



Richtlijn wegen voor GGD-epidemiologen



Beter voorkomen
Kwaliteitsprogramma preventie



Richtlijn wegen voor GGD-epidemiologen

Carolien van den Brink; RIVM

Daan Uitenbroek; GGD Amsterdam en www.quantitativeskills.com

Mandy van den Brink; GGD Hollands Noorden

Elly de Boer; GGD Hart voor Brabant

Klaasjan Hajema; GGD Zuid Limburg

Maarten Schipper; RIVM

Colofon

© GGD Nederland, januari 2009

Richtlijn wegen voor GGD-epidemiologen

Eindredactie: Communicatie GGD Nederland
2009/04

Inhoudsopgave

<i>Hoofdstuk</i>	<i>pagina</i>
1. Achtergrond	9
2. Inleiding	11
3. Methode	13
4. Richtlijn weging	15
5. Discussie	21
6. Handleiding	23
Dankwoord	41
Bijlage 1. Methode	43
Bijlage 2. Resultaten inventarisatie	45
Bijlage 3. Pilot	47
Bijlage 4. Rekenvoorbeeld	49
Bijlage 5. Steekproef gestratificeerd	51
Bijlage 6. Lokale en Nationale Monitor Gezondheid	53

1. Achtergrond

Dit document beschrijft een richtlijn voor het wegen van gegevens die door GGD'en zijn verzameld in het kader van hun periodieke lokale gezondheidsmonitor. Vanuit het Programma Beter Voorkomen is de wens naar voren gekomen voor een dergelijke richtlijn. Beter Voorkomen wordt uitgevoerd door GGD Nederland, de VNG en ActiZ, in opdracht van het ministerie van VWS en onder regie van ZonMw. Het programma wil een extra impuls geven om kwaliteit gezamenlijk aan te pakken en stimuleert en ondersteunt GGD'en bij het opzetten van kwaliteitsbeleid in hun eigen organisatie. In de klankbordgroep Beter Voorkomen bestond de behoefte aan richtlijnen voor verschillende statistische methoden, waarbij weging de hoogste prioriteit had.

De doelstelling voor de richtlijn is 3-ledig:

- een hulpmiddel voor GGD-epidemiologen voor de weging van hun lokale en regionale cijfers;
- streven naar uniformiteit in weging tussen GGD'en wat de vergelijkbaarheid van cijfers verhoogt;
- een uitgangspunt voor weging bij de berekening van landelijke cijfers.

Er wordt aangesloten bij het project van de Lokale en Nationale Monitor Gezondheid. In dit project vindt namelijk al standaardisatie plaats van de gegevensverzamelingen bij GGD'en, onder andere door het vaststellen van standaardvraagstellingen.

Andere methodologische vraagstukken rondom het verzamelen en analyseren van gegevens vallen buiten deze richtlijn.

2. Inleiding

De verzamelde gegevens van de GGD-enquêtes moeten een representatief beeld geven van de populatie. Er kunnen verschillende redenen zijn waarom de gegevens niet representatief zijn. Ten eerste kunnen bepaalde groepen oververtegenwoordigd zijn in de groep respondenten wanneer de steekproef gestratificeerd is getrokken. Sommige GGD'en hogen groepen op om in die groepen voldoende aantallen te hebben om betrouwbare uitspraken te kunnen doen. De gegevens zullen dan moeten worden gewogen voor de variabelen die bij de stratificatie van de steekproef zijn gebruikt. Ten tweede kan de respons selectief zijn, wat betekent dat bepaalde groepen uit de populatie minder geneigd zijn om deel te nemen. Uit de praktijk blijkt dat de respons op gezondheidsenquêtes onder volwassenen tussen de 45 en 65% ligt, wat vooral een probleem is als de respons selectief is. Als oplossing voor deze vertekening door selectieve respons wordt vaak gewogen naar bepaalde kenmerken van de populatie die samenhangen met zowel de respons als de uitkomstvariabelen. 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. Hierbij wordt de (niet verifieerbare) aanname gedaan dat de mannen die wel deelnemen niet verschillen in uitkomstvariabelen van mannen die niet deelnemen. Omdat deze aanname meestal niet volledig klopt, zal weging van de resultaten de vertekening door non-respons niet wegnemen, maar slechts verminderen.

Om zo goed mogelijk voor vertekening door non-respons te corrigeren is informatie nodig over achtergrondkenmerken die samenhangen met de respons en met de uitkomstvariabelen. Van deze kenmerken moet de verdeling bekend zijn in zowel de respondentengroep als de totale (lokale) populatie. Na weging naar die kenmerken zullen de resultaten van de gewogen steekproef een betere afspiegeling zijn van de populatie.

Hoewel weging leidt tot vermindering van vertekening door non-respons, kan de betrouwbaarheid van de uitkomsten afnemen ten gevolge van het wegen. Bij weging maak je kleine groepen namelijk relatief belangrijk (precisie verlies) en grotere groepen relatief minder belangrijk (precisie winst). Omdat het verlies van precisie groter kan zijn dan de winst, kunnen de betrouwbaarheid minder en de betrouwbaarheidsintervallen groter worden. Verschillen zullen hierdoor minder snel significant zijn. De mate waarin de variantie verandert ten gevolge van het wegen, wordt het designeffect genoemd. Hiermee moet rekening worden gehouden bij het berekenen van betrouwbaarheidsintervallen en het toetsen van verschillen.

Deze richtlijn beschrijft hoe GGD'en cijfers kunnen berekenen voor hun gemeenten en regio. Het gaat hier om de vraag naar welke factoren gewogen moet worden, hoe die factoren moeten worden berekend en hoe die te gebruiken in de analyses.

In hoofdstuk 6 staat een handleiding voor de toepassing van de richtlijn in SPSS.

3. Methode

De richtlijn is als volgt tot stand gekomen: Allereerst is er geïnventariseerd bij GGD-epidemiologen hoe zij doorgaans met wegen omgaan. Daarnaast is er een pilot uitgevoerd op data om te bekijken wat de effecten zijn van verschillende weegmethoden.

Een werkgroep bestaande uit GGD-epidemiologen en onderzoekers van het RIVM zijn vier keer bij elkaar geweest om de richtlijn te schrijven. De andere GGD-epidemiologen is om input gevraagd via GGD-kennisnet en er is een themamiddag over de richtlijn gehouden. Ook is het CBS geraadpleegd.

De resultaten van de inventarisatie, de pilot en de werkgroep staan beschreven in de bijlagen 1, 2 en 3.

4. Richtlijn weging

In deze paragraaf worden overwegingen besproken en aanbevelingen gedaan voor de weging.

Steekproeftrekking

De steekproeven dienen aselekt te worden getrokken (= kanssteekproef) binnen de strata.

Verder maakt het voor de weegmethode niet uit hoe de steekproef wordt getrokken. Zolang de GGD'en hun resultaten wegen naar de variabelen waarvoor de steekproef is gestratificeerd, is het niet noodzakelijk dat alle GGD'en op gelijke manier stratificeren. Elke extra stratificatie brengt echter ook extra werk bij het wegen met zich mee.

Weegvariabelen

De verzamelde gegevens worden in ieder geval gewogen naar de variabelen waarnaar de steekproef is gestratificeerd. Op basis van de uitkomsten van de pilot en ervaringen door GGD-epidemiologen luidt het advies van deze richtlijn om de vertekening door non-respons te verminderen door te wegen naar geslacht, leeftijd en etniciteit. Als het uitsplitsen naar etniciteit te kleine aantallen in de cellen oplevert (celaantallen kleiner dan 5), kan beter niet worden gewogen naar etniciteit.

De variabelen worden bij de weging gekruist in een tabel.

Voor het berekenen van (sub)regionale cijfers wordt ook rekening gehouden met het aantal inwoners per gemeente.

Geslacht – wordt ingedeeld in mannen en vrouwen.

Leeftijd – wordt ingedeeld in meerdere groepen. Hiervoor wordt aangesloten bij de leeftijdsgrenzen die gehanteerd worden in de Lokale en Nationale Monitor Gezondheid.

Voor jeugd in de volgende groepen:

- 0 t/m 3 jaar: per leeftijdsjaar;
- 4 t/m 11 jaar: 4 t/m 8; 9 t/m 11;
- 12 t/m 18 jaar: 12 t/m 14; 15 t/m 18.

Voor volwassenen in 3 groepen: 19 t/m 34; 35 t/m 49; 50 t/m 64 jaar.

Voor ouderen in 3 groepen: 65 t/m 74; 75 t/m 84; 85 jaar en ouder.

Wanneer GGD'en andere leeftijdsgrenzen hanteren bij het uitvoeren van hun monitor, is de aanbeveling om voor de weging zoveel mogelijk aan te sluiten bij bovengenoemde grenzen.

Etniciteit – Wanneer de steekproef niet gestratificeerd is naar verschillende etnische groepen, wordt voor de weging etniciteit ingedeeld in autochtonen en allochtonen.

Hiervoor wordt de definitie van het CBS gebruikt waarbij iemand tot de allochtone bevolking wordt gerekend als hij tenminste één ouder heeft die in het buitenland is

geboren. Als het uitsplitsen naar etniciteit te kleine aantallen in de cellen oplevert (celaantallen kleiner dan 5), wordt er niet gewogen naar etniciteit.

Gemeente – bij het berekenen van regionale cijfers moeten de gegevens worden gewogen naar gemeentegrootte (= aantal inwoners).

Andere variabelen die in overweging zijn genomen

Hieronder volgt een beschrijving van de variabelen die in de richtlijn niet als weegvariabele worden opgenomen, maar wel zijn overwogen:

Burgerlijke staat – Volgens de pilot heeft het wegen naar burgerlijke staat een klein effect op de resultaten voor volwassenen. Vanwege het kleine effect en om praktische redenen wordt in deze richtlijn niet geadviseerd om te wegen naar burgerlijke staat. Het toevoegen van burgerlijke staat als weegfactor zorgt voor veel meer cellen en dus extra werk. Daarnaast is het verkrijgen van aantallen op populatieniveau voor de combinatie van etniciteit en burgerlijke staat vaak lastig.

SES – Het is bekend dat sociaal economische status (SES) gerelateerd is aan de respons en aan uitkomstvariabelen. Over het algemeen is de respons onder mensen met een lagere SES lager en hebben mensen met een lagere SES een slechtere gezondheid en een minder gezonde leefstijl dan mensen met een hogere SES. Hoewel SES daarom een potentiële variabele is om naar te wegen, is het lastig om een geschikte proxy te vinden voor SES.

