

# Regionale verschillen in sterfte verklaard

Rob Loke en Andries de Jong<sup>1</sup>

*Het PBL en het CBS hebben in 2011 voor de vierde keer de Regionale bevolkings- en huishoudensprognose met het model PEARL uitgebracht. In deze prognose spelen veronderstellingen over de gemeentelijke sterfte een belangrijke rol. In dit kader is een analyse uitgevoerd naar de verklaring van gemeentelijke verschillen in de levensverwachting aan de hand van een multivariaat regressiemodel. Uit de analyse blijkt dat demografische, culturele/leefstijl, sociaaleconomische en geografische variabelen alle een belangrijke rol spelen bij de verklaring van de gemeentelijke levensverwachting. Van de sociaaleconomische variabelen is met name het percentage mensen met een uitkering van belang. In gemeenten waar dit percentage hoog ligt, is de levensverwachting lager. Religie (cultuur) heeft een klein positief effect: in gemeenten met een hoog percentage stemmers op de kleine christelijke partijen ligt de levensverwachting van vrouwen iets hoger. De leefstijl is geoperationaliseerd aan de hand van sterfte aan diverse doodsoorzaken. In gemeenten met hoge sterftecijfers aan ziekten van het hart- en vaatstelsel en ziekten van het ademhalingsstelsel ligt de levensverwachting lager. Van de demografische variabelen is enkel voor de mannen sprake van een significant verband: in gemeenten met een laag percentage gehuwde mannen ligt de levensverwachting lager*

## 1. Inleiding

In de PBL/CBS regionale prognose vormt sterfte een belangrijk onderdeel. In Nederland overlijden de laatste jaren gemiddeld rond 135 duizend mensen. De sterftegevallen zullen de komende decennia nog fors toenemen door de acceleratie van de vergrijzing, vanwege het 'op leeftijd komen' van de naoorlogse babyboom. In de regionale prognose speelt niet alleen deze groei van de sterfteaantallen een belangrijke rol, maar ook de regionale verschillen in het aantal sterftegevallen. Om dit goed te kunnen voorspellen wordt in de prognose gekeken naar regionale verschillen in de sterftekansen. Hierbij wordt gebruik gemaakt van een samenvattende indicator van leeftijdsspecifieke sterftekansen, te weten de levensverwachting bij de geboorte. Deze indicator geeft een indruk van het aantal jaren dat iemand bij geboorte naar verwachting te leven heeft (indien de in een bepaald kalenderjaar gemeten leeftijdsspecifieke sterftekansen voor deze persoon gedurende zijn leven van toepassing zijn). In 2009 was de levensverwachting voor mannen 78,5 jaar en voor vrouwen 82,7 jaar.

In de afgelopen vier decennia is de levensverwachting in Nederland aanzienlijk gestegen. Tussen 1970 en 2009 nam de verwachting voor mannen met zo'n 7,7 jaar toe, voor vrouwen met zo'n 6,2 jaar. Voor de mannen geldt een min of meer continue stijging, terwijl de stijging bij de vrouwen na 1985 nog maar vrij gering was. Volgens de Jong (2004) kan de stijging van de levensverwachting voor een belangrijk deel op het conto van een daling van hart- en vaatziekten worden geschreven. Deze daling leverde voor mannen en vrouwen een stijging op van bijna drie jaar. Daarnaast speelt de daling van de sterfte door nieuwvormingen (kanker e.d.), niet-natuurlijke dood en overige doodsoorzaken een rol, zij het veel geringer. Het valt te verwachten dat doodsoorzaken niet alleen trends in de tijd kunnen verklaren, maar ook regionale verschillen in de sterfte.

In de PBL/CBS regionale prognose (zie: [www.regionalebevolkingsprognose.nl](http://www.regionalebevolkingsprognose.nl) en [statline.cbs.nl](http://statline.cbs.nl)) worden ten behoeve van het modelleren van de sterfte veronderstellingen geformuleerd over regionale verschillen in de levensverwachting. Hiertoe wordt gebruik gemaakt van een multivariaat regressiemodel, waarmee wordt getracht de levensverwachting op gemeentelijk niveau te relateren aan diverse achtergrondvariabelen. In het model dat bij de vorige prognose is gebruikt (zie de Jong

---

<sup>1</sup> Rob Loke was tot november 2011 werkzaam bij het Planbureau voor de Leefomgeving en Andries de Jong is hier ook werkzaam.

e.a., 2005), werden deze achtergrondvariabelen onderscheiden in vier categorieën, te weten demografische, cultuur/leefstijl, sociaaleconomische en geografische. In dit artikel wordt het model opnieuw geschat aan de hand van recente data. Voorts is het model aangepast doordat nu ook variabelen die betrekking hebben op doodsoorzaken zijn meegenomen.

## **2. Achtergronden bij regionale verschillen in de sterfte**

Uit diverse studies (o.a. Marmot Review, 2010; Commission on Social Determinants of Health, 2008; Smith e.a. 1992) blijkt dat er een relatie bestaat tussen sociaaleconomische deprivatie, slechte gezondheid en voortijdige sterfte. Smits e.a. (2001) laat voor Nederland op postcode-niveau zien dat er een relatie bestaat tussen sociaaleconomische status en de sterftetekans: in postcodegebieden met een lage sociaaleconomische status ligt de sterftetekans beduidend hoger. Anson (2003) laat zien dat in België de (gestandaardiseerde) sterfte in gemeenten met een hoge sociaaleconomische status lager ligt. Van der Gaag en de Jong (1997) melden dat in arme regio's in Schotland en het noorden van Frankrijk de sterftekansen hoger liggen.

