizoenen. Allereerst was het voorkomen van beroerte en coronaire hartziekten hoger in de winter dan in de zomer.^(21;22) Een andere maat voor gezondheid is het aantal ziekenhuisopnamen. Voor totale ziekenhuisopnamen en de diagnose hartfalen was het aantal opnamen hoger in de winter dan in de zomer.^(23;24) Ook was totale sterfte en sterfte aan myocard infarct en coronaire hartziekten gerelateerd aan seizoenen, met een piek in de wintermaanden en een daling in de zomermaanden.^(21;25-27)

Het aantal ongevallen bleek ook seizoensafhankelijk. Onder jonge fietsers en voetgangers in de Verenigde Staten was het aantal verkeersongevallen hoger tijdens zomer- dan wintermaanden.⁽²⁸⁾ Dit geldt ook voor het aantal ongevallen met chemische stoffen in Groot-Brittannië.⁽²⁹⁾ Het aantal heupfracturen onder Noorse ouderen was hoger in de wintermaanden.⁽³⁰⁾

Samengevat kunnen we zeggen dat gezondheid afhankelijk is van het seizoen waarin het gemeten wordt. Over het algemeen hebben mensen een slechtere gezondheid in de zomer dan in de winter.

Leefstijlfactoren

Ook diverse leefstijlfactoren bleken af te hangen van het seizoen waarin gemeten wordt. Verschillende studies lieten zien dat mensen in de zomer meer lichamelijke actief zijn en een hoger energieverbruik hebben dan in de winter.^(21;31-33) Dit werd bevestigd in longitudinale analyses.⁽³⁴⁾ Volgens twee studies werd het hogere niveau in de zomer voornamelijk veroorzaakt door de meest actieve mensen, die dat niveau in de winter niet halen, en niet door het percentage inactieve mensen.^(31;32)

Wat voedingsinname betreft hadden mensen in de winter een hogere vetinname^(21;35;36), terwijl in de zomer de inname van vitamine C, fruit en groente het hoogst was.^(21;37;38)

Voor alcoholconsumptie vonden de meeste studies pieken in december (feestdagen) en de zomermaanden.⁽³⁹⁻⁴²⁾ Een andere studie vond alleen seizoensvariatie in de hoeveelheid alcohol onder de dagelijkse drinkers.⁽⁴³⁾

Ook leefstijlfactoren zijn seizoensgerelateerd. Onderzoek uitgevoerd in de zomer zal tot andere resultaten leiden dan onderzoek dat in de winter uitgevoerd is.

4.4.2 Uitkomsten data-analyse GG&GD^a

Gezondheidsindicatoren

Onder de respondenten van de gezondheidsenquête van de GG&GD Utrecht varieerde het percentage dat beperkingen rapporteerde over de seizoenen tussen de 10 en 26%. Hoewel er wel enige variatie tussen de seizoenen was, was dit niet consistent over de jaren (Figuur 4.4). In 1998 werden in de winter significant meer beperkingen gerapporteerd dan in het voorjaar ($p=0,02$), terwijl in de andere jaren in de zomer de meeste beperkingen gerapporteerd leken te worden.

Voor het percentage respondenten dat een goede ervaren gezondheid rapporteerde leek er wel sprake van consistente seizoensvariatie te zijn (Figuur 4.5). Allereerst werd in alle jaren in het najaar het minst vaak een goede ervaren gezondheid gerapporteerd. In 1996 en 1997 werd er het vaakst in de zomermaanden een goede ervaren gezondheid gerapporteerd. Indien de jaren samen werden geanalyseerd, werd echter geen significant effect van seizoen gevonden.

Ook psychische problemen leken consistent te variëren over de seizoenen (Figuur 4.6). In 1996 en 1997 rapporteerden mensen in het najaar vaker psychische problemen dan in de rest van het jaar. In zowel 1997 als 1998 rapporteerden mensen in de zomermaanden minder psychische problemen dan in het voor- en najaar. Als alle jaren samen werden geanalyseerd, bleek het effect van seizoen borderline significant te zijn ($p=0,06$).

In Figuur 4.7 staat het percentage deelnemers dat aangaf in de afgelopen 12 maanden een ontsteking aan de holtten te hebben gehad. In alle 3 jaren werden in het voorjaar en de zomer minder ontstekingen gerapporteerd dan in de winter. Het verschil was statistisch significant tussen de zomer en de winter als alle jaren samen werden geanalyseerd ($p=0,03$).

^a Bij GG&GD Utrecht is een artikel geschreven over seizoensinvloeden aan de hand van hun eigen data⁽⁶⁷⁾. De doelstelling van dat artikel is anders dan die van dit rapport. Ook de resultaten komen niet geheel overeen, hetgeen verklaard kan worden door verschillen in leeftijdsgroep, periode, indeling seizoenen, en correctiemethoden.

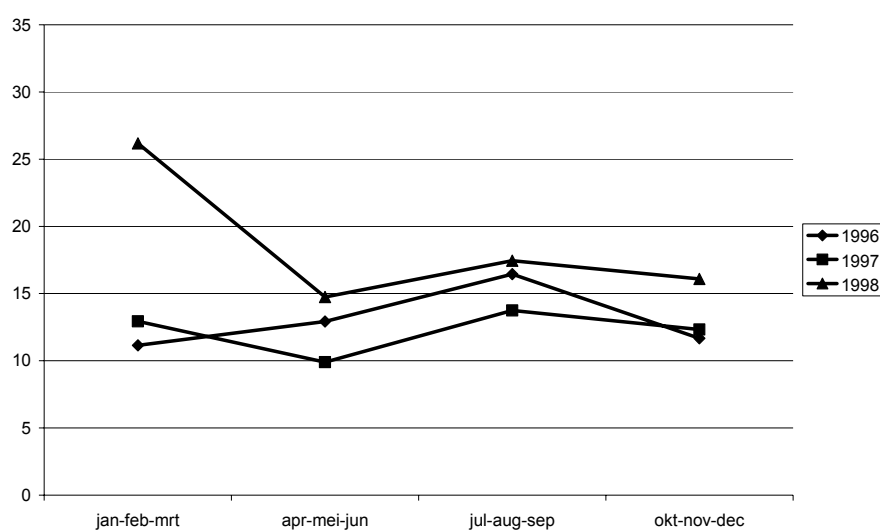
Leefstijlfactoren

In het voorjaar en de zomer waren mensen meer lichamelijk actief dan in de wintermaanden (Figuur 4.8). Dit geldt voor de jaren 1996 en 1998. In 1997 was er nauwelijks variatie over de seizoenen. Indien de jaren samengevoegd werden, was lichamelijke activiteit in het voorjaar en de zomer significant hoger dan in het najaar en de winter ($p=0,004$).

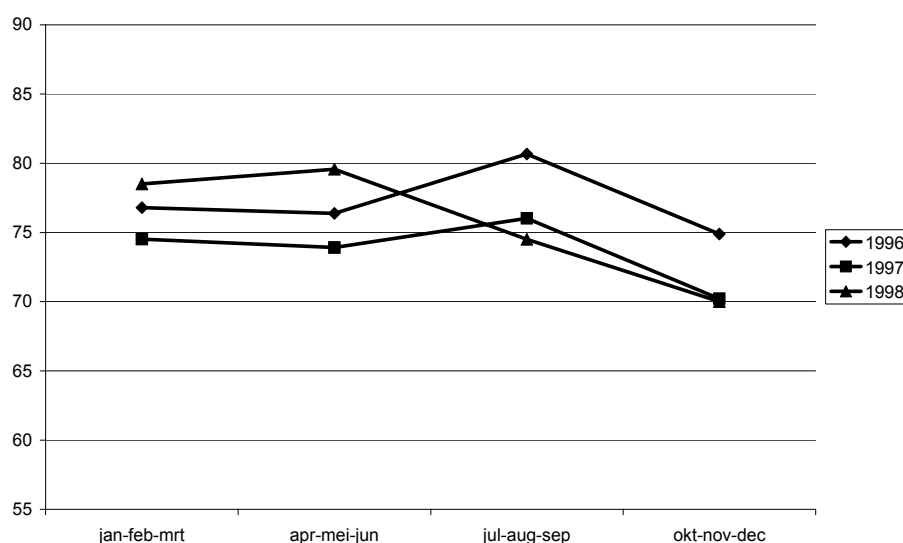
In de jaren 1997 en 1998 was de fruitconsumptie het hoogst in de zomer (Figuur 4.9).

De groenteconsumptie was in 1997 en 1998 in het voorjaar wat lager dan in de andere seizoenen. Ook hier liet 1996 een tegengesteld patroon zien (Figuur 4.10).