Inkomen wordt soms gebruikt als indicator voor SES. Om te kunnen wegen naar inkomen zijn voor zowel de respondenten als de totale populatie inkomensgegevens nodig. Voor de totale populatie zijn die (nog) niet beschikbaar. Daarnaast nemen veel GGD'en de vraag naar netto inkomen niet op in de vragenlijst omdat die vraag vaak slecht wordt ingevuld.

Opleiding is een andere indicator voor SES. Gegevens over opleiding zitten niet in de GBA, maar zijn wel beschikbaar bij het CBS. De CBS gegevens zijn echter niet gebaseerd op de hele populatie, maar op een steekproef, en zijn niet helemaal vergelijkbaar met de standaardvraagstelling uit de Monitor. Daarom wordt opleiding ook niet als weegvariabele opgenomen in de richtlijn.

De *statusscore* is een maat voor de sociale status van postcodegebieden die is ontwikkeld door het Sociaal en Cultureel Planbureau. Deze score is gebaseerd op o.a. opleiding, inkomen en woningprijs en is beschikbaar op postcode 4-niveau. Omdat deze score op het niveau van de respondenten niet te bepalen is, is het niet zinvol om te wegen naar statusscore. Weging naar postcode-4 zou dan gelijk staan aan wegen naar statusscore. Daarnaast is postcode-4 als maat voor SES niet erg nauwkeurig.

Populatietotalen

De populatietotalen voor geslacht, leeftijd, etniciteit en gemeente zijn beschikbaar via Statline van het CBS.

Berekenen van weegfactoren

Er zijn verschillende manieren om weegfactoren te berekenen. In deze richtlijn worden twee manieren beschreven, die worden onderscheiden als 'kleine weegfactoren' (w_i) en 'grote weegfactoren' (W_i). Beide weegmethoden leveren dezelfde resultaten op, maar de interpretatie van de gewichten verschilt. De kleine gewichten worden berekend door de proporties van de strata in de totale populatie te delen door de proporties van diezelfde strata in de respondentengroep. Dit gewicht geeft aan hoe zwaar iemand meetelt bij het berekenen van cijfers. De grote gewichten worden berekend door de aantallen in de strata in de totale populatie te delen door de aantallen van diezelfde strata in de respondentengroep. Het gewicht geeft per respondent aan hoeveel personen hij vertegenwoordigt in de totale populatie.

Het voordeel van de kleine gewichten is dat zij inzicht geven in wat er gebeurt met het wegen (welke groepen wegen zwaarder mee en welke juist niet). Ook worden deze gebruikt voor het afkappen van weegfactoren (zie volgende paragraaf). De grote gewichten liggen dicht bij de werkelijkheid en kunnen bij allerlei analyses worden gebruikt zonder dat de gewichten opnieuw moeten worden geschaald. In de meeste statistische pakketten wordt om die reden gebruik gemaakt van de grote gewichten.

Grootte van weegfactoren

Soms zijn de aantallen in de cellen te klein en/of zijn er eigenlijk te veel weegvariabelen (of categorieën) meegenomen. Dit betekent dat de respondenten in die cellen extreem zouden moeten worden opgehoogd. Omdat dit tot vertekening van de resultaten kan leiden, worden de extreem hoge gewichten afgekapt. Er is gekozen voor een afkapwaarde 5. Het gaat daarbij om de zogenaamde kleine gewichten (w_i). Alle gewichten groter dan 5 worden dan op 5 gezet. Hetzelfde geldt voor extreem kleine gewichten. Gewichten kleiner dan 0,2 worden op 0,2 gezet. De uiteindelijke gewichten moeten na het afkappen opnieuw worden geschaald, zodat de totale steekproefomvang voor en na wegen hetzelfde is. De grote gewichten (W_i) veranderen mee.

Missende waarden

Missende waarden op de weegvariabelen

De gegevens van respondenten met missende waarden op een van de weegvariabelen kunnen verschillend worden behandeld. De beste manier zou zijn om voor die respondenten de variabele te imputeren op basis van uitkomsten op andere variabelen. Dit is echter veel werk.

Daarom wordt geadviseerd om mensen met een missende waarde op leeftijd en geslacht te verwijderen. Omdat verwacht wordt dat er bij etniciteit vaker sprake zal zijn van missende waarden kunt u er voor kiezen niet te wegen naar etniciteit als er te veel missende waarden zijn.

Berekenen van prevalentiecijfers

Voor de gemeente

Prevalentiecijfers worden berekend met behulp van weegfactoren. Bij het berekenen van prevalentiecijfers voor een gemeente wordt in ieder geval gewogen naar de variabelen waarvoor de steekproef is gestratificeerd en naar bovengenoemde variabelen. Voor het berekenen van gewogen prevalentiecijfers kan de procedure 'Weight cases' in SPSS worden gebruikt. Dit kan zowel met de kleine als de grote weegfactoren.

Voor de regio

Bij het berekenen van prevalentiecijfers voor de regio moet rekening worden gehouden met gemeentegrootte. Daarom worden hiervoor de grote weegfactoren gebruikt. Ook moet rekening worden gehouden met het steekproefdesign.

Een alternatief is om de kleine gewichten te herberekenen. De formule hiervoor staat op pagina 26.

Jeugd

Als de steekproef voor jeugd uit de GBA is getrokken, kan er op dezelfde manier als hierboven beschreven worden gewogen.

De meeste GGD'en voeren hun monitor onder jongeren echter via scholen uit. Omdat deze richtlijn zich beperkt tot weging van GBA-gebaseerde steekproeven worden geen uitspraken gedaan over weging in onderzoek onder scholen.

Betrouwbaarheidsintervallen en het toetsen van verschillen tussen gemeenten of regio's

Het toetsen van verschillen tussen gemeenten gebeurt door de gewogen cijfers met elkaar te vergelijken. Dit kan door middel van een t-toets waarbij eerst de standaardfouten van de gewogen cijfers moeten worden berekend. Het berekenen van die standaardfout is echter lastig. Voor het berekenen van de standaardfout of van betrouwbaarheidsintervallen rondom gewogen prevalentiecijfers moet namelijk rekening worden gehouden met een grotere onbetrouwbaarheid als gevolg van het wegen. Er zijn statistische pakketten zoals Epi-info, SPSS-module 'Complex Samples', R, Stata en SAS die rekening houden met het zogenaamde designeffect. Omdat GGD-epidemiologen voornamelijk met SPSS werken, wordt de module 'Complex Samples'

geadviseerd. Het gebruik van 'Complex Samples' wordt in de handleiding uitgewerkt. Er zijn wel kosten verbonden aan deze module.

Andere analyses op de data

Voor het toetsen van verschillen tussen andere groepen dan gemeenten of regio's, moeten de grote weegfactoren worden gebruikt. Als de analyses bijvoorbeeld alleen op vrouwen worden gedaan, zouden de kleine weegfactoren opnieuw moeten worden berekend. De grote weegfactoren blijven echter gelijk.

Daarnaast kunnen er allerlei andere analyses voor bijvoorbeeld wetenschappelijke vraagstellingen worden uitgevoerd. Het wege bij deze analyses valt buiten deze richtlijn en is afhankelijk van de vraagstelling die u wilt beantwoorden.

5. Discussie

Dit document geeft een richtlijn voor het wegen van gegevens die door GGD-epidemiologen zijn verzameld. Het doel van de weging is om vertekening van de resultaten door stratificatie en door non-respons, zoveel mogelijk te verminderen. Daarbij realiseren we ons dat de vertekening nooit kan worden uitgesloten.

Het effect van wegen zal voor elke uitkomstvariabele en voor elke meting verschillen. De resultaten van de pilot gelden dan ook alleen voor de provincie Flevoland in 2005 voor die uitkomstvariabelen die we hebben bekeken. Theoretisch gezien zou er bij elk onderzoek voor elke uitkomstvariabele opnieuw moeten worden bekeken naar welke factoren er gewogen moet worden en wat het effect daarvan is. Omdat dat praktisch niet haalbaar is, worden er in de richtlijn variabelen genoemd waarnaar altijd, dus voor elke uitkomstmaat en in elk onderzoek, zou moeten worden gewogen. Daarbij is een afweging gemaakt tussen wat theoretisch gezien het beste zou zijn en de praktische haalbaarheid van de uitvoering (hoeveelheid werk, beschikbaarheid van gegevens).

De vraag kan rijzen of er niet net zo goed niet gewogen kan worden als het ideaalplaatje toch niet haalbaar is. Met weging worden de resultaten echter vrijwel altijd beter en wordt de werkelijkheid dus beter benaderd, tenzij de weegvariabelen niet samenhangen met de respons en de uitkomstvariabelen of de respondenten in een bepaalde groep totaal niet lijken op de non-respondenten. Eenzelfde vraag kan gesteld worden over de analyses in 'complex samples'. Hoeveel maakt het eigenlijk uit of je wel of niet rekening houdt met het designeffect? Dit zal echter per onderzoek verschillen en is vooraf niet te voorspellen. Het advies luidt dan ook om wel te wegen, dan gaat het in ieder geval goed.

Een andere vraag is of eerst onderzocht moet worden of de respons wel selectief is. De respons is echter altijd wel enigszins selectief. En het is lastig om te zeggen waar de grens dan ligt van wel en niet selectief. Daarom wordt aanbevolen om altijd te wegen. Bij het wegen voor non-respons wordt de aanname gedaan dat de uitkomsten van de respondenten in een bepaalde groep vergelijkbaar zijn met de non-respondenten in diezelfde groep. Hoewel dat niet voor alle uitkomstvariabelen zal gelden, wordt met de weging dat gedeelte dat verklaard wordt door leeftijd, geslacht en etniciteit in ieder geval wel meegenomen.

Weging naar etniciteit wordt in sommige onderzoeken niet gedaan, omdat allochtone respondenten te veel zouden verschillen van de non-respondenten. Allochtonen gaan

echter steeds meer op autochtonen lijken, ook wat hun respons-gedrag betreft. Verder zal de eventuele ophoging van de allochtone groep respondenten nooit tot heel veel vertekening leiden, omdat extreem hoge weegfactoren worden afgekapt.

6. Handleiding

Het berekenen van prevalentie cijfers en betrouwbaarheidsintervallen per gemeente.

Stap 1. Verwijderen van respondenten met missende waarden

Respondenten met missende waarden op leeftijd of geslacht (en eventueel etniciteit) worden verwijderd uit het bestand.