Het effect van de sociaaleconomische status op de sterftetekans wordt voor een belangrijk gedeelte veroorzaakt door de leefstijl (Minkler e.a., 2006; van der Lucht en Picavet, 2006; Mackenbach e.a., 2008). Mensen met een lage sociaaleconomische status leven vaak ongezonder dan mensen met een hoge status. Hierbij spelen vooral roken, (verkeerd) eten en (alcohol) drinken een belangrijke rol. Het roken is een belangrijke oorzaak van diverse kankers (zoals long- en keelkanker). Het rookgedrag is gerelateerd aan sociale klasse. Mede onder invloed van overheids campagnes om met roken te stoppen, zijn mensen met een hogere sociaaleconomische status vaker gestopt met roken. Het roken is ook een belangrijke verklaring voor het geslachtsverschil in de levensverwachting (Valkonen en Van Poppel, 1997). Mannen zijn eerder gaan roken dan vrouwen zodat de sterfte aan longkanker bij mannen veel sterker was dan bij vrouwen. Grote hoeveelheden alcohol kunnen leiden tot bepaalde typen van kanker, leverziekten en ongelukken. In landen die gekenmerkt worden door een hoge alcoholconsumptie, zoals de voormalige Sovjet Unie (Dinkel, 1985) en Finland ligt de levensverwachting laag. Een leefstijlfactor die snel in belang groeit, betreft overgewicht/obesitas. Zo rapporteert McPherson e.a. (2009) een sterke groei hiervan in het Verenigd Koninkrijk. Obesitas vergroot het risico op diverse ziektes, en in het bijzonder op diabetes (zie o.a. Prospective Studies Consortium, 2009). Uit de studie van Caselli e.a. (2003) komt naar voren dat dieet een belangrijke factor is voor de verklaring van ontwikkelingen in regionale sterfteverschillen. In het noorden van Italië is de sterfte in snel tempo gedaald na het overnemen van het 'mediterrane dieet' (veel vis en olijfolie) dat in het zuiden al gangbaar was.

Directe operationalisatie van leefstijlvariabelen aan de hand van roken, alcoholgebruik en overgewicht is in de analyse van gemeentelijke verschillen in de levensverwachting niet mogelijk, omdat deze gegevens slechts aanwezig zijn op GGD niveau (30 regio's) in de databank StatLine van het CBS. Aangezien het aannemelijk is dat er een oorzakelijk verband bestaat tussen leefstijl enerzijds en doodsoorzaken anderzijds, is ervoor gekozen om leefstijl te operationaliseren aan de hand van de belangrijkste doodsoorzaken, die ontleend zijn aan de classificatielijst van de International Classification of Diseases ICD-10 (World Health Organization, 2007). Ook in de studie van Pos en Bouwens (2004) wordt er een verband gelegd tussen leefstijl en doodsoorzaken; zij geven aan dat de kans op hart- en vaatziekten groter wordt door ongezonde voeding, overgewicht en roken.

Niet alleen de sociaaleconomische status en de leefstijl, maar ook de burgerlijke staat heeft invloed op de sterftetekansen. Het huwelijk heeft vooral gunstige effecten op de levensverwachting van mannen (Zick en Smith, 1991; Rogers, 1995). Uit de Jong (2002) blijkt dat de levensverwachting van vijftigjarige gehuwde mannen rond 4 jaar hoger ligt dan die van ongehuwde mannen. Bij vijftigjarige gehuwde vrouwen ligt de levensverwachting rond 2 jaar hoger ligt dan die van

ongehuwde vrouwen. Anson (2003) brengt de hogere levensverwachting van gehuwden in verband met het hebben van sociale bindingen en geeft aan dat mannen hier 'gevoeliger' voor zijn.

Volgens Mackenbach e.a. (1991) en Mackenbach (1992) kunnen regionale sterfteverschillen in Nederland in belangrijke mate worden herleid tot verschillen in een drietal sociaaldemografische kenmerken, namelijk het gemiddeld inkomen, het percentage inwoners in grote steden en het percentage mensen met een rooms-katholieke achtergrond. Godsdienst heeft geen directe invloed op gezondheid, maar gaat samen met verschillen in culturele gedragspatronen die van invloed zijn op de gezondheid. Zo is er sprake van een hogere prevalentie van roken onder rooms-katholieken. Het gemiddeld inkomen hangt negatief samen met de sterfte aan maagkanker, ischemische hartziekten en verkeersongevallen. In Italië constateren Caselli e.a. (2003) een negatief verband tussen sterfte en het BNP per hoofd van de bevolking. In het welvarende noorden van Italië ligt de sterfte (aan hart- en vaatziekten en longkanker) echter hoger dan in het minder welvarende zuiden. De achterliggende verklaring wordt gezocht in het feit dat de vervuiling in het noorden groter is. Wilkinson en Pickett (2009) relateren sterfteverschillen tussen landen aan de mate van inkomensongelijkheid: hoe meer ongelijkheid, hoe hoger de sterfte.

In de steden ligt de levensverwachting doorgaans onder het landelijke gemiddelde. Ook in de vier grote gemeenten van Nederland is een lage levensverwachting van toepassing, zowel voor mannen als voor vrouwen. In Amsterdam, 's-Gravenhage en Utrecht ligt de levensverwachting rond één jaar onder het landelijke gemiddelde en in Rotterdam zelfs ruim anderhalf jaar. Ook op provinciaal niveau zijn er duidelijke verschillen in de levensverwachting zichtbaar. In de provincie Zeeland ligt de levensverwachting vrij hoog en in Limburg en Groningen vrij laag. De verhoogde sterfte in Limburg hangt samen met een hoge sterfte aan hart- en vaatziekten, terwijl de lage sterfte in Zeeland samenhangt met lagere sterfte aan hart- en vaatziekten en longkanker (Mackenbach e.a., 1991).