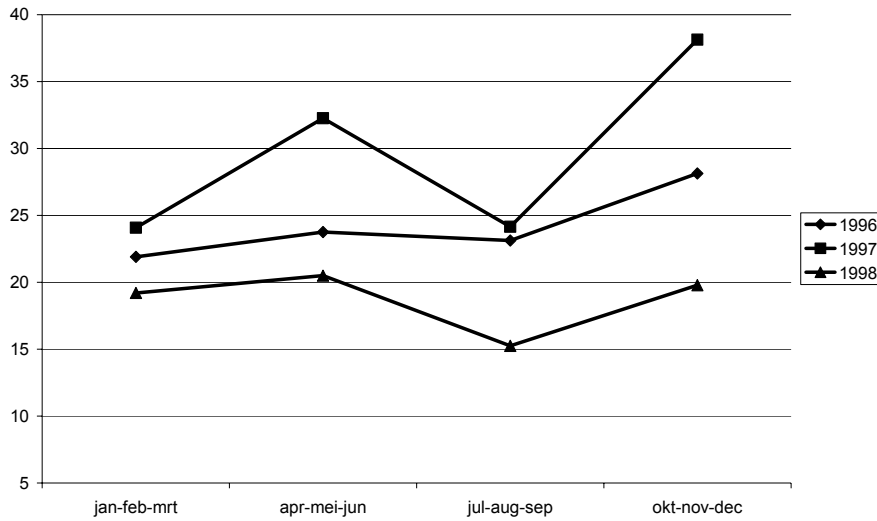
Frequentie van ontbijtgebruik vertoonde een grillig verloop over de seizoenen en de verschillende jaren (Figuur 4.11). Hier was geen sprake van consistente seizoensvariatie.



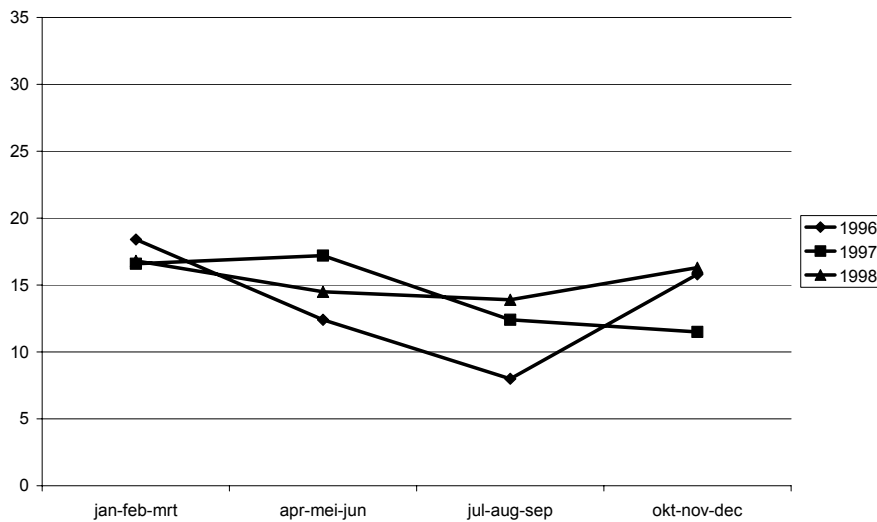
Figuur 4.4 Prevalentie lichamelijke beperkingen per seizoen voor de drie onderzoeksjaren ($n=1494$)



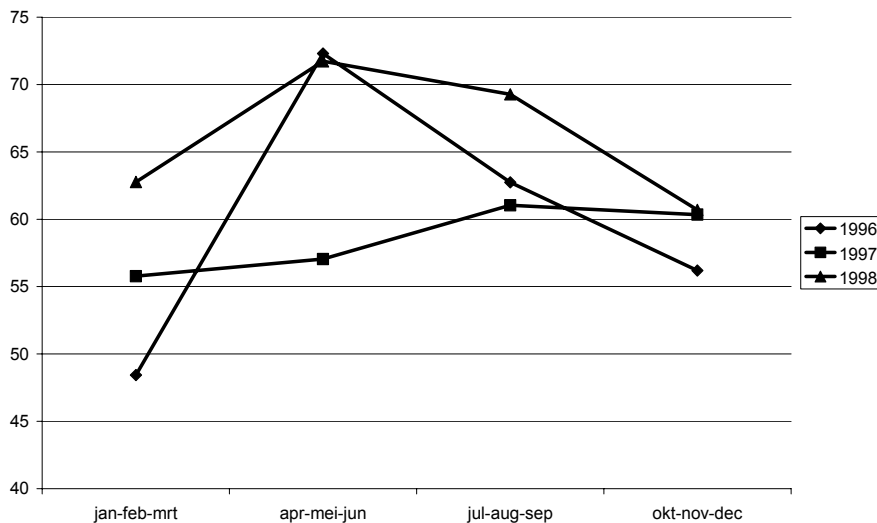
Figuur 4.5 Prevalentie goede ervaren gezondheid per seizoen voor de drie onderzoeksjaren ($n=1470$)



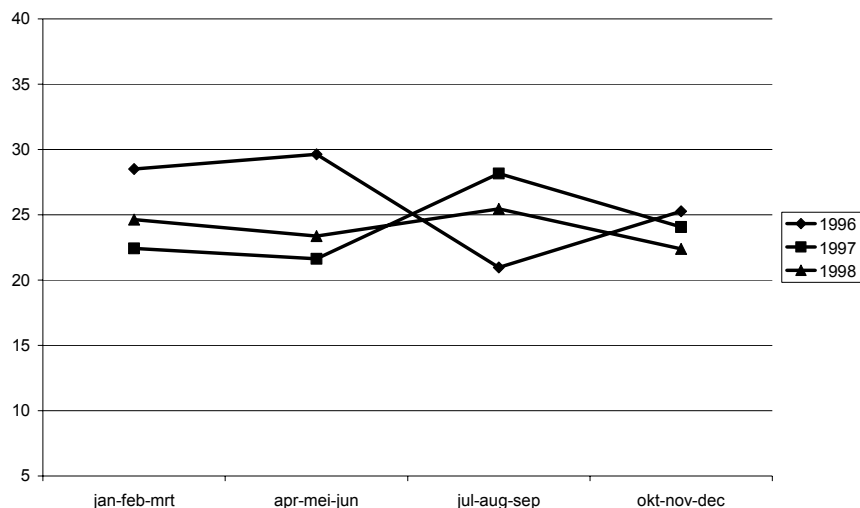
Figuur 4.6 Percentage respondenten met psychische problemen per seizoen voor de drie onderzoeksjaren (n=1455)



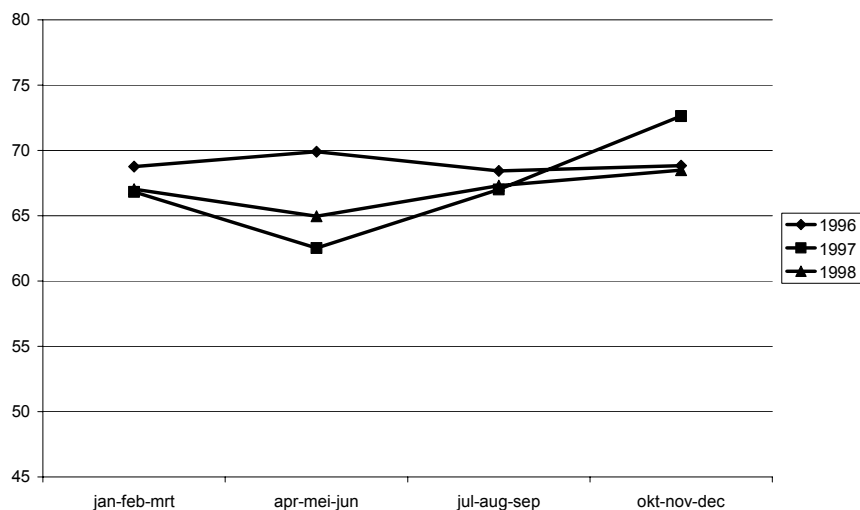
Figuur 4.7 Percentage respondenten met ontsteking holten door een arts vastgesteld