Stap 2. Berekenen van kleine (w_i) en grote (W_i) weegfactoren per gemeente

In box 1 staat een syntax waarmee u de gegevens van de respondenten voor leeftijd kunt indelen in 3 groepen en voor etniciteit in 2 groepen. Bekijk vervolgens voor elke gemeente het aantal allochtone respondenten gekruist naar leeftijd en geslacht. Als dit celaantallen kleiner dan 5 oplevert, wordt er niet gewogen naar etniciteit. Onderstaande tabellen worden dan niet uitgesplitst naar etniciteit.

Maak 6 tabellen met daarin per gemeente leeftijd (3 groepen), geslacht (2 groepen) en etniciteit (2 groepen). Zie voorbeeld op de volgende pagina.

In de eerste tabel zet u per cel het aantal mensen in de totale populatie. Deze gegevens kunt u uit de GBA of uit Statline halen (zie box 2).

In de tweede tabel berekent u proporties (%) van de verschillende strata voor de totale populatie door de celaantallen te delen door het totale inwoneraantal en te vermenigvuldigen met 100. In de derde tabel staat per cel het aantal respondenten. In tabel 4 worden de proporties voor de respondenten berekend zoals in tabel 2. In tabel 5 komen de kleine weegfactoren, door de proporties uit de totale populatie (tabel 2) te delen door de proporties onder de respondenten (tabel 4). In tabel 6 komen de grote weegfactoren, door de aantallen uit de totale populatie (tabel 1) te delen door de aantallen onder de respondenten (tabel 3).

Bij stap 3 staat hoe u de berekening van de weegfactoren kunt laten uitvoeren via een speciaal programma.

Als u de steekproef gestratificeerd had getrokken naar leeftijd of etniciteit, breidt u de tabel uit met extra categorieën van leeftijd of etniciteit indien van toepassing. Heeft u gestratificeerd naar andere variabelen, dan kunt u in bijlage 5 lezen hoe u daarmee om moet gaan.

Een rekenvoorbeeld voor gemeenten zonder allochtonen staat in bijlage 4.

Tabel 1. Aantal inwoners voor de gemeente

Gemeente	Autochtoon		Allochtoon	
	man	Vrouw	man	Vrouw
19-34				
35-49				
50-64				

Tabel 2. Proporties per cel ten opzichte van de totale populatie

Gemeente	Autochtoon		Allochtoon	
	Man	Vrouw	man	Vrouw
19-34				
35-49				
50-64				

Tabel 3. Aantal respondenten in de gemeente

Gemeente	Autochtoon		Allochtoon	
	Man	Vrouw	man	Vrouw
19-34				
35-49				
50-64				

Tabel 4. Proporties per cel ten opzichte van het totaal aantal respondenten

Gemeente	Autochtoon		Allochtoon	
	Man	Vrouw	man	Vrouw
19-34				
35-49				
50-64				

Tabel 5. Kleine weegfactoren (w_i)

Gemeente	Autochtoon		Allochtoon	
	man	Vrouw	man	Vrouw
19-34				
35-49				
50-64				

Tabel 6. Grote weegfactoren (W_i)

Gemeente	Autochtoon		Allochtoon	
	man	Vrouw	man	Vrouw
Leeftijd				
19-34				
35-49				
50-64				

Box 1. Syntax voor indeling etniciteit en leeftijd in andere categorieën

```

IF AGETS201=1 ALLOCHTOON=1.
IF AGETS202=1 ALLOCHTOON=1.
IF (AGETS201=2 AND AGETS202=2) ALLOCHTOON=2.

VARIABLE LABELS
  ALLOCHTOON 'allochtoon ja/nee'.

VALUE LABELS
  ALLOCHTOON 1 'ja' 2 'nee' 9 'onbekend'.

MISSING VALUES
  ALLOCHTOON (9).

IF MISSING (ALLOCHTOON) ALLOCHTOON=9.

EXECUTE.

IF AGLFS201=19 LEEFTIJD=1.
IF AGLFS202=1 LEEFTIJD=1.
IF AGLFS202=2 LEEFTIJD=1.
IF AGLFS202=3 LEEFTIJD=1.
IF AGLFS202=4 LEEFTIJD=2.
IF AGLFS202=5 LEEFTIJD=2.
IF AGLFS202=6 LEEFTIJD=2.
IF AGLFS202=7 LEEFTIJD=3.
IF AGLFS202=8 LEEFTIJD=3.
IF AGLFS202=9 LEEFTIJD=3.

VARIABLE LABELS
  LEEFTIJD 'leeftijd 3 categorieën'.

VALUE LABELS
  LEEFTIJD 1 '19-34 jaar' 2 '35-49 jaar' 3 '50-64 jaar'.

MISSING VALUES
  LEEFTIJD (9).

IF MISSING (LEEFTIJD) LEEFTIJD=9.

(AGETS201=eerste generatie allochtoon ja/nee; AGETS202=tweede generatie allochtoon ja/nee;
AGLFS201=5-jaars leeftijd klasse)

```

Box 2. Populatietotalen in Statline (Inclusief etniciteit (autochtoon/allochtoon))

- Ga naar www.statline.nl.
- Klik op 'Thema'.
- Klik onder het kopje 'Thema's' op 'Bevolking'.
- Klik daarnaast onder het kopje 'Bevolking' op 'Allochtonen/herkomstgroepering'.
- Klik daarnaast onder het kopje 'Allochtonen/herkomstgroepering' op 'bevolking per regio; herkomst'.
- Klik in het tabblad 'Onderwerpen' op 'Totale bevolking' en selecteer in de kolom daarnaast mannen en vrouwen apart.
- Kies in het tabblad 'Herkomstgroepering' voor 'Alle herkomstgroeperingen' en selecteer vervolgens 'Allochtoon' en 'Autochtoon'.
- Kies in het tabblad 'Regio's' voor 'Gemeenten van het jaar ' (afhankelijk van uw onderzoeksjaar). Klik in de kolom daarnaast op 'Gemeentelijst van ...'. Selecteer in de gemeentelijst de gemeenten die voor uw regio van toepassing zijn.
- Kies in het tabblad 'Leeftijd' de categorieën die voor uw onderzoek van toepassing zijn. Het is helaas niet mogelijk om in combinatie met etniciteit de aantallen van alleen de 19-jarigen op te vragen. Maak de selectie daarom vanaf 20 jaar. De aantallen in de cellen voor 20 tot 25-jarigen vermenigvuldigt u vervolgens met 1,2. De aanname die hierbij wordt gedaan is dat de aantallen mannen en vrouwen en allochtonen niet verschillen tussen 19-jarigen en 20 tot 25-jarigen.
- Klik in het tabblad 'Perioden' op 'Groepen van jaren' en selecteer uw onderzoeksjaar.
- Klik op de grote pijl met 'toon gegevens'.
- De getoonde tabel kan vervolgens gedownload worden naar Excel.

Exclusief etniciteit (autochtoon/allochtoon)

- Ga naar www.statline.nl.
- Klik op 'Thema'.
- Klik onder het kopje 'Thema's' op 'Bevolking'.
- Klik onder het kopje 'Bevolking' op 'Bevolking; leeftijd, regio'.
- Klik in het tabblad 'Onderwerpen' op 'Bevolking naar geslacht' en selecteer in de kolommen daarnaast mannen en vrouwen apart.
- Kies daarna bij 'Leeftijd' voor 'Afzonderlijke leeftijden' en klik dan in de kolom daarnaast op 'Afzonderlijke leeftijden' en vink in de lijst alleen 19 jaar aan.
- Kies daarna bij 'Leeftijd' voor 'Totalen per leeftijdsgroep', klik in de kolom daarnaast op 'Vijfjaars leeftijdsgroepen' en selecteer daar de categorieën van 20 tot 65 jaar.
- Kies in het tabblad 'Regio's' voor 'Gemeenten van het jaar ' (afhankelijk van uw onderzoeksjaar). Klik in de kolom daarnaast op 'Gemeentelijst van '. Selecteer in de gemeentelijst de gemeenten die voor uw regio van toepassing zijn.
- Klik in het tabblad 'Perioden' op 'Groepen van jaren' en selecteer uw onderzoeksjaar.
- Klik op de grote pijl met 'toon gegevens'.
- De getoonde tabel kan vervolgens gedownload worden naar Excel.

Stap 3. Afkappen van weegfactoren

Voor het afkappen van weegfactoren worden de kleine weegfactoren gebruikt. Kleine weegfactoren (w_i) hoger dan 5 of lager dan 0,2 worden afgekapt op respectievelijk 5 en 0,2. De overige kleine weegfactoren in dezelfde gemeente en de bijbehorende grote gewichten (W_i) worden opnieuw berekend, zodat de totale steekproefomvang voor en na het afkappen van de gewichten even groot is.

Er is een speciaal programma geschreven om de weegfactoren na afkappen te laten herberekenen. Dit is te downloaden via

<http://www.quantitativeskills.com/manuals/weighting.htm> en werkt als volgt.

Er moeten per gemeente 2 kolommen worden ingevoerd. Deze kunnen worden gekopieerd uit bijvoorbeeld Excel. In de eerste kolom staan de aantallen uit de populatietabel. De cellen van de verschillende strata (geslacht, leeftijd en etniciteit) worden onder elkaar gezet. In de tweede kolom gebeurt hetzelfde voor de respondent aantallen. Voor elke gemeente moet dit apart gebeuren. In figuur 1 ziet u hoe de gegevens moeten worden ingevoerd.