### **3. Beschrijving van methode en analyse**

#### *3.1 Levensverwachting als indicator voor regionale verschillen in sterfte*

Als indicator van het niveau van de sterfte is, net als in de Jong e.a. (2005), gekozen voor de levensverwachting bij geboorte. Deze maat is ongevoelig voor leeftijdsopbouweffecten, waardoor het zowel geschikt is voor het volgen van ontwikkelingen in de tijd, als het in kaart brengen van verschillen tussen regio's. Er is niet gekozen voor een meer 'technische' maat als het indirect gestandaardiseerde bruto sterftecijfer (die eveneens niet gevoelig is voor leeftijdsopbouweffecten), omdat de levensverwachting intuïtief beter te duiden valt. Een groot nadeel van het meten van de levensverwachting op regionaal niveau is echter dat de bevolkingsomvang in sommige regio's klein kan zijn en dat hierdoor het aantal sterftegevallen erg laag kan uitvallen (zoals onder de 50 overledenen). Hierdoor kan de levensverwachting in bepaalde regio's slechts gebrekkig worden berekend, mede omdat voor de berekening de sterftetekansen op de verschillende leeftijden benodigd zijn en die kansen door de geringe aantallen overledenen onbetrouwbaar kunnen zijn. Door het toepassen van filtertechnieken die zijn ontwikkeld in de ruimtelijke wetenschappen kan hier echter voor worden gecorrigeerd (zie hieronder). Door het filteren van de ruwe data is het dan mogelijk om een vrij betrouwbaar beeld te krijgen van ruimtelijke verschillen in de levensverwachting.

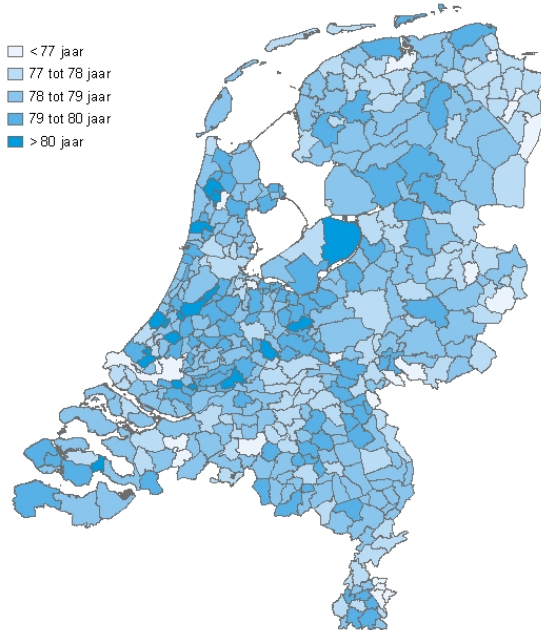
In de analyse is gebruik gemaakt van de levensverwachting van mannen en vrouwen per gemeente voor het kalenderjaar 2009. Via het filteren van de ruwe data is getracht de toevalsfluctuaties in de data terug te dringen om zodoende de betrouwbaarheid te vergroten. De ruwe data is hiertoe op drie manieren gefilterd. De eerste stap was het aanpassen van de levensverwachting in gemeenten met minder dan 50 sterfgevallen. Hierbij is gebruik gemaakt van de levensverwachting van de COROP waartoe de betreffende gemeenten behoren. Er is namelijk een gewogen gemiddelde berekend van de levensverwachting van elke betreffende gemeente en die van de COROP, op basis van de betreffende inwonertallen. Deze stap is gangbaar in de demografie. De tweede stap was het

'ruimtelijk' filteren van de regionale levensverwachtingen, wederom met het doel ruimtelijke toevalsfluctuaties te verwijderen in gemeenten met kleine inwoneraantallen. Hiervoor is een lokale variant van empirische Bayes schattingen gebruikt (Bailey en Gatrell, 1995; Waller en Gotway, 2004). In deze methode wordt de data voor een gemeente veranderd indien een lokaal berekende betrouwbaarheid te klein is; in dat geval wordt de score aangepast aan het lokaal berekende gemiddelde. In de derde stap werd gefilterd voor gemeentelijke uitbijters die te maken hebben met het gekozen kalenderjaar 2009. Met behulp van een temporeel filter is het namelijk mogelijk te corrigeren voor toevalsfluctuaties in de tijd. Hiertoe is gebruik gemaakt van een algoritme dat het voortschrijdend gemiddelde over de tijd iteratief berekend, totdat de zogenaamde signaal-ruisverhouding aan bepaalde normen voldoet (Loke, 2011). Deze stap maakt gebruik van de volgens de eerste twee stappen gefilterde data voor de periode 1997-2009<sup>2</sup>. Merk op dat de data in bovenstaand raamwerk eerst ruimtelijk en dan temporeel wordt gefilterd; deze volgorde is gekozen omdat de onzekerheid in de ruimte (de populatiegrootte varieert aanzienlijk tussen kleine en grote gemeenten) in principe groter is dan die in de tijd (de populatiegrootte varieert minder sterk per gemeente).