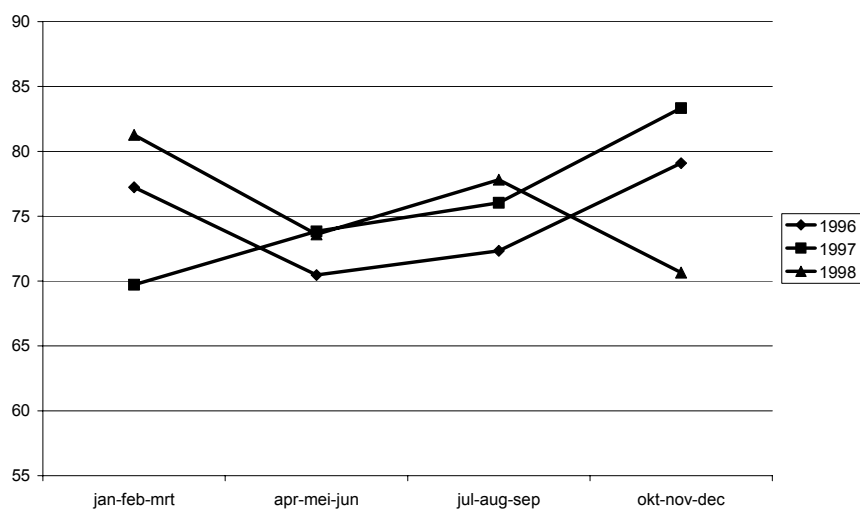
Figuur 4.8 Percentage respondenten dat minstens 30 minuten per dag fietst of wandelt, per seizoen voor de drie onderzoeksjaren (n=1446)



Figuur 4.9 Percentage respondenten dat minstens 8 stuks fruit per week eet, per seizoen voor de drie onderzoeksjaren (n=1520)



Figuur 4.10 Percentage respondenten dat 7 dagen per week groente eet, per seizoen voor de drie onderzoeksjaren (n=1507)

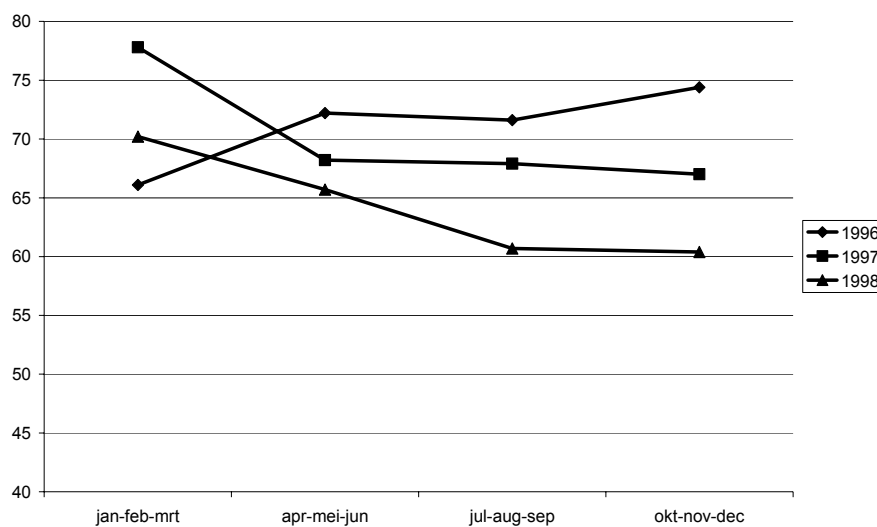


Figuur 4.11 Percentage respondenten dat minstens 5 dagen per week ontbijt, per seizoen voor de drie onderzoeksjaren (n=1496)

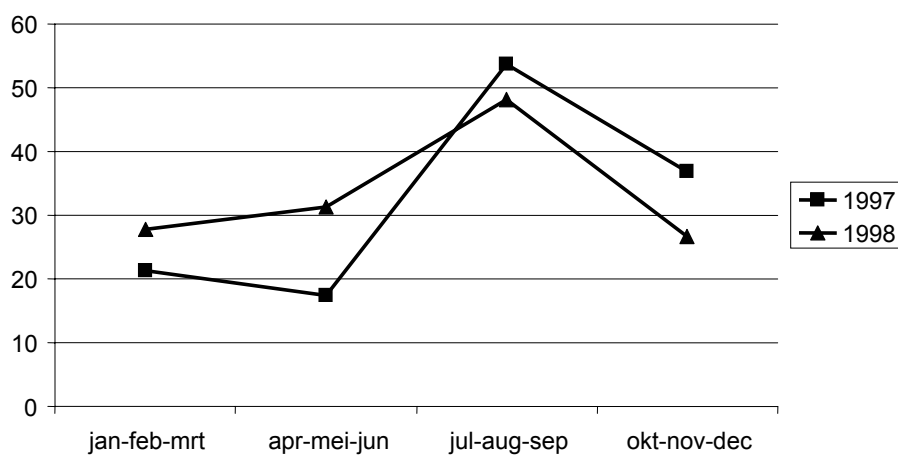
Respons

Het gemiddelde responspercentage was 71,3 voor 1996, 70,2 voor 1997 en 64,2 voor 1998. Hoewel er wel variatie zat in het responspercentage, hing dit niet consistent samen met de verschillende seizoenen (Figuur 4.12).

Voor de responstijd was er duidelijk wel variatie tussen de seizoenen (Figuur 4.13). In de zomermaanden was de gemiddelde responstijd het hoogst ($p < 0,0001$). Tussen de andere seizoenen waren geen opmerkelijke verschillen te zien.



Figuur 4.12 Responspercentage per seizoen voor de drie onderzoeksjaren



Figuur 4.13 Gemiddelde responstijd (dagen) per seizoen voor de jaren 1997 en 1998

4.5 Gevolgen voor vergelijkbaarheid

De resultaten uit het literatuuronderzoek wezen duidelijker op seizoensvariatie dan de analyses uitgevoerd op de GG&GD data. In beide onderzoeken zijn echter aanwijzingen gevonden voor seizoensvariatie in psychische gezondheid, lichamelijke activiteit en fruitconsumptie. Volgens het literatuuronderzoek hangen ook endogene factoren en alcoholconsumptie samen met seizoenen. Over endogene factoren waren in de dataset geen gegevens beschikbaar. Voor alcoholconsumptie konden alleen nooit-drinkers en zware drinkers worden onderscheiden, waarvoor we geen seizoensinvloeden verwachten. Ook de incidentie van ongevallen, ziekte, en sterfte, verkregen uit registraties, hangt volgens het literatuuronderzoek samen met seizoenen. In de vraagstelling van de GG&GD naar ziekte/aandoeningen is het afgelopen jaar als referentieperiode gebruikt, waardoor we verwachtten dat de aanwezigheid van ziekte in de afgelopen 12 maanden gelijk zou moeten zijn voor elk seizoen. Toch bleken mensen in de zomer minder ontstekingen aan de holten te rapporteren dan in de winter. Kennelijk moet ook bij vragen die refereren naar een langere periode rekening gehouden worden met verschillen tussen seizoenen.

De bevestiging van de aanwezigheid van seizoensvariatie in gezondheidsindicatoren en leefstijlfactoren in de literatuur en de analyses van de GG&GD data betekent dat de resultaten van die indicatoren niet vergelijkbaar zijn als er in verschillende seizoenen gemeten is. Voor zover we weten is er in de literatuur niets bekend over corrigeren of wegen indien er sprake is van verschillende seizoenen waarin gemeten is. Theoretisch gezien zou dit mogelijk zijn door bijvoorbeeld lichamelijke activiteit die in de ene populatie in de winter gemeten is, te vergelijken met die in een andere populatie die in de zomer gemeten is, door er een vermenigvuldigingsfactor van 1,4 tussen te zetten. Voor elke variabele zou echter een specifieke vermenigvuldigingsfactor gebruikt moeten worden. Voor het bepalen van de grootte van deze factoren zouden onderzoeken uitgevoerd moeten worden waarin dezelfde mensen in verschillende seizoenen benaderd worden. Tevens moeten deze resultaten door verschillende studies bevestigd worden. Het vergelijken van gegevens tussen GGD'en die in verschillende seizoenen gemeten hebben, is dus eigenlijk niet mogelijk.

Voor het bepalen van landelijke referentiecijfers zou bij het samenvoegen van de gegevens gewogen kunnen worden naar seizoen; seizoenen die minder vaak vertegenwoordigd zijn, tellen dan zwaarder mee. Probleem hierbij is echter dat GGD'en zich dan niet kunnen spiegelen aan het referentiecijfer, hetgeen ook geen optie is.

Op basis van dit onderzoek kunnen we geen uitspraak doen over de vergelijkbaarheid van gegevens die in verschillende jaren gemeten zijn. Daarbij zou er namelijk invloed kunnen zijn van bijvoorbeeld mediacampagnes of veranderingen in wetgeving (denk aan rookvrije werkplek). Het bepalen van een landelijk referentiecijfer zou eventueel over een aantal jaren kunnen gebeuren, waarbij jaren die minder vertegenwoordigd zijn zwaarder meewegen.