Figuur 1. Input voor berekening van weegfactoren

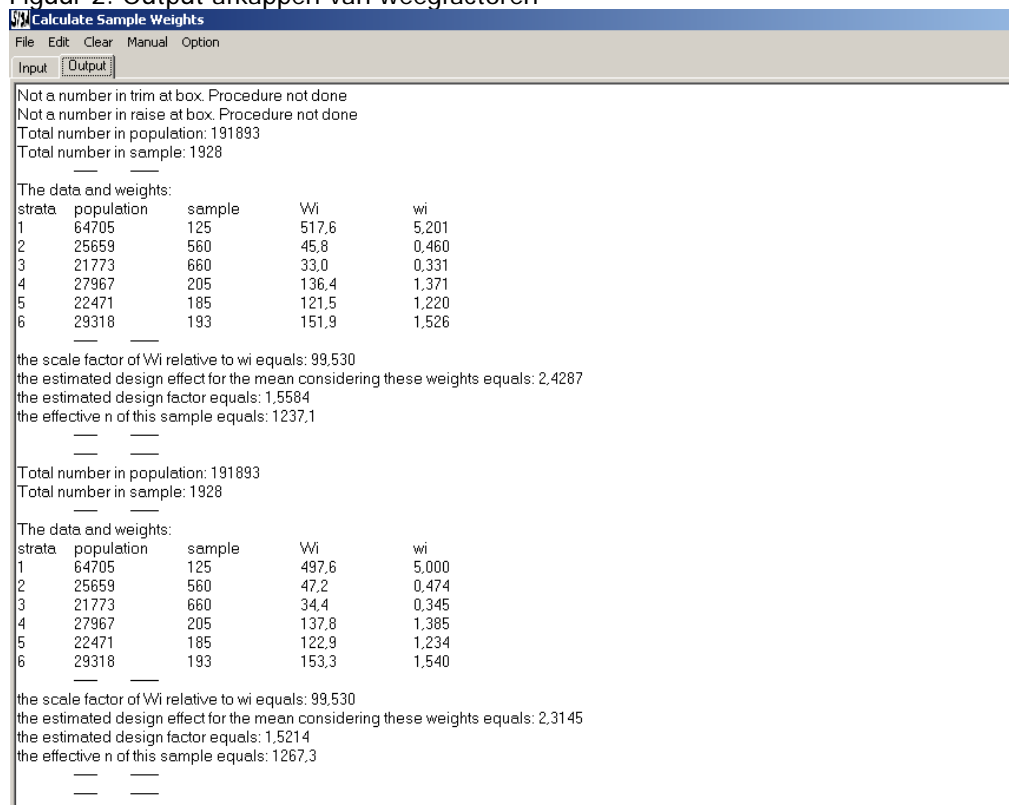
64705	125
25659	560
21773	660
27967	205
22471	185
29318	193

raise weights below:

trim weights above:

[Quantitative Skills](#)

Figuur 2. Output afkappen van weegfactoren



In figuur 2 staat de output. In de derde kolom staan de grote weegfactoren (W_i) die aangeven hoeveel personen 1 respondent vertegenwoordigt. In de vierde kolom staan de weegfactoren (w_i) die de verhouding van de proporties tussen de populatie en de respondenten aangeven.

In het onderste deel staan de weegfactoren nadat ze zijn afgekapt op 5 en 0,2 en opnieuw geschaald. Zoals te zien is, veranderen de grote weegfactoren mee. De designfactor is ook terug te vinden in de output.

Het is ook mogelijk om de weegfactoren voor alle gemeenten in één keer te berekenen. Hiervoor klikt u in het programma op option – split file. Vervolgens kunt u aangeven na hoeveel cellen er een nieuwe gemeente begint (bijvoorbeeld na 6 cellen, als er is uitgesplitst naar geslacht en 3 leeftijdsgroepen).

Stap 4. Toevoegen van weegfactoren aan bestand

Voeg de weegfactoren toe voor de verschillende combinaties van leeftijd, geslacht en etniciteit per gemeente. Dit kunt u voor zowel de grote als de kleine weegfactoren doen.

De SPSS-syntax die u daarvoor voor elke gemeente kunt gebruiken staat in box 3.

Box 3. Syntax voor toevoegen van weegfactoren in databestand

```
IF (LEEFTIJD=1 AND AGGSB201=1 AND BBGMM408=108 AND ALLOCHTOON=1)
MIWV201=a.
IF (LEEFTIJD=2 AND AGGSB201=1 AND BBGMM408=108 AND ALLOCHTOON=1)
MIWV201=b.
IF (LEEFTIJD=3 AND AGGSB201=1 AND BBGMM408=108 AND ALLOCHTOON=1)
MIWV201=c.
IF (LEEFTIJD=1 AND AGGSB201=2 AND BBGMM408=108 AND ALLOCHTOON=1)
MIWV201=d.
IF (LEEFTIJD=2 AND AGGSB201=2 AND BBGMM408=108 AND ALLOCHTOON=1)
MIWV201=e.
IF (LEEFTIJD=3 AND AGGSB201=2 AND BBGMM408=108 AND ALLOCHTOON=1)
MIWV201=f.
IF (LEEFTIJD=1 AND AGGSB201=1 AND BBGMM408=108 AND ALLOCHTOON=2)
MIWV201=g.
IF (LEEFTIJD=2 AND AGGSB201=1 AND BBGMM408=108 AND ALLOCHTOON=2)
MIWV201=h.
IF (LEEFTIJD=3 AND AGGSB201=1 AND BBGMM408=108 AND ALLOCHTOON=2)
MIWV201=i.
IF (LEEFTIJD=1 AND AGGSB201=2 AND BBGMM408=108 AND ALLOCHTOON=2)
MIWV201=j.
IF (LEEFTIJD=2 AND AGGSB201=2 AND BBGMM408=108 AND ALLOCHTOON=2)
MIWV201=k.
IF (LEEFTIJD=3 AND AGGSB201=2 AND BBGMM408=108 AND ALLOCHTOON=2)
MIWV201=l.
(AGGSB201=geslacht, BGMM408=gemeente).
```

Stap 5. Berekenen van prevalentiecijfers

Het berekenen van frequenties of gemiddelden per gemeente kunt u in SPSS doen. Eerst moet de weegfactor worden aanzet. Klik daarvoor op data – weight cases en selecteer de variabele van de weegfactor. Dit kan zowel de kleine als de grote weegfactor zijn. Vervolgens kunt u de gewenste cijfers uitdraaien via analyse – descriptive statistics of via analyse-tables. De weegfactor wordt dan automatisch meegenomen. Deze kunt u weer uitzetten via data – weight cases. Doet u hier geen significantie toets, want er wordt geen rekening gehouden met designeffecten. Het toetsen kunt u in complex samples doen, wat in stap 6 verder wordt uitgewerkt.

Stap 6. Berekenen van betrouwbaarheidsintervallen en verschillen tussen gemeenten

Bij het berekenen van betrouwbaarheidsintervallen moet rekening worden gehouden met het designeffect. Er zijn twee methoden:

1) U kunt de output bij stap 3 gebruiken. Na weging neemt de standaardfout toe met de design factor. De betrouwbaarheidsintervallen nemen toe met diezelfde factor.

2) Complex samples

Complex samples berekent niet zelf de gewichten. Deze kunt u berekenen in Excel of met behulp van het eerder genoemde programma op www.quantitativeskills.com. Naast de gewichten verwacht Complex samples ook een variabele voor de strata. Deze geeft aan in welk stratum (cel) iemand hoort. Zie box 4.

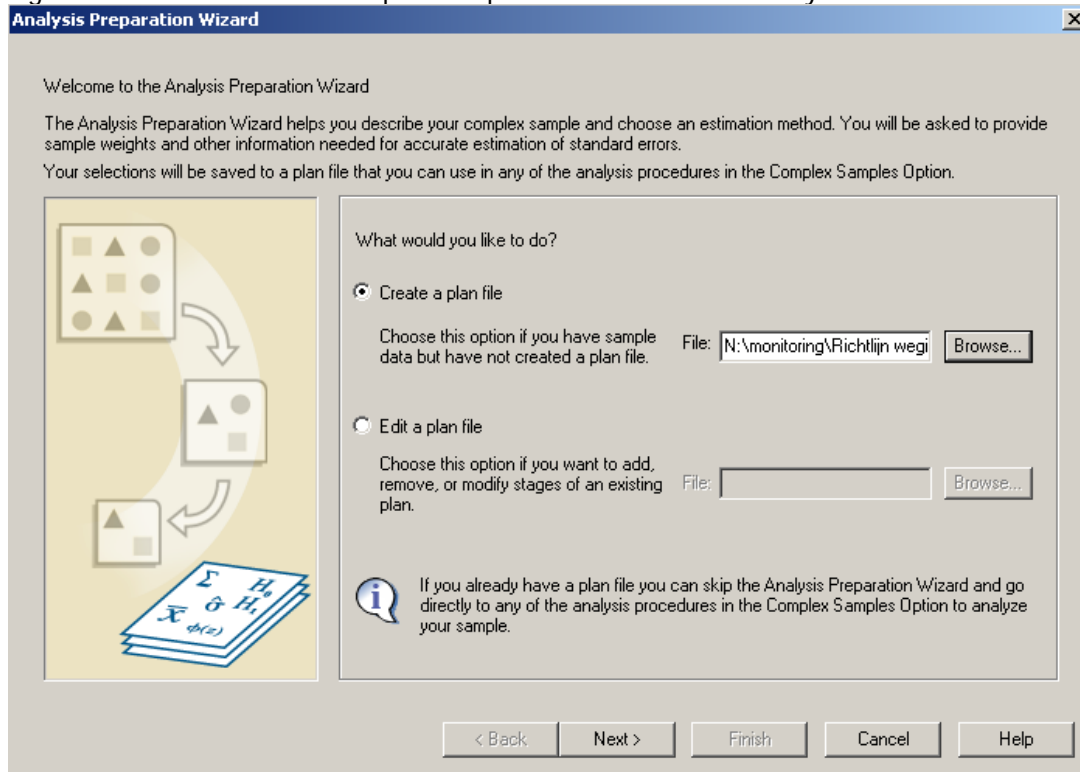
Box 4. Syntax voor toevoegen van strata aan databestand

```
IF (LEEFTIJD=1 AND AGGSB201=1 AND BBGMM408=108 AND ALLOCHTOON=1)
STRATA=1.
IF (LEEFTIJD=2 AND AGGSB201=1 AND BBGMM408=108 AND ALLOCHTOON=1)
STRATA=2.
IF (LEEFTIJD=3 AND AGGSB201=1 AND BBGMM408=108 AND ALLOCHTOON=1)
STRATA=3.
IF (LEEFTIJD=1 AND AGGSB201=2 AND BBGMM408=108 AND ALLOCHTOON=1)
STRATA=4.
IF (LEEFTIJD=2 AND AGGSB201=2 AND BBGMM408=108 AND ALLOCHTOON=1)
STRATA=5.
IF (LEEFTIJD=3 AND AGGSB201=2 AND BBGMM408=108 AND ALLOCHTOON=1)
STRATA=6.
IF (LEEFTIJD=1 AND AGGSB201=1 AND BBGMM408=108 AND ALLOCHTOON=2)
STRATA=7.
IF (LEEFTIJD=2 AND AGGSB201=1 AND BBGMM408=108 AND ALLOCHTOON=2)
STRATA=8.
IF (LEEFTIJD=3 AND AGGSB201=1 AND BBGMM408=108 AND ALLOCHTOON=2)
STRATA=9.
IF (LEEFTIJD=1 AND AGGSB201=2 AND BBGMM408=108 AND ALLOCHTOON=2)
STRATA=10.
IF (LEEFTIJD=2 AND AGGSB201=2 AND BBGMM408=108 AND ALLOCHTOON=2)
STRATA=11.
IF (LEEFTIJD=3 AND AGGSB201=2 AND BBGMM408=108 AND ALLOCHTOON=2)
STRATA=12.
(AGGSB201=geslacht, BBGMM408=gemeente)
```

a) Bestand klaarmaken voor analyses (figuur 3)

- Ga in SPSS naar Analyze – Complex Samples – Prepare for Analysis.
- Kies voor 'Create a plan file' en browse naar de plek waar dit moet worden opgeslagen en geef het een naam.
- Klik op 'Next'.