**Kaart 1**

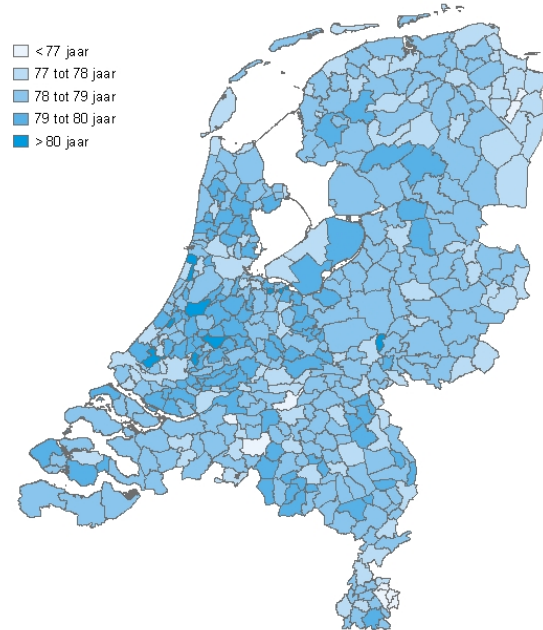
**Kaart 2**

Levensverwachting voor mannen per gemeente, 2009



Planbureau voor de Leefomgeving

Gefilterde levensverwachting voor mannen per gemeente, 2009



Planbureau voor de Leefomgeving

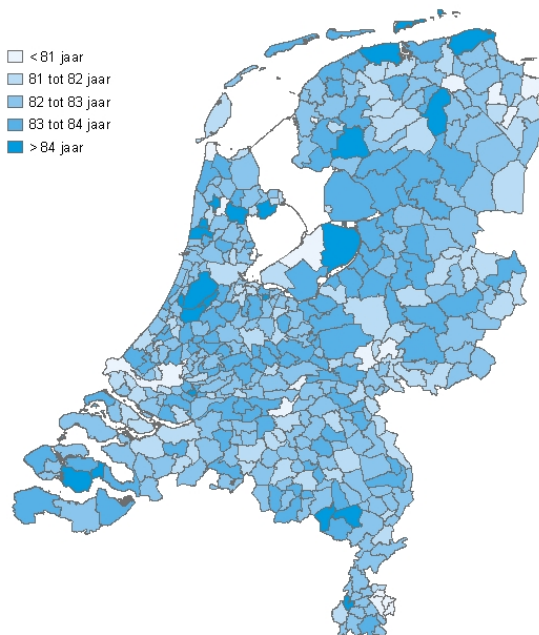
<sup>2</sup> Hierbij is gebruik gemaakt van de lijst van gemeenten die aanwezig waren in het jaar 2009; in dat jaar waren er 441 gemeenten. Gemeenten die in voorgaande jaren door gemeentelijke herindelingen zijn opgeheven, zijn in de jaren dat ze nog bestonden toegevoegd aan de nieuw gevormde gemeenten danwel aan al bestaande gemeenten waarmee ze zijn gefuseerd. Dit is gedaan door in die jaren de levensverwachting van de betrokken gemeenten gewogen te middelen aan de hand van de inwonertallen.

*Kaarten 1-4* geven een beeld van de regionale verschillen in de levensverwachting voor mannen en vrouwen voor het jaar 2009 voor en na de filtering. Er is sprake van grote regionale verschillen in de gemeentelijke levensverwachting. Voor mannen zijn de hoogste waarden van de gefilterde levensverwachting (rond 81 jaar) te vinden in de gemeenten Voorschoten, Kaag en Braassem, Langedijk, Dronten en Leiderdorp, en de laagste waarden (rond 76 jaar) in Kerkrade, Rucphen, Onderbanken en Tilburg. Voor vrouwen zijn de hoogste waarden (rond 85 jaar) te vinden in de gemeenten Langedijk, Dronten, Lisse, Dongeradeel, Schiermonnikoog, Meerssen en Borsele, en de laagste waarden (rond 80 jaar) in Kerkrade, Brunssum, Menterwolde, Pekela en Heerlen.

**Kaart 3**

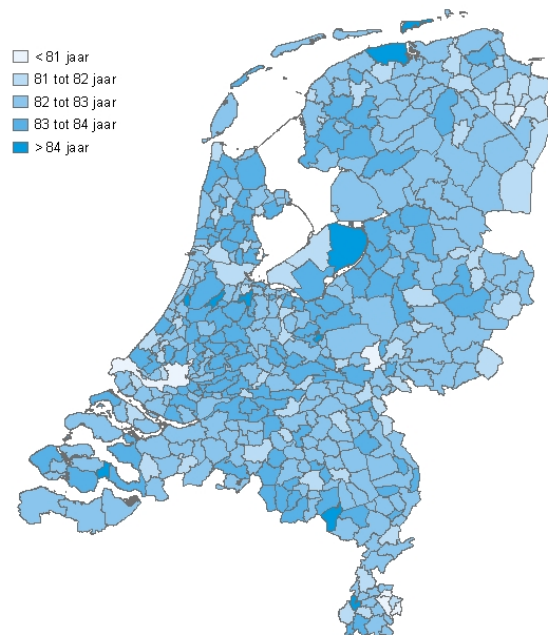
**Kaart 4**

Levensverwachting voor vrouwen per gemeente, 2009



Planbureau voor de Leefomgeving

Gefilterde levensverwachting voor vrouwen per gemeente, 2009



Planbureau voor de Leefomgeving

### 3.2 Analyse met behulp van multivariate regressie

Ter verklaring van regionale verschillen in de levensverwachting is een multivariaat regressiemodel opgesteld. In de Jong e.a. (2005) werd een model geschat met Ordinary Least Squares (OLS). In deze analyse is echter de voorkeur gegeven aan een weging van het model (apart voor mannen en vrouwen) aan de hand van het aantal personen van 65 jaar en ouder. In SPSS kan deze weging automatisch worden uitgevoerd met de optie Weighted Least Squares (WLS). Door deze weging gaan gemeenten met grote inwonertallen (toevalsfluctuatie in data is klein) zwaarder meetellen dan gemeenten met kleinere inwoneraantallen (toevalsfluctuatie in data is groot); zie ook Strutz (2010). De weging komt de precisie van de modellering dus ten goede.

Bij de toepassing van het regressiemodel is de volgende gangbare werkwijze gehanteerd (zie Chatterjee en Hadi, 2006). Naast de afhankelijke variabele (de levensverwachting) zijn alle onafhankelijke variabelen (univariaat) gecheckt op normaliteit, namelijk door inspectie van histogrammen. Enkel het histogram voor de zogenaamde bible belt (zie de volgende paragraaf) bleek scheef verdeeld. Per onafhankelijke variabele is gekeken (met behulp van scatterplots) naar de (bivariate) correlatie met de levensverwachting; op basis hiervan is een voorselectie van verklarende variabelen gemaakt. Vervolgens is de analyse met de geselecteerde variabelen uitgevoerd. Hierna zijn nog diverse variabelen uit het model verwijderd, namelijk de variabelen met een niet-significante bijdrage aan de verklaarde variantie of een (volgens de theorie) verkeerd teken. Ook variabelen met hoge VIFs (Variance Inflation Factors) werden verwijderd om problemen met multicollineariteit te voorkomen. Hierdoor resulteerde uiteindelijk een beknopt model met enkel significante variabelen die een plausibel teken hadden voor het effect op de levensverwachting.