4.6 Conclusie en aanbevelingen

Bij het vergelijken van resultaten moet rekening gehouden worden met seizoensinvloeden. Op basis van onze analyses concluderen we dat gegevens die in verschillende seizoenen verzameld zijn, niet altijd met elkaar te vergelijken zijn. Dit geldt in ieder geval voor bepaalde gezondheidsindicatoren, zoals psychische gezondheid en ziekten, en leefstijlfactoren, zoals lichamelijke activiteit. Omdat wegen voor seizoen geen optie is, zouden voor vergelijkbare gegevens GGD'en in hetzelfde seizoen moeten meten.

Aangezien de responspercentages niet consistent samenhangen met seizoen, lijkt er niet een bepaalde periode te zijn die het beste is om de enquêtes af te nemen. Wanneer we naar de responstijd kijken, is in elk geval de zomerperiode het minst geschikt voor de enquête. De responstijd is dan langer, hetgeen meer kosten met zich meebrengt voor GGD'en. Op welk ander tijdstip in het jaar de enquête wel uitgevoerd moet worden, zou af kunnen hangen van praktische overwegingen, zoals vakanties, feestdagen en beschikbaarheid van medewerkers.

5. Manier van enquêteren

5.1 Inleiding

Dataverzameling kan plaatsvinden door middel van een interview (face-to-face), telefonisch interview, schriftelijke enquête via de post of een combinatie van methoden. Keuze hiervoor hangt onder meer af van de verwachte respons en het beschikbare budget. Interviews zijn bijvoorbeeld kostbaarder dan schriftelijke enquêtes.

De manier van enquêteren kan gevolgen hebben voor de totale respons, de respons binnen bepaalde groepen, en de antwoorden op vragen betreffende de gezondheidssituatie en leefstijlfactoren van de bevolking. Dat laatste heeft voornamelijk te maken met de verschillen in ervaren anonimiteit tussen de manieren van enquêteren, wat samenhangt met het geven van sociaal wenselijke antwoorden. Met schriftelijke vragenlijsten zullen mensen meer anonimiteit ervaren dan bij een telefonisch interview, terwijl bij die laatste meer anonimiteit ervaren zal worden dan bij een face-to-face interview. We verwachten dat deze verschillen in ervaren anonimiteit terug te zien zijn in de uitkomstmaten.

Tegenwoordig bestaat ook de mogelijkheid elektronische vragenlijsten in te vullen. Hoewel het invullen van vragenlijsten via internet qua anonimiteit vergeleken kan worden met een schriftelijke vragenlijst, verwachten we dat mensen bij het invullen via internet sneller antwoorden (impulsief) dan bij een schriftelijke vragenlijst, en dat dit dus betrouwbaarder is. Mensen zouden dan minder geneigd zijn tot het geven van sociaal wenselijke antwoorden. Dit hoofdstuk beschrijft welke prevalentieschattingen beïnvloed worden door de manier van enquêteren en welke consequenties dat heeft voor het samenvoegen van en de vergelijking tussen gegevens van de verschillende GGD'en.

5.2 Situatieschets

De meeste GGD'en sturen schriftelijke enquêtes via de post. In Amsterdam worden mensen echter uitgenodigd om naar de GG&GD te komen voor een interview. Dit blijkt namelijk met name onder allochtone mensen een hogere respons op te leveren dan een schriftelijke enquête. Naast het afnemen van mondelinge en schriftelijke enquêtes is de elektronische enquête in opkomst. Hierbij krijgen mensen de mogelijkheid om de enquête via internet in te vullen.

5.3 Methode

Er is een literatuuronderzoek uitgevoerd om het effect van de manier van enquêteren op gezondheidsindicatoren en leefstijlfactoren te bestuderen. Tevens hadden we de mogelijkheid om analyses naar het verschil tussen schriftelijk en elektronisch enquêteren uit te voeren op een bestand van de GGD'en uit de regio Limburg. De methoden en resultaten van beide onderzoeken zullen apart besproken worden.

5.3.1 Literatuuronderzoek

Literatuuronderzoek is uitgevoerd om te bepalen wat de effecten zijn van de verschillende manieren van enquêteren op de respons en op prevalentieschattingen van gezondheidsindicatoren en leefstijlfactoren. In Pubmed is gezocht naar Engelstalige literatuur (uit Nederland en andere landen) en in Google naar Nederlandstalige “grijze” literatuur. Hiervoor zijn de volgende trefwoorden gebruikt: “mode”, “interview”, “respons”, “questionnaire”, “mail”, “internet”, “disease”, “self-rated health”, “depression”, “life style”, “alcohol”, “smoking”, “nutrition”, “fruit”, “vegetables”, “physical activity”. Indien interessante artikelen werden gevonden, is daarbij ook naar de gebruikte referenties gekeken en naar gerelateerde artikelen.

In eerste instantie is de vergelijking gemaakt tussen face-to-face interview en schriftelijke vragenlijst. Omdat er ook literatuur beschikbaar was over de vergelijking tussen telefonisch interview en schriftelijke vragenlijst is dat ook beschreven om de conclusies te bevestigen. Tenslotte zijn de resultaten van schriftelijke vragenlijsten vergeleken met die van elektronische.

5.3.2 Data-analyse op bestand GGD'en regio Limburg

In het najaar van 2003 hebben de GGD'en in Limburg hun 4-jaarlijkse gezondheidsenquête uitgevoerd. Mensen kregen de mogelijkheid om de vragenlijst elektronisch (via internet) in te vullen of schriftelijk. Van de 19717 deelnemers hebben 1504 (7,6 %) de enquête via internet ingevuld.

Met behulp van lineaire modellen is bekeken of de verschillen in manier van enquêteren effect hebben op verschillende gezondheidsindicatoren en leefstijlfactoren. In deze analyses is gecorrigeerd voor de demografische kenmerken die verschillen tussen mensen die via internet en die schriftelijk hebben ingevuld.

5.4 Resultaten

5.4.1 Uitkomsten literatuuronderzoek

Schriftelijke enquête versus interview

Respons

Uit eerder onderzoek bleek dat zowel onder adolescenten als volwassenen de respons hoger was wanneer de gegevens door middel van een interview verzameld werden dan wanneer er een schriftelijke vragenlijst gebruikt werd.⁽⁴⁴⁻⁴⁶⁾ Twee van de drie genoemde studies beschreven in Nederland uitgevoerde surveys.^(45;46) Hierbij moet de kanttekening geplaatst worden dat er in de studie van Picavet⁽⁴⁵⁾ verschillende surveys met elkaar vergeleken zijn waarbij het doel, de focus en de studieopzet verschillen. Daardoor kunnen er tussen de surveys ook andere verschillen zijn die samenhangen met de respons, die niet terug te voeren zijn op verschillende manieren van enquêteren.

Invloed op onderzoeksvariabelen

Voor wat betreft de onderzoeksvariabelen had de manier van enquêteren allereerst invloed op het rapporteren van ziekten en symptomen. Mensen uit de algemene populatie die geïnterviewd werden bleken een betere gezondheidstoestand te rapporteren dan mensen die een vragenlijst hebben ingevuld.^(47;48) Het probleem bij de studie van Bowling et al. is echter dat vergelijken moeilijk is, omdat er drie verschillende surveys gebruikt zijn voor de vergelijking tussen manier van enquêteren, terwijl de surveys verschilden in populaties, doelen en studieopzet. Ook het rapporteren van angst en depressie bleek geassocieerd met de manier van enquêteren. Zowel bij face-to-face als telefonisch interview werden minder symptomen gerapporteerd dan bij een schriftelijke enquête.⁽⁴⁹⁾

Ook bij surveys onder patiënten had de manier van enquêteren invloed. Door zowel dialyse- als astmapatiënten werd een betere kwaliteit van leven gerapporteerd wanneer de patiënten geïnterviewd werden vergeleken met het invullen van een schriftelijke vragenlijst.^(50;51) In de studie onder astmapatiënten hebben alle deelnemers op beide manieren deelgenomen, met een periode van twee weken ertussen en de volgorde random bepaald. In de studie onder dialysepatiënten bestond de geïnterviewde groep uit mensen die om gezondheidsredenen niet schriftelijk konden invullen. Hoewel er gecorrigeerd is voor verschillen in morbiditeit, is deze groep mogelijk ook op andere fronten niet vergelijkbaar met de patiënten die wel een schriftelijke vragenlijst konden invullen.

De manier van enquêteren heeft tevens effect op het rapporteren van diverse leefstijlfactoren. Ook hier lijken mensen over het algemeen in een interview meer geneigd sociaal wenselijk gedrag (= gezonder gedrag) te rapporteren dan in een schriftelijke vragenlijst. Bij rapportage door middel van een vragenlijst bleken mensen minder lichamelijk actief te zijn dan volgens rapportage in een interview.^(52;53) In beide studies hadden de deelnemers zowel via vragenlijst als interview deelgenomen, in een random volgorde met een tussenperiode van respectievelijk 2 weken en 8 dagen.