Figuur 3. Schermvoorbeeld complex samples: klaarmaken voor analyses

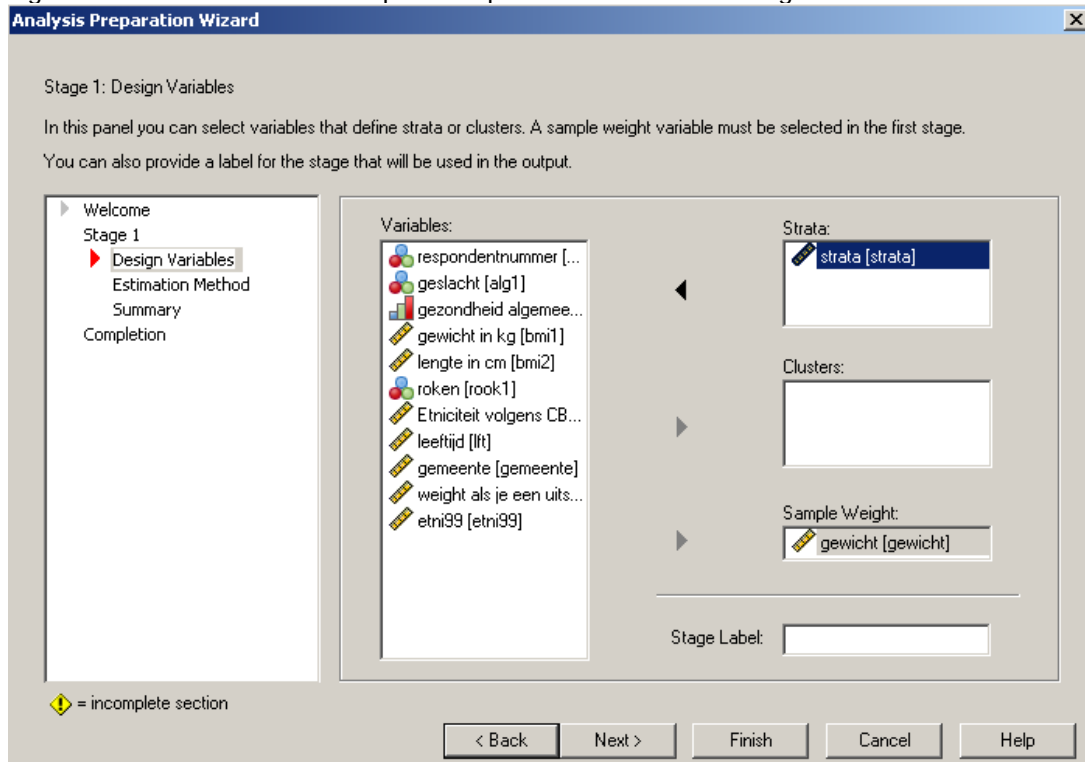


- Selecteer de variabele die de strata weergeeft (figuur 4).
- Selecteer de variabele die het gewicht weergeeft (gebruik hier de grote weegfactor W_i). (Clusters hoeft alleen te worden ingevuld als u bijvoorbeeld uit elke klas een steekproef heeft getrokken).
- Klik op 'Next'.
- Kies voor 'WR' (sampling with replacement) (figuur 5).
- Klik op 'Next'.

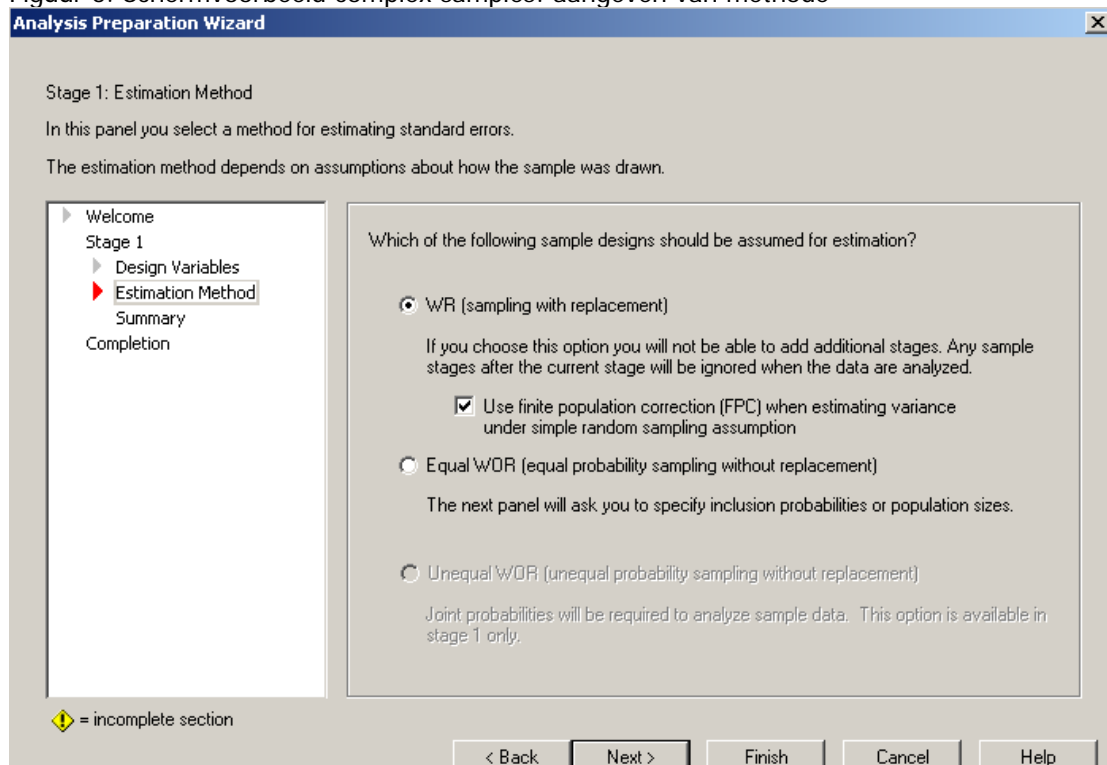
Het volgende scherm geeft een samenvatting van de structuur.

- Klik op 'Next'.
- Plak de tekst in de syntax en klik op 'Finish' (figuur 6).

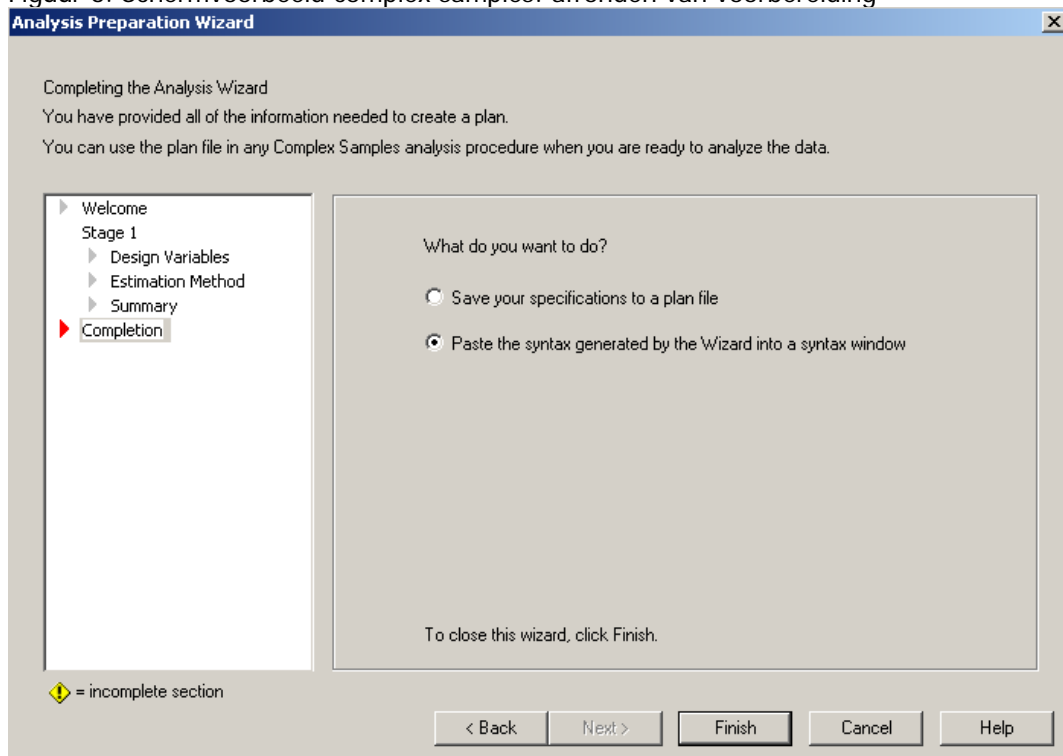
Figuur 4. Schermvoorbeeld complex samples: selecteren van weegfactoren en strata



Figuur 5. Schermvoorbeeld complex samples: aangeven van methode



Figuur 6. Schermvoorbeeld complex samples: afronden van voorbereiding

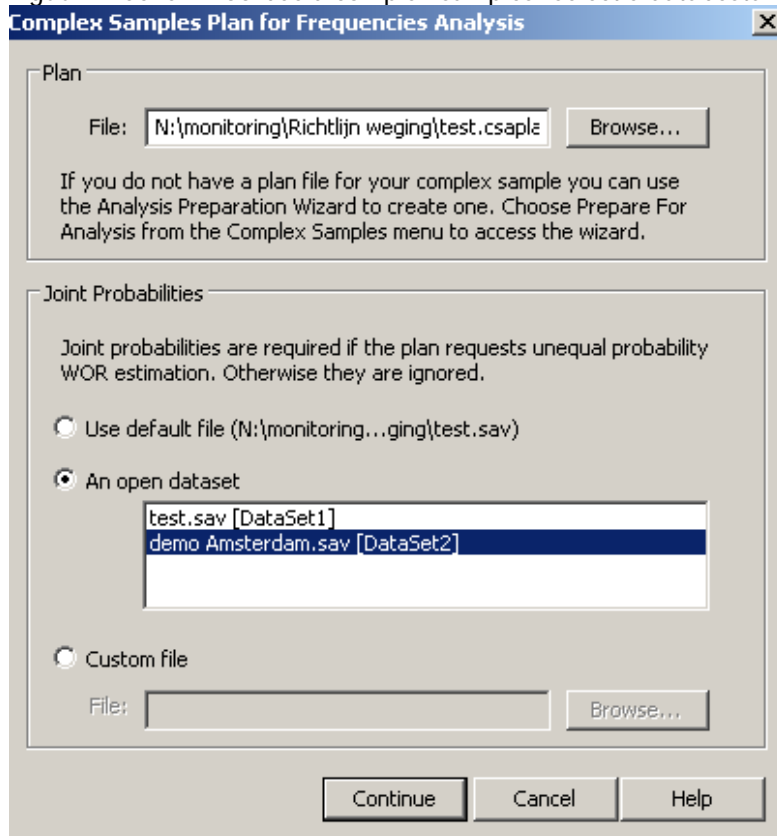


- Run de syntax.