Om het resulterende model te valideren, zijn ten slotte de residuen (de verschillen tussen waargenomen en voorspelde waarden) nog bekeken. Zowel bij het model voor mannen als voor vrouwen, is er sprake van één forse uitbijter (bij respectievelijk de gemeente Rozendaal en de gemeente Abcoude). Zonder deze uitbijters zijn de residuen normaal verdeeld (zowel de Kolmogorov-Smirnov test als de Shapiro-Wilk test leverde niet significante waarden op). Daarnaast is met behulp van het globale Moran's I algoritme (zie Griffith, 1987; Mitchell, 2005) getoetst of de residuen ruimtelijk zijn geclusterd (in welk geval er sprake zou zijn van ruimtelijke autocorrelatie). Dit bleek niet het geval te zijn (door in de analyse gebruik te maken van zogenaamde ruimtelijke dummies, zie hieronder, is ook hiervoor gecorrigeerd). Er kan dus worden geconcludeerd dat er wordt voldaan aan de algemene regressie-eisen (Chatterjee en Hadi, 2006) en dat de opgestelde globale regressiemodellen geldig zijn.

### **3.3 Verklarende variabelen voor verschillen in de regionale levensverwachting**

De verklarende variabelen voor het multivariate regressiemodel kunnen ingedeeld worden in vier categorieën.

#### *Demografische variabelen*

Uit het literatuuroverzicht van de vorige paragraaf kwam een effect van de burgerlijke staat naar voren: gehuwden leven langer dan ongehuwden. In het model zijn om deze reden diverse variabelen die betrekking hebben op de burgerlijke staat opgenomen. Hiertoe zijn op basis van cijfers in de databank StatLine van het CBS zowel voor mannen als vrouwen in de leeftijdsklasse 65 jaar en ouder de volgende indicatoren berekend voor het kalenderjaar 2009: het percentage personen dat getrouwd is, het percentage personen dat verweduwd is en het percentage personen dat woont in een institutioneel huishouden (zoals verpleeg- en verzorgingshuizen). De laatste variabele is geselecteerd in de veronderstelling dat dit via een selectie-effect (vooral mensen met een zwakke gezondheid zullen hier wonen) een negatieve invloed zal uitoefenen op de levensverwachting van een bepaalde gemeente.

#### *Culturele en leefstijlvariabelen*

In het literatuuroverzicht kwamen religie en leefstijl naar voren als belangrijke verklarende variabelen voor sterfteverschillen. Voor de operationalisatie van religie wordt gebruik gemaakt van een proxy, namelijk het percentage stemmers op de Christen Unie (CU) plus de Staatkundig Gereformeerde Partij (SGP) bij de Tweede Kamerverkiezingen van 2006. In ruimtelijk opzicht is dit een operationalisatie van de zogenaamde 'bible belt': een strook gemeenten die schuin door Nederland loopt vanaf de provincie Zeeland naar de provincie Overijssel en waar relatief veel gereformeerden en (orthodox) hervormden wonen.

Leefstijl kan geoperationaliseerd worden aan de hand van sterfte aan bepaalde doodsoorzaken. Voor de doodsoorzaken is gebruik gemaakt van de sterftcijfers in 2009 volgens de classificatielijst International Classification of Diseases ICD-10 (codes C00-D48), zie World Health Organization (2007), die door de databank Statline van het CBS worden verstrekt. De volgende doodsoorzaken zijn in de analyse meegenomen:

- ziekten van hart- en vaatstelsel (bijvoorbeeld hartziekten en hersen- en vaatletsel);
- nieuwvormingen (goedaardige of kwaadaardige gezwellen – ook wel met kanker aangeduid -);
- ziekten van het ademhalingsstelsel;
- overige doodsoorzaken (alle doodsoorzaken met uitzondering van nieuwvormingen, hart- en vaatziekten, ademhalingsziekten en uitwendige doodsoorzaken).

Uitwendige doodsoorzaken (ongevallen, zelfdoding, moord en doodslag, gebeurtenissen waarvan opzet onbekend is en overige uitwendige oorzaken van ziekte en sterfte) lijken in principe willekeurig en worden daarom niet meegenomen in de analyse. Per doodsoorzaak is een procentuele sterftkans bepaald door het aantal overledenen aan de betreffende doodsoorzaak (op alle leeftijden) te relateren aan alle personen in de leeftijdsklasse 65 jaar en ouder en vervolgens te vermenigvuldigen met 100; hierbij is verondersteld dat de meeste overledenen vallen in de leeftijdscategorie 65-plus.

### *Sociaaleconomische variabelen*

De sociaaleconomische status komt uit de literatuur naar voren als een belangrijke verklarende variabele voor sterfteverschillen. Dit lijkt in verband te staan met verschillen in leefstijl tussen verschillende bevolkingsgroepen. De sociaaleconomische status is niet rechtstreeks te ontleen aan de beschikbare data. Het is echter wel mogelijk om bepaalde economische variabelen te gebruiken als operationalisatie van sociaaleconomische status. Op basis van de beschikbare data in de databank StatLine van het CBS is gekozen voor het percentage personen met een WAO-, WW- of bijstandsuitkering in het kalenderjaar 2009 en het gemiddeld besteedbaar inkomen in het kalenderjaar 2007.

### *Geografische variabelen*

In de studie van Mackenbach e.a. (1991) kwam stedelijkheid naar voren als een belangrijke verklarende variabele voor sterfteverschillen. Deze variabele is geoperationaliseerd als de omgevingsadressendichtheid per vierkante kilometer in het jaar 2009, zoals gedefinieerd in de databank StatLine van het CBS.