Voor roken en alcoholconsumptie zijn de gepubliceerde resultaten niet consistent. Terwijl de ene studie onder adolescenten liet zien dat er geen effect van de manier van enquêteren is op het rapporteren van rookgedrag⁽⁴⁴⁾, bleek in een andere studie dat adolescenten in een vragenlijst vaker aangeven te roken dan in een interview.⁽⁵⁴⁾ Dezelfde studie meldde dat voor volwassenen de manier van enquêteren geen invloed had op rapportage van het rookgedrag⁽⁵⁴⁾, evenals een andere studie waarin volwassenen op beide manieren deelnamen.⁽⁵²⁾

Ook de resultaten van het rapporteren van alcoholconsumptie zijn niet consistent. In een studie uit Rotterdam bleek dat er geen verschil was in het rapporteren van alcoholconsumptie tussen een vragenlijst en een interview.⁽⁴⁶⁾ Een studie uit Japan liet echter zien dat meer mensen voldoen aan de definitie van “gezond drinken” volgens een interview vergeleken met een vragenlijst.⁽⁵²⁾ Over het algemeen concludeert het merendeel dat er geen effect is van de manier van enquêteren op rookgedrag en alcoholconsumptie.

Schriftelijke enquête versus telefonisch interview

De verschillen tussen telefonisch interview en schriftelijke vragenlijst leken wat minder sterk aanwezig, maar wezen wel in de richting van meer sociaal wenselijk gedrag bij het rapporteren via het telefonische interview vergeleken met een schriftelijke vragenlijst. In de studies die hierna beschreven worden, zijn deelnemers random ingedeeld voor ofwel een schriftelijke vragenlijst ofwel een telefonisch interview. Vergeleken met een telefonisch interview werd in een schriftelijke vragenlijst een slechtere gezondheid gerapporteerd, zowel betreffende lichamelijke gezondheid^(55;56) als psychische gezondheid.⁽⁵⁷⁾ Voor gemiddelde alcoholconsumptie werd er geen verschil gevonden tussen telefonisch interview en vragenlijst, hoewel zwaar drinken en alcohol-gerelateerde problemen in de schriftelijke vragenlijst vaker gerapporteerd werden.^(58;59)

Wanneer telefonisch interview vergeleken werd met face-to-face interview bleken er voor lichamelijke activiteit en voeding geen verschillen te zijn tussen beide manieren van enquêteren.⁽⁶⁰⁻⁶²⁾

Schriftelijke enquête versus internet

Hoewel het gebruik van elektronische vragenlijsten steeds vaker voorkomt, zijn er nog relatief weinig studies naar het effect hiervan op het rapporteren van gezondheid en leefstijl. De studies die dit wel hebben onderzocht lieten zien dat er nauwelijks verschil was tussen beide manieren van enquêteren (schriftelijk versus elektronisch). Allereerst bleek er tussen mensen die random ingedeeld werden voor een schriftelijke vragenlijst of een elektronische vragenlijst geen verschil te zijn in het rapporteren van onder anderen beperkingen, ademhalingsproblemen, lichamelijke activiteit, en het aantal ziekenhuisopnamen in het afgelopen half jaar.⁽⁶³⁾ In een studie waaraan zowel gezonde mensen als patiënten deelnamen werden er nauwelijks verschillen gevonden in de rapportages van de algemene gezondheid (SF-36) tussen schriftelijke en elektronische manier van invullen.⁽⁶⁴⁾ De kanttekening die bij deze studie gemaakt moet worden is dat er een cross-over opzet is gebruikt met slechts 5 minuten tussen het invullen van de twee lijsten, waardoor mensen mogelijk onthouden hebben wat ze de eerste keer hebben ingevuld. Voor wat betreft leefstijlfactoren was er onder adolescenten en studenten geen verschil in het rapporteren van alcoholconsumptie en andere aspecten van drinkgedrag.^(65;66)

5.4.2 Uitkomsten analyses op bestand GGD'en regio Limburg

Schriftelijke enquête versus internet

Deelnemers die de enquête van de GGD'en Limburg via internet hebben ingevuld verschilden in demografische kenmerken van deelnemers die een schriftelijke enquête ingevuld hebben (Tabel 5.1). Degenen die via internet invulden waren vaker van het mannelijk geslacht, jonger, hoger opgeleid, en woonden vaker in meer stedelijke gebieden (alle p-waarden < 0,0001). Het percentage allochtonen verschilde niet tussen beide groepen (p=0,53).

Voor wat betreft de gezondheidsindicatoren rapporteerden mensen die de schriftelijke vragenlijst hadden ingevuld een slechtere gezondheid dan mensen die via internet invulden (Tabel 5.2). Bijna alle verschillen verdwenen, of werden in ieder geval kleiner na correctie voor de demografische variabelen.

Tabel 5.1 Kenmerken van mensen die vragenlijst schriftelijk of via internet hebben ingevuld

	schriftelijk	internet
geslacht (% mannen)	44%	60%*
leeftijd (gem. in jaren)	48,7	40,9*
score opleiding (range 1-8 [†])	4,6	5,3*
urbanisatiegraad (range 1-5 [‡])	3,4	3,2*
etniciteit (% autochtonen)	94%	94%

* significant verschillend van schriftelijke vragenlijst

† hogere score betekent hogere opleiding

‡ hogere score betekent minder stedelijk gebied

Bron: Limburgse Gezondheidsenquête 2003

Tabel 5.2 Gezondheidsindicatoren gerapporteerd via schriftelijke vragenlijst en via internet

	ruw		gecorrigeerd*	
	schriftelijk	internet	schriftelijk	internet
ervaren gezondheid (range 1-5) [‡]	3,0	2,8 [†]	3,0	3,0
maagklachten (% ja)	26%	23% [†]	26%	25%
pijn botten/spieren (% ja)	46%	34% [†]	44%	41% [†]
diabetes (% ja)	5%	3% [†]	5%	5%
hartinfarct (% ja)	4%	2% [†]	4%	3%
kanker (% ja)	5%	3% [†]	5%	4%
migraine (% ja)	22%	21%	22%	22%
hoge bloeddruk (% ja)	16%	10% [†]	16%	17%
urineverlies (% ja)	8%	4% [†]	7%	8%
beperkingen thuis (% ja)	65%	63%	63%	65%
depressiescore (range 1-3) [§]	1,55	1,43 [†]	1,55	1,50 [†]
beperkingen bewegingsvrijheid (%)	23%	13% [†]	24%	25%
beperkingen communicatie (%)	7,2%	5,3% [†]	8,5%	8,7%

* gecorrigeerd voor leeftijd, geslacht, opleiding en urbanisatiegraad

† significant verschillend van schriftelijke vragenlijst

‡ 1=uitstekend; 2=zeer goed; 3=goed; 4=matig; 5=slecht

§ 1=geen depressieve symptomen; 2=matige depressieve klachten; 3=indicatie voor depressie

Bron: Limburgse Gezondheidsenquête 2003

Tabel 5.3 Leefstijlfactoren gerapporteerd via schriftelijke vragenlijst en via internet

	ruw		gecorrigeerd*	
	schriftelijk	internet	schriftelijk	internet
alcohol				
# glazen door de week	5,6	5,3	5,5	5,3
# glazen weekend	6,3	7,2 [†]	6,3	6,3
% huidige rokers	28%	33% [†]	28%	31% [†]
richtlijn groente (200 gr)	27%	22% [†]	27%	25%
richtlijn groente (150 gr)	52%	44% [†]	52%	47% [†]
richtlijn fruit (200 gr)	29%	18% [†]	28%	25% [†]
richtlijn fruit (incl. sap)	41%	28% [†]	41%	35% [†]
sport (range 1-3) [‡]	1,8	2,0 [†]	1,8	1,9

* gecorrigeerd voor leeftijd, geslacht, opleiding en urbanisatiegraad

[†] significant verschillend van schriftelijke vragenlijst

[‡] 1=sport niet; 2=minder dan 1 keer per week; 3=1 keer per week of vaker

Bron: Limburgse Gezondheidsenquête 2003

De mensen die de schriftelijke vragenlijst hebben ingevuld rapporteerden een gezondere leefstijl dan degenen die via internet invulden (Tabel 5.3). Een deel van die verschillen werd verklaard door verschillen in demografische factoren, hetgeen bleek uit de kleinere verschillen na correctie voor demografische factoren. De rest van de verschillen zou veroorzaakt kunnen worden door resterende verschillen in demografische kenmerken, of doordat mensen bij het invullen van een schriftelijke lijst meer geneigd zijn tot sociaal wenselijk antwoorden.