b) Betrouwbaarheidsintervallen en toetsen van verschillen tussen gemeenten

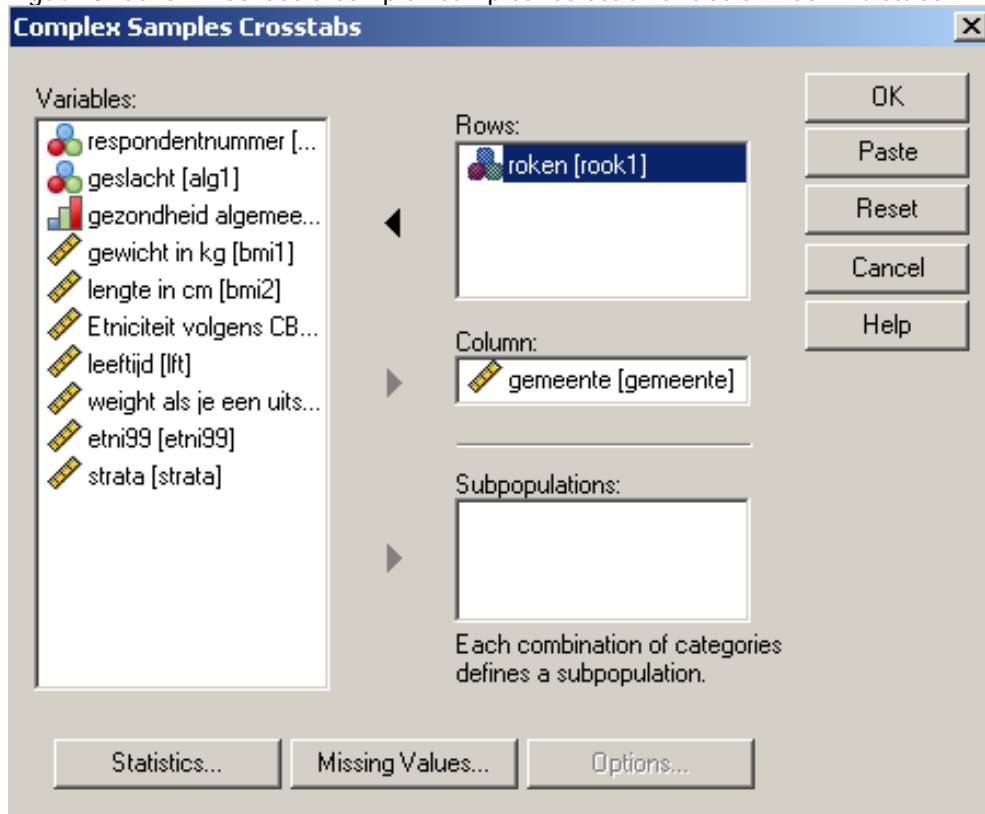
- Ga in SPSS naar Analyze – Complex Samples – Crosstabs.
- Selecteer het bestand waarin de structuur is aangegeven (figuur 7).
- Selecteer het databestand.
- Klik op 'Continue'.

Figuur 7. Schermvoorbeeld complex samples: selectie databestand



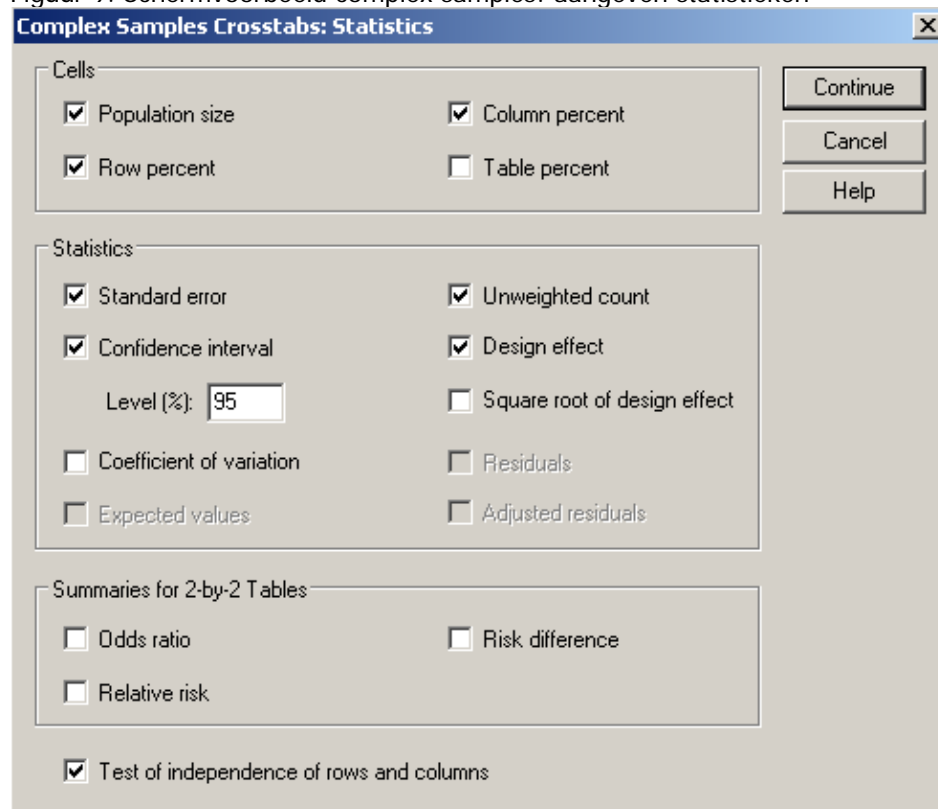
- Selecteer de variabele(n) waarvoor u frequenties wilt uitdraaien (figuur 8).
- Selecteer de groepen die u wilt vergelijken (in dit geval 'gemeente').
- Klik op 'Statistics'.

Figuur 8. Schermvoorbeeld complex samples: selectie variabelen voor kruistabel



- Selecteer 'Row percent', 'Column percent', 'Confidence Interval', 'Unweighted count', 'Design effect' en 'Test of independence' (= Chi kwadraat) (figuur 9).

Figuur 9. Schermvoorbeeld complex samples: aangeven statistieken



Interpretatie van de output:

- De resultaten (frequenties per gemeente met betrouwbaarheidsintervallen) zijn te vinden achter '% within gemeente' (zie figuur 10).

Figuur 10. Output complex samples: verschillen tussen groepen (1)

			roken * gemeente					Total	
			gemeente						
			gemeente a	gemeente b	gemeente c	gemeente d	gemeente e		
Population Size	Estimate		69046,000	49346,000	35878,000	18352,000	15898,000	188520,0	
	Standard Error		7407,262	4733,293	3179,612	2094,140	2287,182	9848,764	
	95% Confidence Interval	Lower		54517,566	40062,227	29641,582	14244,602	11411,973	169202,9
		Upper		83574,434	58629,773	42114,418	22459,398	20384,027	207837,1
	Design Effect			2,530	1,390	,841	,690	,946	2,158
	Unweighted Count			86	121	129	74	51	461
% within roken	Estimate		36,6%	26,2%	19,0%	9,7%	8,4%	100,0%	
	Standard Error		2,8%	2,2%	1,7%	1,1%	1,2%	,0%	
	95% Confidence Interval	Lower		31,3%	22,1%	16,0%	7,7%	6,4%	100,0%
		Upper		42,3%	30,7%	22,5%	12,2%	11,1%	100,0%
	Design Effect			1,768	1,329	,940	,751	,967	.
	Unweighted Count			86	121	129	74	51	461
% within gemeente	Estimate		34,4%	37,0%	37,6%	26,9%	23,7%	33,4%	
	Standard Error		3,7%	3,5%	3,3%	3,1%	3,4%	1,7%	
	95% Confidence Interval	Lower		27,6%	30,4%	31,4%	21,3%	17,7%	30,1%
		Upper		41,9%	44,1%	44,2%	33,3%	31,0%	36,9%
	Design Effect			3,375	1,956	1,225	,910	1,195	2,133
	Unweighted Count			86	121	129	74	51	461

- De p-waarde staat in de kolom 'Sig' en laat zien dat de verschillen tussen de gemeenten significant zijn.

Figuur 11. Output complex samples: verschillen tussen groepen (2)

Tests of Independence						
		Chi-Square	Adjusted F	df1	df2	Sig.
roken * gemeente	Pearson	104,723	8,299	6,688	11235,778	,000
	Likelihood Ratio	101,934	8,078	6,688	11235,778	,000

Ter controle:

De frequenties per gemeente kunt u ook uitrekenen door middel van Crosstabs (buiten Complex Samples om) en dan de weegfactor aanzetten via 'Data – Weight cases'. De frequenties zijn hetzelfde als wanneer deze via Complex Samples worden berekend. Het significantieniveau verandert echter; p-waardes worden groter wanneer het designeffect wordt meegenomen.