In de Jong e.a. (2005) werden regionale variabelen in het multivariate regressiemodel toegevoegd door middel van dummies voor landsdelen, provincies en COROP regio's. Deze variabelen zijn een indicatie van structurele verschillen die niet kunnen worden verklaard door de overige variabelen. In de huidige analyse is gekozen voor een andere operationalisatie op gemeenteniveau, die minder grof is. Er worden wederom dummies in het model opgenomen, maar eerst is de ruimtelijke structuur van de dummies op gemeenteniveau onderzocht via een algoritme dat bekend is in de literatuur van de ruimtelijke analyse. Zowel voor de analyse van de levensverwachting van mannen als voor vrouwen is een zogenaamde 'lokale hot spot' analyse uitgevoerd, met behulp van het Anselin Moran's I algoritme (Anselin, 1995; Mitchell, 2005). Hierbij wordt een analyse gepleegd op de gemeentelijke residuen, berekend als het verschil tussen de waargenomen levensverwachting en de door het regressiemodel voorspelde levensverwachting, waarbij nog geen ruimtelijke dummies zijn meegenomen. In gemeenten waar significante waarden met het algoritme worden gevonden, is er sprake van een ruimtelijke clustering en moet het desbetreffende regressiemodel worden bijgesteld. Er is voor gekozen om de gevonden clusters te operationaliseren aan de hand van twee samengestelde dummies: de eerste dummy heeft betrekking op alle gemeenten met lage residuwaarden en de tweede dummy op alle gemeenten met hoge residuwaarden. Door het opnemen

van deze samengestelde dummies in het model wordt voldaan aan de algemene regressie-eisen (zie paragraaf 3.2).

#### 4. Uitkomsten van het multivariate model voor gemeentelijke sterfteverschillen

In deze paragraaf wordt het uiteindelijk geselecteerde multivariate model voor de verklaring van gemeentelijke sterfteverschillen beschreven. Door de werkwijze die is beschreven in paragraaf 3.2 zijn binnen de diverse categorieën slechts een beperkt aantal variabelen overgebleven. Het uiteindelijk afgeleide regressiemodel wordt in de *staten 1 en 2* gepresenteerd (één model voor mannen en één model voor vrouwen).

In de staten staan zowel de B waarde als de Bèta waarde weergegeven. Aan de hand van de B waarden kan de hoogte van de levensverwachting in een bepaalde gemeente worden berekend door vermenigvuldiging van de waarde van deze parameters met de hoogte van de betreffende verklarende variabelen en hierbij de waarde van de constante op te tellen. Bij de Bèta waarden is er gestandaardiseerd voor verschillen in de standaarddeviatie van de geselecteerde variabelen, zodat de verklaringskracht van de verschillende variabelen onderling kan worden vergeleken

Het regressiemodel met de overgebleven inhoudelijke variabelen is zowel voor de mannen als voor de vrouwen in staat 71% van de variantie in de gemeentelijke levensverwachting te verklaren. Alle variabelen zijn significant op 1% significantieniveau.

##### Staat 1

##### *Multivariaat regressiemodel voor de verklaring van de 'gefilterde' gemeentelijke levensverwachting bij geboorte van mannen in 2009<sup>1</sup>*

Variabele	Multivariaat ( $R^2 = 0,71$ )	
	B	Bèta
Constante	78,082 ***	
<i>Demografische variabelen</i>		
% gehuwde mannen	0,0310 ***	0,164
<i>Culturele en leefstijl variabelen</i>		
% ziekten van hart- en vaatstelsel	-0,816 ***	-0,233
% ziekten van ademhalingsstelsel	-1,138 ***	-0,171
<i>Sociaaleconomische variabelen</i>		
% uitkeringen (WAO, WW, bijstand) in totale bevolking	-0,104 ***	-0,358
Gemiddeld besteedbaar inkomen	0,084 ***	0,183
<i>Regionale dummies</i>		
Gemeente met onverwacht lage levensverwachting	-1,364 ***	-0,260
Gemeente met onverwacht hoge levensverwachting	1,044 ***	0,179

Significantieniveau: \* = 0,10; \*\* = 0,05; \*\*\* = 0,01

1. De variabelen zijn in elke gemeente gewogen met het aantal mannen van 65 jaar en ouder



## Staat 2

### *Multivariaat regressiemodel voor de verklaring van de 'gefilterde' gemeentelijke levensverwachting bij de geboorte van vrouwen in 2009<sup>2</sup>*

Variabele	Multivariaat ( $R^2 = 0,70$ )	
	B	Bèta
Constante	86,661 ***	
<i>Demografische variabelen</i>		
<i>Culturele en leefstijl variabelen</i>		
% stemmen op CU en SGP bij Tweede Kamerverkiezingen 200	0.013***	0,086
% ziekten van hart- en vaatstelsel	-0,546 ***	-0,159
% ziekten van ademhalingsstelsel	-0,959***	-0,146
% nieuwvormingen	-0,601***	-0,143
% overige doodsoorzaken	-0,435 ***	-0,156
<i>Sociaaleconomische variabelen</i>		
% uitkeringen (WAO, WW, bijstand) in totale bevolking	-0,098 ***	-0,349
<i>Regionale dummies</i>		
Gemeente met onverwacht lage levensverwachting	-0,807 ***	-0,213
Gemeente met onverwacht hoge levensverwachting	1,178 ***	0,224