Gezien de resultaten lijkt het er op dat de verschillen in gerapporteerde gezondheidsindicatoren bijna helemaal verklaard kunnen worden door verschillen in demografische kenmerken, hetgeen suggereert dat er geen effect is van schriftelijk of elektronisch enquêteren op het rapporteren van gezondheidsindicatoren. De verschillen in leefstijlfactoren verdwenen echter niet helemaal na correctie voor demografische kenmerken. Mogelijk wordt de rapportage beïnvloed door het geven van sociaal wenselijke antwoorden.

Gevolgen voor vergelijkbaarheid

Volgens de literatuur verschilden de uitkomsten van schriftelijke vragenlijsten van die van een interview. Ondanks dat niet voor alle uitkomstmaten verschillen werden gevonden, wezen de verschillen die er waren richting meer sociaal wenselijke rapportage bij een interview vergeleken met een vragenlijst. Doordat veel van de genoemde verschillen gebaseerd zijn op slechts één studie, kunnen er geen harde uitspraken gedaan worden over welke variabele wel beïnvloed wordt door de manier van enquêteren en welke variabele niet. Over het algemeen lijken uitkomstvariabelen als ziekte, gezondheid, en leefstijlfactoren niet vergelijkbaar als mensen niet op dezelfde manier benaderd zijn.

Het is de vraag in hoeverre de manier van face-to-face interviewen beschreven in de literatuur vergelijkbaar is met die van de GG&GD Amsterdam. Bij de beschreven studies waarin mensen geïnterviewd werden, kwam de interviewer bij de mensen thuis, terwijl bij de GG&GD in Amsterdam de mensen uitgenodigd werden om zelf naar de GG&GD te komen. Hoewel uit de genoemde studies bleek dat de respons bij een interview hoger was dan bij een schriftelijke vragenlijst, is dat mogelijk anders wanneer mensen zelf naar een locatie toe moeten gaan. Voor wat betreft de uitkomsten verwachten we niet veel verschil tussen een interviewer die bij mensen thuis komt of een interviewer op locatie.

Voor zover we weten is er in de literatuur niets bekend over corrigeren of wegen indien er sprake is van verschillende manieren van enquêteren. Ons inziens is het ook niet mogelijk om hiervoor te wegen, aangezien er dan voor elke variabele een specifieke wegingsfactor gebruikt zou moeten worden om gelijke waarden voor schriftelijke en mondelinge vragenlijsten te krijgen. Voor het bepalen van de grootte van deze wegingsfactoren zouden onderzoeken uitgevoerd moeten worden waarin dezelfde mensen op verschillende manieren benaderd worden. De resultaten moeten door verschillende studies bevestigd worden.

Het gebruik van een elektronische vragenlijst of een schriftelijke vragenlijst leidde volgens de (beperkt beschikbare) literatuur niet tot verschillende resultaten. De aanvullende analyses op het bestand van de GGD'en in Limburg liet echter wel verschillen zien. Mensen leken bij het invullen van een schriftelijke vragenlijst meer geneigd tot het rapporteren van sociaal wenselijk gedrag dan bij de elektronische enquête. Dit gold voor leefstijlfactoren, maar niet voor gezondheidsindicatoren. Er zal meer onderzoek gedaan moeten worden om de resultaten te bevestigen.

Het zal voorlopig nog niet mogelijk zijn om de schriftelijke vragenlijst te vervangen door een elektronische, aangezien de mensen die daarmee bereikt worden niet geheel representatief zijn voor de totale populatie. Verder zou de vergelijkbaarheid met gegevens die schriftelijk verzameld zijn dan tot problemen kunnen leiden.

5.5 Conclusie en aanbeveling

De manier van gegevensverzameling heeft invloed op de vergelijkbaarheid van resultaten. Gegevens die door middel van een schriftelijke vragenlijst zijn verzameld zijn niet vergelijkbaar met gegevens verzameld via een interview. Dit geldt in elk geval voor vragen naar gezondheidsindicatoren. Gegevens die schriftelijk verzameld zijn lijken meer valide doordat er minder sprake is van het rapporteren van sociaal wenselijk gedrag.

Om gegevens tussen GGD'en te kunnen vergelijken moet er op dezelfde manier geënquêteerd worden. De voorkeur gaat uit naar schriftelijke vragenlijsten boven interviews, omdat mensen dan minder geneigd zijn tot sociaal wenselijk antwoorden.

Voor het trekken van conclusies over de vergelijkbaarheid tussen gegevens die via internet verzameld worden of schriftelijk, is aanvullend onderzoek nodig.

6. Conclusie

Op basis van bovenstaand onderzoek kan geconcludeerd worden dat verschillen in methoden van gegevensverzamelingen kunnen leiden tot verschillen in uitkomstmaten. Met deze verschillen dient rekening gehouden te worden wanneer gegevens van GGD'en vergeleken of samengevoegd worden. Bij het invoeren van lokale gegevens in de landelijke database zal een lijst met meta-informatie moeten worden ingevuld met methodologische informatie over de dataverzameling van GGD'en.

Voor het berekenen van landelijke representatieve cijfers zal een weegmodel toegepast worden dat bestaat uit de volgende variabelen: leeftijd, geslacht, burgerlijke staat en urbanisatiegraad. GGD'en dienen deze variabelen voor alle respondenten aan te leveren. GGD'en leveren tevens een door hen gebruikte weegfactor aan die ze gebruiken voor de berekening van lokaal representatieve resultaten. Ze vermelden daarbij op basis van welke variabelen die weegfactor is samengesteld.

We bevelen een uniform meetmoment en manier van enquêteren aan, omdat het vergelijken of samenvoegen van gegevens alleen mogelijk is als het seizoen waarin gemeten is en de manier van enquêteren niet verschillen.

Aanvullend onderzoek zal moeten uitwijzen wat het effect is van elektronisch (via internet) enquêteren, ten opzichte van schriftelijk.

Samengevat zouden GGD'en voor elke respondent gegevens moeten leveren over leeftijd, geslacht, burgerlijke staat, urbanisatiegraad en de weegfactor die ze gebruiken voor regionale cijfers. Daarbij vermelden GGD'en ook uit welke variabelen die weegfactor bestaat. Tevens moet vermeld worden in welk jaar, in welk seizoen en op welke manier (schriftelijk, mondeling, internet) er geënquêteerd is. Indien deze aspecten binnen een onderzoek verschillen tussen respondenten, moet op individueel niveau informatie meegeleverd worden. Daarnaast geven informatie over steekproefomvang, -fractie en stratificatie de gebruikers meer inzicht in de gegevensverzameling.

Dankwoord

Bij het tot stand komen van dit rapport zijn verschillende mensen betrokken geweest. Wij willen hen allen heel hartelijk danken. Robbert Renssen en Ferdy Otten van het CBS hebben hun expertise gedeeld en adviezen gegeven. Daarnaast danken we Harriët van Veldhuizen en Kaspar Bams van GGD Nederland, en Jeanet Bruil van TNO voor het meedenken en kritisch bekijken van de geschreven stukken.

Literatuur

1. Cavelaars, A. Cross-national comparisons of socio-economic differences in health indicators [thesis]. Rotterdam; 1998.
2. de Bruin A, Picavet HS, Nossikov A. Health interview surveys. Towards international harmonization of methods and instruments. WHO Reg Publ Eur Ser 1996;58:i-xiii, 1-161.
3. Bethlehem J, Schouten B. Nonresponse analyses of the Integrated Survey on Living Conditions (POLS). Voorburg/ Heerlen: Statistics Netherlands; 2004.
4. Ronckers C, Land C, Hayes R, Verduijn P, van Leeuwen F. Factors impacting questionnaire response in a Dutch retrospective cohort study. *Ann Epidemiol* 2004;14(1):66-72.
5. Loon AJM van, Tjihuis M, Picavet HSJ, Surtees PG, Ormel J. Survey Non-response in the Netherlands. Effects on Prevalence Estimates and Associations. *Ann Epidemiol* 2003;13(2):105-10.
6. Vollebregt, M. Non-respons in POLS: een verkenning. Voorburg: CBS; 2002.
7. Reijneveld SA, Stronks K. The impact of response bias on estimates of health care utilization in a metropolitan area: the use of administrative data. *Int J Epidemiol* 1999;28(6):1134-40.
8. Edwards P, Roberts I, Clarke M, DiGuseppi C, Pratap S, Wentz R, Kwan I. Increasing response rates to postal questionnaires: systematic review. *BMJ* 2002;324(7347):1183.
9. Voortman JK, Uitenbroek DG, Verhoeff AP, et al. Amsterdamse Gezondheidsmonitor 2003/2004. Pilot Endogene Factoren. Amsterdam: GG&GD Amsterdam; 2003.
10. Bethlehem J, Schouten B. Nonresponse adjustment in household surveys. Voorburg/ Heerlen: Statistics Netherlands; 2004.
11. Te Riele S. Vertekening door non-respons: Hoe nauwkeurig zijn de uitkomsten van persoonsenquetes? Centraal Bureau voor de Statistiek; 2002.
12. Viet AL, Fiolet D, Voortman JK, et al. Evaluatie pilot Endogene Factoren. Bilthoven: RIVM; 2003. Rapportnummer 260854006.
13. Koeter MWJ, Ormel J, composers. General Health Questionnaire. Nederlandse bewerking. Handleiding. Lisse: Swets&Zeitlinger; 1991.