Berekenen van prevalentie cijfers voor de regio

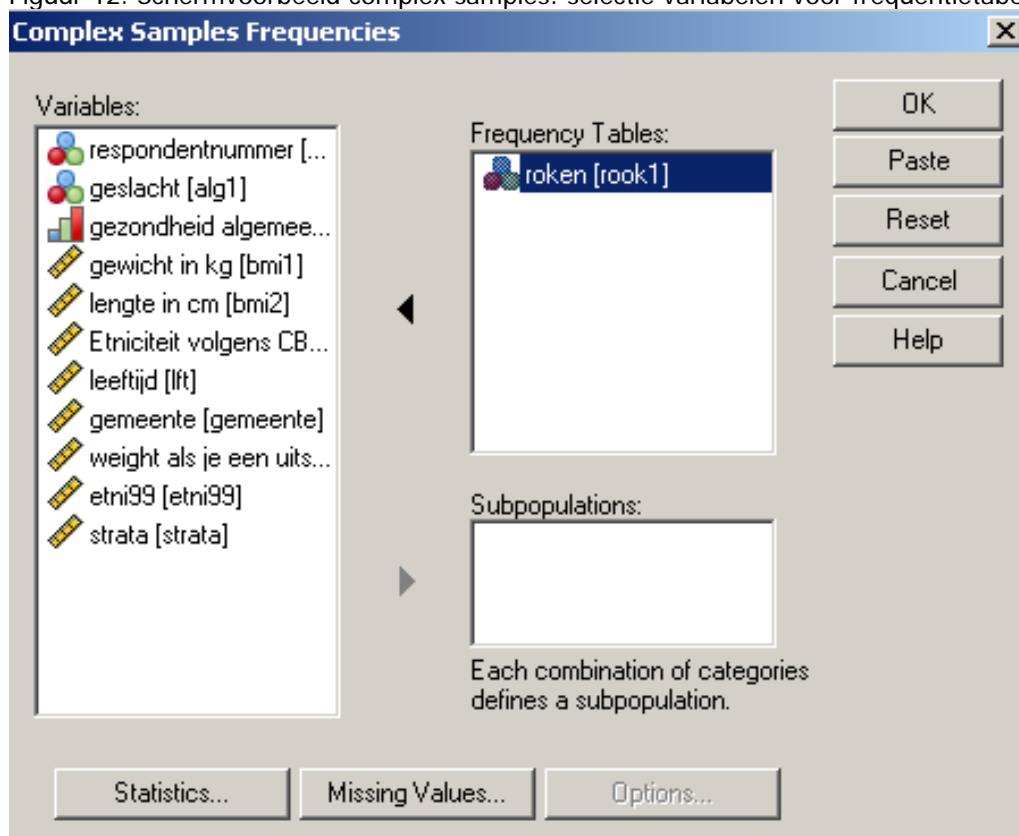
Zorg ervoor dat het bestand is klaargemaakt voor analyses, zoals dat in stap 6a staat beschreven.

- Ga vervolgens in SPSS naar Analyze – Complex Samples – Frequencies.
- Selecteer het bestand waarin de structuur is aangegeven.
- Selecteer het databestand.
- Klik op 'Continue'.

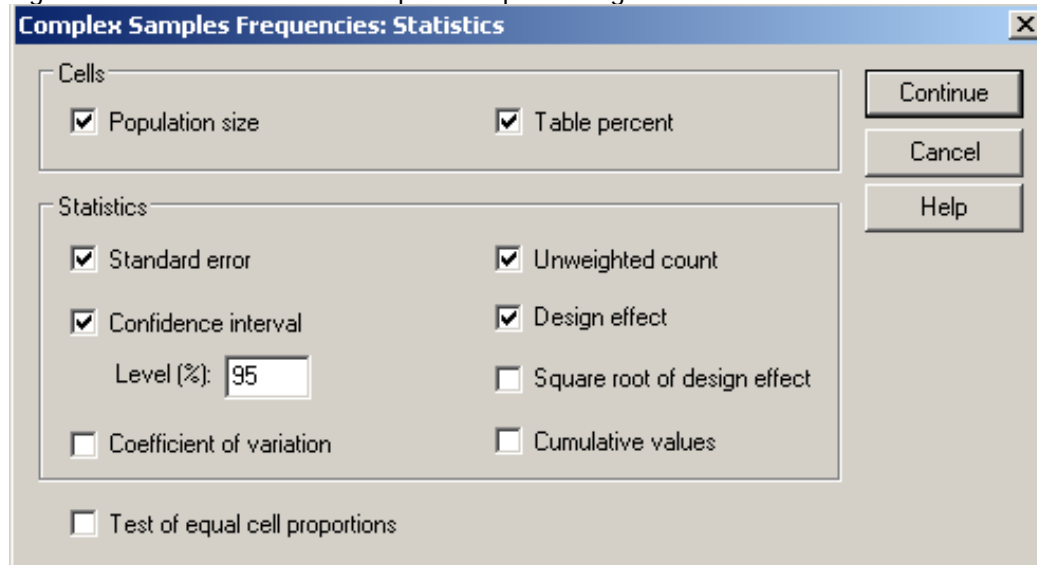
- Selecteer de variabele(n) waarvoor u frequenties wilt uitdraaien (figuur 12).
- Klik op 'Statistics'.

- Selecteer 'Table percent', 'Confidence interval', 'Unweighted count' en 'Design effect' (figuur 13).

Figuur 12. Schermvoorbeeld complex samples: selectie variabelen voor frequentietabel



Figuur 13. Schermvoorbeeld complex samples: aangeven statistieken



Interpretatie van de output:

- De kolom 'Estimate' geeft het geschatte aantal rokers in de totale populatie (figuur 14). De totale populatie (in kolom 'Estimate', achter 'Total') moet gelijk zijn aan het aantal personen in de totale onderzoekspopulatie.
- De kolom 'Unweighted Count' geeft het aantal rokers weer in de respondenten groep.
- De resultaten staan in de kolom Estimate, achter '% of Total'.

Figuur 14. Output complex samples: prevalentiecijfers en betrouwbaarheidsintervallen

		roken					
		Estimate	Standard Error	95% Confidence Interval		Design Effect	Unweighted Count
				Lower	Upper		
Population Size	ja	188520,0	9848,764	169202,9	207837,1	2,158	461
	nee, vroeger wel	143219,0	8102,880	127326,2	159111,8	1,716	419
	nee, nooit gerookt	233039,0	10079,016	213270,2	252807,8	2,074	695
	Total	564778,0	3324,658	558257,1	571298,9	.	1575
% of Total	ja	33,4%	1,7%	30,1%	36,9%	2,133	461
	nee, vroeger wel	25,4%	1,4%	22,7%	28,3%	1,697	419
	nee, nooit gerookt	41,3%	1,8%	37,8%	44,8%	2,034	695
	Total	100,0%	,0%	100,0%	100,0%	.	1575

Wanneer met de grote weegfactoren wordt gewerkt, wordt dus automatisch al rekening gehouden met het aantal inwoners per gemeente.

Bij het gebruik van de kleine weegfactoren voor de berekening van regionale cijfers, moet rekening worden gehouden met het aantal inwoners per gemeente. Daartoe worden de kleine weegfactoren die eerder zijn berekend op gemeenteniveau, vermenigvuldigd volgens onderstaande formule.

Kleine gewicht regio = kleine gewicht gemeente * ((bevolking gemeente/steekproef gemeente)/(bevolking regio/steekproef regio)).

Voor het berekenen van betrouwbaarheidsintervallen moet de standaardfout worden vermenigvuldigd met de designfactor (zie output figuur 2).

Dankwoord

Bij het tot stand komen van deze richtlijn zijn verschillende mensen betrokken geweest. Wij willen hen allen heel hartelijk danken. Dr. Hendriek Boshuizen van het Expertisecentrum voor Methodes en Informatie van het RIVM heeft haar expertise gedeeld en adviezen gegeven. Daarnaast danken we prof. Jelke Bethlehem van het CBS voor het kritisch bekijken van het geschreven stuk. En ten slotte danken we alle GGD-epidemiologen die input hebben geleverd voor deze richtlijn.

Bijlage 1. Methode

De richtlijn is als volgt tot stand gekomen: Allereerst is er geïnventariseerd bij GGD-epidemiologen hoe zij doorgaans met wegen omgaan. Daarnaast is er een pilot uitgevoerd op data om te bekijken wat de effecten zijn van verschillende weegmethoden.

Een werkgroep bestaande uit GGD-epidemiologen en onderzoekers van het RIVM zijn vier keer bij elkaar geweest om de richtlijn te schrijven. De andere GGD-epidemiologen is om input gevraagd via GGD-kennisnet en er is een themamiddag over de richtlijn gehouden. Ook is het CBS geraadpleegd.

Inventarisatie bij GGD'en

Bij GGD-epidemiologen is nagevraagd hoe zij doorgaans met wegen omgaan. Hiertoe is een e-mail bericht verstuurd naar alle hoofden Epidemiologie en is een bericht geplaatst op GGD-kennisnet met de vraag hoe GGD-epidemiologen de steekproef doorgaans trekken (stratificatie/ fractie) en hoe hun weegfactor tot stand komt. Daarnaast is gevraagd wat bekend is over de non-respondenten. De GGD-epidemiologen konden ook aangeven tegen welke problemen zij aanlopen bij het wegen en welke vragen zij graag beantwoord zien in de richtlijn.

Pilot

Door het RIVM is een pilot uitgevoerd. Het doel van de pilot was om te onderzoeken wat er met de data gebeurt als je weegt. Eerst is onderzocht voor welke factoren zou moeten worden gewogen. Daarvoor is voor bepaalde factoren (zoals leeftijd, geslacht, etniciteit, burgerlijke staat) bekeken of de respons verschilt in de verschillende categorieën. Ook is gekeken naar verbanden tussen die factoren en een aantal uitkomstvariabelen.

Voor de pilot worden de gegevens van de gezondheidspeiling voor volwassenen van Flevoland 2005 gebruikt. Er is gekozen voor Flevoland, omdat die GGD de hele provincie dekt en op de website van het CBS de verdelingen van de potentiële weegfactoren per provincie bekend zijn.

Werkgroep

Er is een werkgroep opgericht om te spreken over ervaringen bij het wegen en na te denken over de richtlijn. Deze werkgroep bestond uit vier GGD-epidemiologen en twee medewerkers van het RIVM (projectleider Lokale en Nationale Monitor Gezondheid en een senior adviseur biostatistiek).

Bijlage 2. Resultaten inventarisatie

Door 15 personen is een reactie gegeven op de oproep informatie te verstrekken over stratificatie, wegen en respons. Deze personen vertegenwoordigen soms meerdere GGD'en. Daarnaast kon voor twee GGD'en de gevraagde informatie gehaald worden uit de informatie die ze ingevuld hadden bij het aanleveren van hun databestand op www.monitoregezondheid.nl. De meeste beschrijvingen gaan over onderzoek onder volwassenen en ouderen. Slechts drie keer gaat het expliciet over jeugd. Vanuit de werkgroep kwamen nog aanvullingen op de verkregen informatie.