Significantieniveau: \* = 0,10; \*\* = 0,05; \*\*\* = 0,01

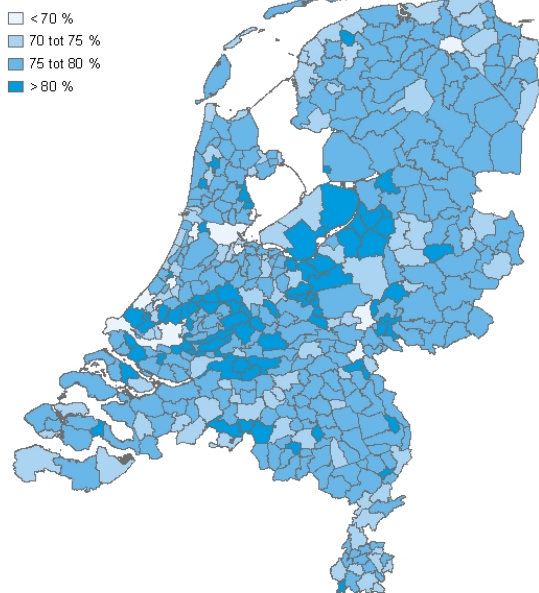
#### *2. De variabelen zijn in elke gemeente gewogen met het aantal vrouwen van 65 jaar en ouder*

Van de groep van demografische variabelen is bij de mannen de variabele percentage gehuwd overgebleven. Bij de vrouwen had het percentage gehuwd echter geen significant effect op de levensverwachting. Dit is in overeenkomst met het in de literatuur geconstateerde grotere effect van burgerlijke staat op mannen dan op vrouwen. Voor de mannen geldt dat een hoger percentage gehuwd samengaat met een hogere levensverwachting. Indien het percentage gehuwde mannen in een bepaalde gemeente 10 procentpunten hoger ligt dan in andere gemeenten, dan mogen de mannen in de betreffende gemeente rekenen op 0,3 jaar extra levensverwachting. In *kaart 5* is het percentage gehuwde mannen per gemeente weergegeven. Dit percentage is in het algemeen in de grotere gemeenten en de 'periferie' (de regio's die grenzen aan Duitsland en België) laag. In de plattelandsgemeenten in de Randstad en de zogenaamde intermediaire zone ligt het percentage gehuwde mannen doorgaans relatief hoog. Dit betekent dus dat de levensverwachting wat gedrukt wordt in de grote steden en de periferie.

Ook voor de culturele en leefstijlvariabelen zijn de opgenomen variabelen voor mannen en vrouwen verschillend. Voor vrouwen ligt in de 'bible-belt' gemeenten de levensverwachting iets hoger, terwijl voor mannen hier geen effect van uitgaat. Indien het percentage stemmers op CU en SGP in een bepaalde gemeente 10 procentpunten hoger ligt dan in andere gemeenten, dan mogen vrouwen in deze gemeente rekenen op 0,1 jaar extra levensverwachting. In *kaart 6* is het percentage stemmers op deze kleine christelijke partijen weergegeven. De 'bible belt' is opvallend goed zichtbaar in de kaart als de strook gemeenten die schuin over Nederland loopt vanuit de provincie Zeeland tot de provincie Overijssel. Ook in het westen van de provincie Groningen en het oosten van de provincie Friesland ligt het percentage stemmers op de twee christelijke partijen relatief hoog.

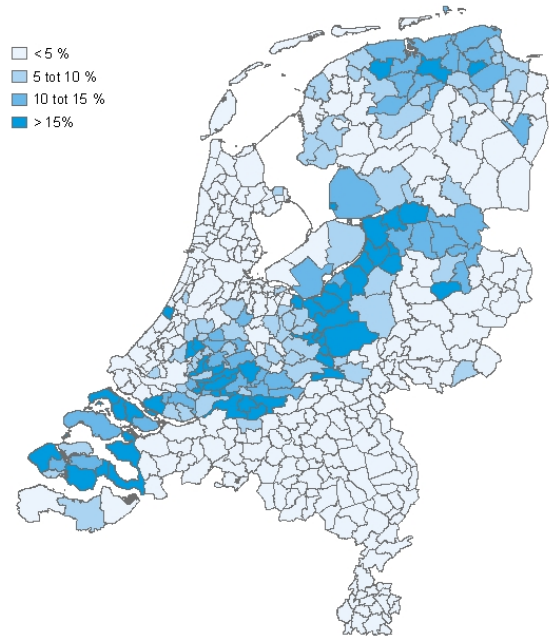
## Kaart 5

% Gehuwde mannen in de leeftijdsgroep 65+ jaar per gemeente, 2009



## Kaart 6

% Stemmers op kleine christelijke partijen per gemeente, Verkiezingen 2006



Wat betreft de verschillende variabelen die betrekking hebben op de doodsoorzaken, zijn de opgenomen variabelen wederom niet gelijk voor mannen en vrouwen. Bij mannen kwamen slechts twee doodsoorzaken als significant uit de bus, te weten sterfte aan ziekten van het hart- en vaatstelsel en sterfte aan ziekten van het ademhalingsstelsel. Bij vrouwen waren deze twee doodsoorzaken ook significant, maar er werden tevens twee extra doodsoorzaken in het model opgenomen, te weten sterfte aan nieuwvormingen en sterfte aan overige doodsoorzaken.