14. Harmatz MG, Well AD, Overtree CE, Kawamura KY, Rosal M, Ockene IS. Seasonal variation of depression and other moods: a longitudinal approach. *J Biol Rhythms* 2000 Aug;15(4):344-50.
15. Murase S, Murase S, Kitabatake M, Yamauchi T, Mathe AA. Seasonal mood variation among Japanese residents of Stockholm. *Acta Psychiatr Scand* 1995;92(1):51-5.
16. Schlager D, Schwartz JE, Bromet EJ. Seasonal variations of current symptoms in a healthy population. *Br J Psychiatry* 1993;163:322-6.
17. Magnusson A, Axelsson J, Karlsson MM, Oskarsson H. Lack of seasonal mood change in the Icelandic population: results of a cross-sectional study. *Am J Psychiatry* 2000;157(2):234-8.
18. Hiltunen P, Jokelainen J, Ebeling H, Szajnberg N, Moilanen I. Seasonal variation in postnatal depression. *J Affect Disord* 2004;78(2):111-8.
19. Ockene IS, Chiriboga DE, Stanek EJ 3rd, Harmatz MG, Nicolosi R, Saperia G, Well AD, Freedson P, Merriam PA, Reed G, et al. Seasonal variation in serum cholesterol levels: treatment implications and possible mechanisms. *Arch Intern Med* 2004;164(8):863-70.
20. Visscher TL, Seidell JC. Time trends (1993-1997) and seasonal variation in body mass index and waist circumference in the Netherlands. *Int J Obes Relat Metab Disord* 2004;28(10):1309-16.
21. Pell JP, Cobbe SM. Seasonal variations in coronary heart disease. *QJM* 1999;92(12):689-96.
22. Wang H, Sekine M, Chen X, Kagamimori S. A study of weekly and seasonal variation of stroke onset. *Int J Biometeorol* 2002;47(1):13-20.
23. Martinez-Selles M, Garcia Robles JA, Prieto L, Serrano JA, Munoz R, Frades E, Almendral J. Annual rates of admission and seasonal variations in hospitalizations for heart failure. *Eur J Heart Fail* 2002;4(6):779-86.
24. Stewart S, McIntyre K, Capewell S, McMurray JJ. Heart failure in a cold climate. Seasonal variation in heart failure-related morbidity and mortality. *J Am Coll Cardiol* 2002;39(5):760-6.
25. Crawford VL, McCann M, Stout RW. Changes in seasonal deaths from myocardial infarction. *QJM* 2003;96(1):45-52.
26. van Rossum CT, Shipley MJ, Hemingway H, Grobbee DE, Mackenbach JP, Marmot MG. Seasonal variation in cause-specific mortality: are there high-risk groups? 25-year follow-up of civil servants from the first Whitehall study. *Int J*

- Epidemiol 2001;30(5):1109-16.
27. Gemmell I, McLoone P, Boddy FA, Dickinson GJ, Watt GC. Seasonal variation in mortality in Scotland. *Int J Epidemiol* 2000;29(2):274-9.
 28. Durkin MS, Laraque D, Lubman I, Barlow B. Epidemiology and prevention of traffic injuries to urban children and adolescents. *Pediatrics* 1999;103(6):e74.
 29. Olowokure B, Saunders PJ, Dyer JA, Kibble AJ. Temporal and seasonal variation in the occurrence of chemical incidents. *Occup Environ Med* 2004;61(2):177-9.
 30. Bulajic-Kopjar M. Seasonal variations in incidence of fractures among elderly people. *Inj Prev* 2000;6(1):16-9.
 31. Plasqui G, Westerterp KR. Seasonal variation in total energy expenditure and physical activity in Dutch young adults. *Obes Res* 2004;12(4):688-94.
 32. Pivarnik JM, Reeves MJ, Rafferty AP. Seasonal variation in adult leisure-time physical activity. *Med Sci Sports Exerc* 2003;35(6):1004-8.
 33. Matthews CE, Hebert JR, Freedson PS, Stanek EJ 3rd, Merriam PA, Ebbeling CB, Ockene IS. Sources of variance in daily physical activity levels in the seasonal variation of blood cholesterol study. *Am J Epidemiol* 2001;153(10):987-95.
 34. Matthews CE, Freedson PS, Hebert JR, Stanek EJ 3rd, Merriam PA, Rosal MC, Ebbeling CB, Ockene IS. Seasonal variation in household, occupational, and leisure time physical activity: longitudinal analyses from the seasonal variation of blood cholesterol study. *Am J Epidemiol* 2001;153(2):172-83.
 35. Shahar DR, Yerushalmi N, Lubin F, Froom P, Shahar A, Kristal-Boneh E. Seasonal variations in dietary intake affect the consistency of dietary assessment. *Eur J Epidemiol* 2001;17(2):129-33.
 36. Woodhouse PR, Khaw KT. Seasonal variation of risk factors for cardiovascular disease and diet in older adults. *Int J Circumpolar Health* 2000;59(3-4):204-9.
 37. Fahey MT, Sasaki S, Kobayashi M, Akabane M, Tsugane S. Seasonal misclassification error and magnitude of true between-person variation in dietary nutrient intake: a random coefficients analysis and implications for the Japan Public Health Center (JPHC) Cohort Study. *Public Health Nutr* 2003;6(4):385-91.
 38. Cox BD, Whichelow MJ, Prevost AT. Seasonal consumption of salad vegetables and fresh fruit in relation to the development of cardiovascular disease and cancer. *Public Health Nutr* 2000;3(1):19-29.
 39. Carpenter C. Seasonal variation in self-reports of recent alcohol consumption: racial

- and ethnic differences. *J Stud Alcohol* 2003;64(3):415-8.
40. Cho YI, Johnson TP, Fendrich M. Monthly variations in self-reports of alcohol consumption. *J Stud Alcohol* 2001;62(2):268-72.
 41. Uitenbroek DG. Seasonal variation in alcohol use. *J Stud Alcohol* 1996;57(1):47-52.
 42. Lemmens PH, Knibbe RA. Seasonal variation in survey and sales estimates of alcohol consumption. *J Stud Alcohol* 1993;54(2):157-63.
 43. Del Rio MC, Prada C, Alvarez FJ. Drinking habits throughout the seasons of the year in the Spanish population. *J Stud Alcohol* 2002;63(5):577-80.
 44. Kaplan CP, Hilton JF, Park-Tanjasiri S, Perez-Stable EJ. The effect of data collection mode on smoking attitudes and behavior in young African American and Latina women. Face-to-face interview versus self-administered questionnaires. *Eval Rev* 2001;25(4):454-73.
 45. Picavet HS. National health surveys by mail or home interview: effects on response. *J Epidemiol Community Health* 2001;55(6):408-13.
 46. Bongers IM, Van Oers JA. Mode effects on self-reported alcohol use and problem drinking: mail questionnaires and personal interviewing compared. *J Stud Alcohol* 1998;59(3):280-5.
 47. Bowling A, Bond M, Jenkinson C, Lamping DL. Short Form 36 (SF-36) Health Survey questionnaire: which normative data should be used? Comparisons between the norms provided by the Omnibus Survey in Britain, the Health Survey for England and the Oxford Healthy Life Survey. *J Public Health Med* 1999;21(3):255-70.
 48. Grootendorst PV, Feeny DH, Furlong W. Does it matter whom and how you ask? inter- and intra-rater agreement in the Ontario Health Survey. *J Clin Epidemiol* 1997;50(2):127-35.
 49. Moum T. Mode of administration and interviewer effects in self-reported symptoms of anxiety and depression. *Social Indicators Research* 1998;45:279-318.
 50. Unruh M, Yan G, Radeva M, Hays RD, Benz R, Athienites NV, Kusek J, Levey AS, Meyer KB. Bias in assessment of health-related quality of life in a hemodialysis population: a comparison of self-administered and interviewer-administered surveys in the HEMO study. *J Am Soc Nephrol* 2003;14(8):2132-41.
 51. Cook DJ, Guyatt GH, Juniper E, Griffith L, McIlroy W, Willan A, Jaeschke R, Epstein R. Interviewer versus self-administered questionnaires in developing a disease-specific, health-related quality of life instrument for asthma. *J Clin Epidemiol*