Stratificatie steekproef

De gebruikte methoden verschillen niet veel tussen de GGD'en. De steekproeven worden vaak gestratificeerd naar gemeente. Sommige GGD'en trekken uit elke gemeente dezelfde aantallen. Soms worden bepaalde gemeenten juist extra opgehoogd om voldoende aantallen te hebben om betrouwbare uitspraken te kunnen doen. Sommige GGD'en stratificeren ook naar leeftijd. Leeftijd wordt dan ingedeeld in groepen. De kleine groep van 85 jaar en ouder wordt bijvoorbeeld opgehoogd. In grotere steden wordt soms ook nog gestratificeerd naar wijk en etniciteit. Bepaalde groepen worden dan opgehoogd.

Wegen

De meeste GGD'en wegen hun gegevens alvorens de resultaten te presenteren. Sommige GGD'en wegen niet, maar presenteren de gegevens uitgesplitst naar leeftijd en geslacht.

De GGD'en wegen in ieder geval voor de variabelen waarnaar de steekproef is gestratificeerd. Geslacht en leeftijd worden daar vaak nog aan toegevoegd. Veel GGD'en delen leeftijd in groepen in. Hiervoor worden verschillende indelingen gebruikt: * 19-49, 50-74, 75+; * 19-34, 35-49, 50-64, 65-74, 75+; en voor ouderen: * 65-75, 75+; * 55-64, 64-74, 75+; * 65-74, 75-84, 85+.

Voor het berekenen van cijfers op regio- of provincieniveau wordt er gewogen naar het inwoneraantal van de gemeenten (gemeente = stratificatie variabele).

Een aantal GGD'en geeft een weegfactor van 1 aan mensen van wie leeftijd en/of geslacht onbekend is.

Bij het berekenen van cijfers op gemeenteniveau worden door de GGD'en meestal 6 groepen gevormd, gebaseerd op 2 categorieën voor geslacht en 3 leeftijdscategorieën.

Jeugd

In E-movo 2003 werd niet gewogen, omdat alle scholen meededen en de respons nauwelijks verschilde tussen scholen of niet selectief was. Ook door een andere GGD werd aangegeven dat er niet werd gewogen bij een onderzoek onder tweede- en vierdeklassers, omdat de verdeling van geslacht, klassen en opleidingsniveau niet bekend was op populatieniveau.

Voor een jongere leeftijdsgroep werd gewogen naar leeftijd met onderscheid tussen 4-8 en 9-12 jaar.

Andere GGD'en geven aan voor jeugd op dezelfde manier te wegen als bij volwassenen, namelijk naar gemeente, geslacht en leeftijd.

Non-respondenten

Het responspercentage ligt ongeveer tussen de 45 en 65%. Non-respondenten zijn vaker van het mannelijke geslacht, in de leeftijdsgroep 19-34 jaar en behorende tot een etnische minderheidsgroep. Over de respons onder de oudste senioren (75+) worden verschillende bevindingen gemeld: de één geeft aan dat de respons in deze groep hoger is, de ander juist lager.

De gegevens van een GGD die verschillende reminders stuurde, lieten zien dat late respondenten een slechtere ervaren gezondheid en leefstijl hebben en meer gebruik maken van geestelijke gezondheidszorg. Ook een non-respons onderzoek bij een andere GGD toonde een slechtere ervaren gezondheid, meer beperkingen en meer gebruik van voorzieningen onder non-respondenten ten opzichte van respondenten.

Bijlage 3. Pilot

Respons

In Flevoland zijn gemiddeld genomen de respondenten uit de jongste leeftijdsgroep (24-30 jaar) ondervertegenwoordigd. De groep 35-50 jarigen is oververtegenwoordigd. Daarnaast zijn respondenten vaker vrouwen (verschil tussen mannen en vrouwen is gemiddeld 9%), hoger opgeleid en vaker van autochtone afkomst ten opzichte van de totale populatie. Ook zijn de respondenten vaker gehuwd en zijn de ongehuwden ondervertegenwoordigd.

Uitkomsten

In Flevoland is voor elke gemeente een zelfde fractie willekeurig uit de populatie getrokken, zonder verdere stratificatie.

In de pilot zijn verschillende uitkomsten gewogen; in eerste instantie naar leeftijd en geslacht. Geslacht is ingedeeld in 2 categorieën en leeftijd in 7 categorieën die elk vijf jaar beslaan. Het blijkt dat wegen naar leeftijd en geslacht bij deze steekproef weinig effect heeft op de schattingen van de regio gemiddeldes en de bijbehorende onzekerheden, bijvoorbeeld voor fruitconsumptie. Wanneer daarnaast ook nog gewogen wordt naar burgerlijke staat, veranderen de schattingen en onzekerheden enigszins. Op gemeenteniveau worden echter wel effecten van weging naar leeftijd en geslacht gevonden. Als bijvoorbeeld in Almere gewogen wordt naar leeftijd en geslacht blijkt dat de schattingen voor de gemiddeldes en de standaardfouten wel degelijk veranderen t.o.v. de ongewogen schattingen. Verder is voor Almere gekeken wat er gebeurt als gewogen wordt naar etniciteit. Etniciteit is voor deze analyse onderverdeeld in 2 klassen: allochtoon en autochtoon. De schattingen veranderen dan opnieuw.

Uit de pilot wordt geconcludeerd dat leeftijd, geslacht, burgerlijke staat en etniciteit een invloed hebben op de wegingsprocedure. Deze variabelen hangen dus samen met de uitkomstvariabelen én er zijn verschillen tussen respondenten en de hele populatie op basis van deze variabelen.

Tenslotte is nagegaan wat het effect is van de onderverdeling van leeftijd in meer of minder categorieën. Wanneer leeftijd wordt opgedeeld in 3 leeftijdsklassen (19-34, 35-49 en 50-55) verandert er niets aan de gevonden schattingen t.o.v. de eerdere opdeling van leeftijd in 5-jaars klassen.

We realiseren ons dat de conclusies uit de pilot alleen gelden voor de onderzochte variabelen in dit databestand. Andere uitkomstmaten en andere databestanden zouden tot andere conclusies kunnen leiden.

Bijlage 4. Rekenvoorbeelden

Zie aparte Excel sheets met rekenvoorbeelden. Deze zijn te vinden op www.ggdkennisnet.nl.

Bijlage 5. Steekproef is gestratificeerd getrokken

Als u de steekproef heeft gestratificeerd naar andere variabelen dan leeftijd, geslacht, etniciteit en gemeente, moet de weegmethode worden uitgebreid.

Stel u heeft de steekproef gestratificeerd naar wijk. Het beste zou dan zijn om per wijk een tabel te maken met de aantallen uitgesplitst naar leeftijd, geslacht en etniciteit.

Dit is alleen mogelijk als de aantallen van de totale populatie per wijk beschikbaar zijn, uitgesplitst naar leeftijd, geslacht en etniciteit.

Prevalentie cijfers voor de totale gemeente kunt u dan op dezelfde manier berekenen als voor de regio.

Het alternatief is om de weegfactor voor elke persoon in wijk x te vermenigvuldigen volgens onderstaande formule, zodat er gewichten voor de gemeente ontstaan:

Kleine gewicht gemeente = kleine gewicht wijk * ((bevolking wijk/steekproef wijk)/(bevolking gemeente/steekproef gemeente)).



Bijlage 6. Gebruik Richtlijn Weging in Lokale en Nationale Monitor Gezondheid

Doel

Dit document beschrijft hoe de Richtlijn Weging, die in 2008 vanuit het Programma Beter Voorkomen is ontwikkeld, gebruikt wordt in de Lokale en Nationale Monitor Gezondheid.

Achtergrond

Het RIVM heeft in opdracht van GGD Nederland samen met een aantal GGD-epidemiologen een richtlijn ontwikkeld voor het wegen van gegevens die door GGD'en zijn verzameld in het kader van hun periodieke lokale gezondheidsmonitor.

Hoewel deze richtlijn in eerste instantie een hulpmiddel is voor GGD-epidemiologen, is de doelstelling van de richtlijn ook dat deze bijdraagt aan uniformiteit tussen GGD'en en dient de richtlijn als uitgangspunt bij de berekening van landelijke cijfers door het RIVM.

Gebruik Richtlijn Weging door GGD-epidemiologen en RIVM

De Richtlijn Weging heeft dezelfde status als de standaardvraagstellingen die voor de Lokale en Nationale Monitor Gezondheid zijn vastgesteld. Dat betekent dat wordt afgesproken dat als GGD'en hun gegevens analyseren, ze de weging uitvoeren zoals beschreven in de richtlijn.

Het RIVM zal bij het berekenen van landelijke cijfers naar dezelfde variabelen wegen als die in de richtlijn staan beschreven. Aanvullend moet echter ook rekening worden gehouden met de verdeling van de beschikbare gegevens over het land, waardoor mogelijk ook naar andere factoren moet worden gewogen, zoals urbanisatiegraad.

Over het naar buiten brengen van regionale cijfers door het RIVM (bijvoorbeeld in Atlas), worden aanvullende afspraken gemaakt in de Samenwerkingsovereenkomst Lokale en Nationale Monitor Volksgezondheid. Als het RIVM regionale cijfers gaat berekenen doet zij dat volgens de richtlijn. Als de GGD'en dat ook doen, zouden de cijfers die door het RIVM worden berekend gelijk moeten zijn aan degene die in de lokale of regionale rapportages staan.

Aanleveren databestanden op www.monitorgezondheid.nl

GGD'en leveren de verzamelde gegevens aan op de website www.monitorgezondheid.nl. In het databestand dat wordt geupload staat een kolom waarin de weegfactor moet worden toegevoegd. Dit moet de 'grote' weegfactor zijn, die per respondent aangeeft hoeveel personen in de populatie hij representeert. Deze weegfactoren moeten volgens de richtlijn worden bepaald. In principe kan het RIVM op basis van de postcodes de weegfactoren zelf berekenen, maar omdat GGD'en deze weegfactoren toch al berekenen voor de lokale analyses wordt aan de GGD'en gevraagd deze aan te leveren.

Vaststelling van het gebruik van de Richtlijn Weging in de Lokale en Nationale Monitor Gezondheid

Dit document over het gebruik van de richtlijn voor de Lokale en Nationale Monitor Gezondheid is op dezelfde manier vastgesteld als de standaardvraagstellingen. Dat betekent dat het eerst is goedgekeurd door de Vakgroep Epidemiologie en vervolgens door het bestuur van GGD Nederland.



GGD Nederland
Adriaen van Ostadelaan 140

Postbus 85300
3508 AH Utrecht

T 030 252 30 04
F 030 251 18 69

postbus@ggd.nl
www.ggd.nl



Het landelijke kwaliteitsprogramma Beter voorkomen wordt uitgevoerd door GGD Nederland, de VNG en ActiZ, organisatie van zorgondernemers, in opdracht van het ministerie van VWS en onder regie van ZonMw.