- 1993;46(6):529-34.
52. Okamoto K, Ohsuka K, Shiraishi T, Hukazawa E, Wakasugi S, Furuta K. Comparability of epidemiological information between self- and interviewer-administered questionnaires. *J Clin Epidemiol* 2002;55(5):505-11.
 53. Vuillemin A, Oppert JM, Guillemin F, Essermeant L, Fontvieille AM, Galan P, Kriska AM, Hercberg S. Self-administered questionnaire compared with interview to assess past-year physical activity. *Med Sci Sports Exerc* 2000;32(6):1119-24.
 54. Brittingham A, Tourangeau R, Kay W. Reports of smoking in a national survey: data from screening and detailed interviews, and from self- and interviewer-administered questions. *Ann Epidemiol* 1998;8(6):393-401.
 55. Brewer NT, Hallman WK, Fiedler N, Kipen HM. Why do people report better health by phone than by mail? *Med Care* 2004;42(9):875-83.
 56. McHorney CA, Kosinski M, Ware JE Jr. Comparisons of the costs and quality of norms for the SF-36 health survey collected by mail versus telephone interview: results from a national survey. *Med Care* 1994;32(6):551-67.
 57. Perkins JJ, Sanson-Fisher RW. An examination of self- and telephone-administered modes of administration for the Australian SF-36. *J Clin Epidemiol* 1998;51(11):969-73.
 58. Kraus L, Augustin R. Measuring alcohol consumption and alcohol-related problems: comparison of responses from self-administered questionnaires and telephone interviews. *Addiction* 2001;96(3):459-71.
 59. Gmel G. The effect of mode of data collection and of non-response on reported alcohol consumption: a split-sample study in Switzerland. *Addiction* 2000;95(1):123-34.
 60. Brustad M, Skeie G, Braaten T, Slimani N, Lund E. Comparison of telephone vs face-to-face interviews in the assessment of dietary intake by the 24 h recall EPIC SOFT program--the Norwegian calibration study. *Eur J Clin Nutr* 2003;57(1):107-13.
 61. Hayden-Wade HA, Coleman KJ, Sallis JF, Armstrong C. Validation of the telephone and in-person interview versions of the 7-day PAR. *Med Sci Sports Exerc* 2003;35(5):801-9.
 62. Tran KM, Johnson RK, Soutanakis RP, Matthews DE. In-person vs telephone-administered multiple-pass 24-hour recalls in women: validation with doubly labeled water. *J Am Diet Assoc* 2000;100(7):777-83.
 63. Ritter P, Lorig K, Laurent D, Matthews K. Internet versus mailed questionnaires: a

randomized comparison. *J Med Internet Res* 2004;6(3):e29.

64. Ryan JM, Corry JR, Attewell R, Smithson MJ. A comparison of an electronic version of the SF-36 General Health Questionnaire to the standard paper version. *Qual Life Res* 2002;11(1):19-26.
65. Kypri K, Gallagher SJ, Cashell-Smith ML. An Internet-based survey method for college student drinking research. *Drug Alcohol Depend* 2004;76(1):45-53.
66. McCabe SE, Boyd CJ, Couper MP, Crawford S, D'Arcy H. Mode effects for collecting alcohol and other drug use data: Web and U.S. mail. *J Stud Alcohol* 2002;63(6):755-61.
67. Bouwman-Notenboom AJ, Toet J, Ameijen EJC van. Seizoensinvloeden in gezondheidsenquêtes: wat is het meest geschikte moment van verzending van vragenlijsten? Seizoensinvloeden in gezondheidsenquêtes. *Tijdschrift voor Gezondheidswetenschappen* *geaccepteerd voor publicatie*.

Bijlage 1. Gebruikte vraagstellingen GG&GD Utrecht; 1996-1998

Gezondheidsindicatoren

Beperkingen

Hebt u een lichamelijke beperking? ja, namelijk
 nee

Ervaren gezondheid

Hoe vindt u uw eigen gezondheid in het algemeen?

- Zeer goed
- Goed
- Gaat wel
- Soms goed, soms slechts
- Slecht

Psychische problemen

Antwoordcategorieën voor de volgende vragen zijn:

- beter dan anders
- hetzelfde als anders
- slechter dan anders
- veel slechter dan anders

Kunt u zich concentreren op de dingen waar u mee bezig bent?

Komt u door zorgen slaap tekort?

Vindt u zich een waardeloos iemand?

Hebt u het gevoel dat u zich nuttig maakt?

Voelt u zich in staat om beslissingen te nemen?

Voelt u zich constant onder spanning staan?

Hebt u het gevoel dat u uw problemen niet de baas kunt?

Hebt u plezier in uw dagelijkse bezigheden?

Bent u in staat om uw problemen onder ogen te zien?

Voelt u zich ongelukkig en neerslachtig?

Raakt u het vertrouwen in uzelf kwijt?

Voelt u zich, alles welbeschouwd, tamelijk gelukkig?

Ziekte en aandoeningen

Hebt u in de afgelopen twaalf maanden één of meer van de volgende ziekten en aandoeningen gehad? Het gaat hierbij om ziekten en aandoeningen die (ooit) *door uw huisarts of specialist zijn vastgesteld*.

	ja	nee
– astma, chronische bronchitis of cara	<input type="checkbox"/>	<input type="checkbox"/>
– ontsteking van de neusbijholte, voorhoofdsholte of kaakholte	<input type="checkbox"/>	<input type="checkbox"/>
– ernstige hartkwaal of hartinfarct	<input type="checkbox"/>	<input type="checkbox"/>
– hoge bloeddruk	<input type="checkbox"/>	<input type="checkbox"/>
– (gevolgen van) een beroerte	<input type="checkbox"/>	<input type="checkbox"/>
– suikerziekte	<input type="checkbox"/>	<input type="checkbox"/>
– (gevolgen van) kanker	<input type="checkbox"/>	<input type="checkbox"/>
– rugaandoening van hardnekkige aard, langer dan drie maanden, of hernia	<input type="checkbox"/>	<input type="checkbox"/>
– gewrichtsslijtage (artrose) van knieën, heupen of handen	<input type="checkbox"/>	<input type="checkbox"/>
– gewrichtsontsteking (chronische reuma, reumatoïde artritis) van handen of voeten	<input type="checkbox"/>	<input type="checkbox"/>
– andere chronische reuma, langer dan drie maanden	<input type="checkbox"/>	<input type="checkbox"/>
– migraine	<input type="checkbox"/>	<input type="checkbox"/>
– depressiviteit, overspannenheid	<input type="checkbox"/>	<input type="checkbox"/>
– overige aandoeningen, namelijk	<input type="checkbox"/>	<input type="checkbox"/>

LeefstijlfactorenBewegen

Hoeveel minuten loopt en fietst u gemiddeld per dag naar het werk, de winkel en dergelijke?

Ongeveer minuten per dag

Fruit

Hoeveel stuks fruit eet u gemiddeld per week? (Uitgeperst vers fruit telt ook mee. Bij klein fruit zoals kersen, kunt u een handjevol voor 1 stuk tellen).

- 0-1 stuks per week
 2-4 stuks per week
 5-7 stuks per week
 8 of meer stuks per week

Groente

Hoeveel dagen per week eet u rauwe of gekookte groente? (groenten uit blik en diepvriesgroenten tellen ook mee).

..... dagen per week

Ontbijt

Hoeveel keer per week ontbijt u? keer per week

Demografische kenmerkenGeslacht

Wat is uw geslacht? man
 vrouw

Leeftijd

Wat is uw leeftijd? jaar

Opleiding

Wat is de *hoogste opleiding* die u met een diploma hebt afgesloten? (U mag maar één antwoord aankruisen)

- lager onderwijs* (basisonderwijs)
- middelbaar algemeen onderwijs* (lavo, ulo/mulo/mavo, enzovoorts)
- hoger algemeen voortgezet onderwijs / voorbereidend wetenschappelijk onderwijs* (havo, mms, hbs, vwo, gymnasium)
- lager of voorbereidend beroepsonderwijs* (lts, lhno, huishoudschool, leao, las, enzovoorts)
- middelbaar beroepsonderwijs* (mts, meao, inas, kvjv, vhbo kleuterleidster, gezinsverzorgster, praktijkdiploma boekhouden, enzovoorts)
- hoger beroepsonderwijs en wetenschappelijk onderwijs* (hts, heao, pa/pabo, lerarenopleiding, sociale academie, hhno, universiteit, hogeschool, enzovoorts)
- anders, namelijk*