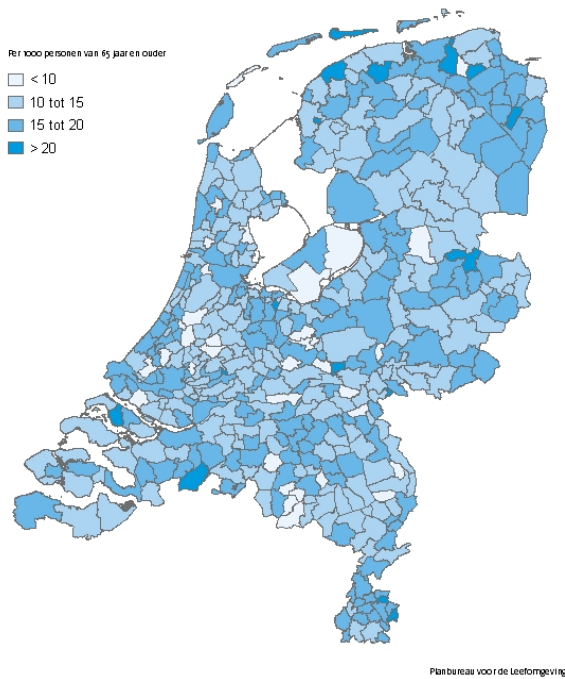
In de kaarten 7-10 worden voor de vier doodsoorzaken de sterftetekansen getoond, per 1000 personen van 65 jaar of ouder. De sterftetekans aan ziekten aan het hart- en vaatstelsel loopt uiteen tussen een 5 procent en 30 promille (merk op dat in de kaarten *alle* sterftegevallen aan een bepaalde doodsoorzaak gerelateerd zijn aan personen van 65 jaar en ouder). Er is geen duidelijk regionaal patroon zichtbaar; wel lijken er in de drie noordelijke provincies Groningen, Friesland en Drenthe en in het zuiden van de provincie Limburg relatief veel gemeenten te zijn met een hoge score op deze doodsoorzaak. Ook bij de sterfte aan nieuwvorming loopt de regionale sterftetekans uiteen van rond een 5 en 30 promille. De kaart met de sterftetekans aan nieuwvorming is echter duidelijk afwijkend van de kaart met de sterftetekans aan ziekten van het hart- en vaatstelsel. In de Randstadprovincies (Noord- en Zuid-Holland, Utrecht) maar ook in Gelderland, Flevoland en Noord-Brabant zijn er wat betreft nieuwvorming veel meer gemeenten die hoog scoren. De regionale variatie in de sterftetekans aan ziekten van het ademhalingsstelsel is minder sterk dan die in de sterftetekans aan de vorige twee doodsoorzaken: het loopt uiteen van 0 tot 10 promille. De regionale fluctuatie is wederom niet goed te duiden, maar met veel voorbehoud kan gezegd worden dat deze sterftetekans wat lager lijkt te liggen in de perifere regio's van Nederland (zoals de provincies Friesland, Groningen, Zeeland en Limburg) terwijl in provincies Gelderland, Overijssel

en Noord-Brabant relatief veel gemeenten hoog scoren. De sterftekans aan de overige doodsoorzaken loopt uiteen van rond 5 tot 20 promille. Wederom met de nodige reserve, kan gezegd worden dat relatief veel gemeenten in Noord- en Zuid-Holland, Utrecht, Friesland en Groningen hoge sterftekansen kennen.

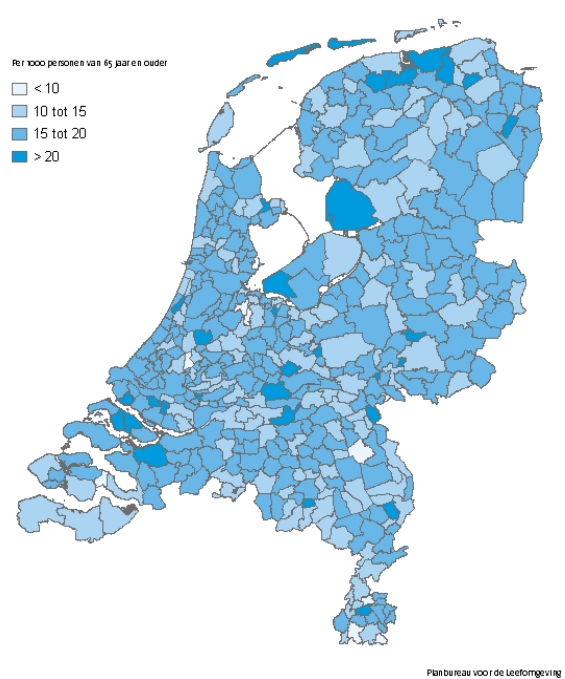
### Kaart 7

### Kaart 8

Sterftekansen aan hart- en vaatstelsel per gemeente, 2009



Sterftekansen aan nieuwvormingen per gemeente, 2009



De verklarende kracht van de sterftekans aan ziekten van het hart- en vaatstelsel, gemeten aan zowel de B waarde als de Bèta waarde, was voor mannen duidelijk hoger dan voor vrouwen. Indien in een bepaalde gemeente de sterfte aan deze doodsoorzaak 1 procentpunt hoger ligt dan in een andere gemeente dan betekent dit voor mannen een verlies van 0,8 levensjaar tegen 0,5 jaar voor vrouwen. Voor de sterftekans aan ziekten van het ademhalingsstelsel was de verklarende kracht voor mannen en vrouwen ongeveer gelijk. Hier gaat een verhoging van de sterftekans aan deze doodsoorzaak met 1 procentpunt zowel voor mannen als vrouwen samen met een verlaging van de levensverwachting van rond 1 jaar. Voor vrouwen geldt bovendien dat een verhoging van de sterfte in een bepaalde gemeente met 1 procentpunt voor zowel de doodsoorzaak nieuwvormingen als overige doodsoorzaken gepaard gaat met een verlies van rond een half levensjaar.

## Kaart 9

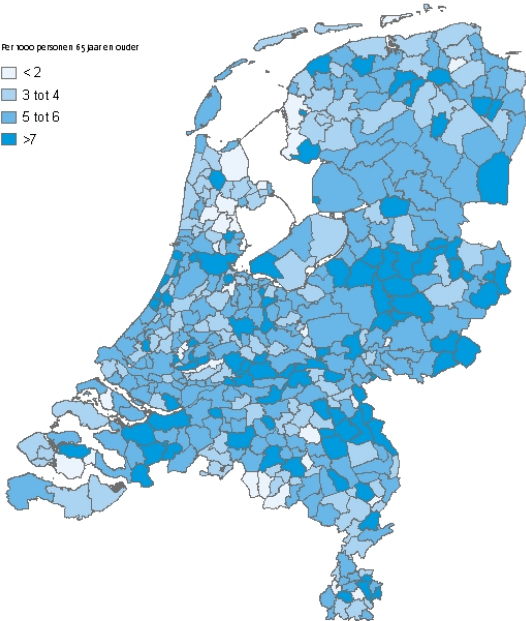
## Kaart 10

1

Sterftekansen aan ademhalingsstelsel per gemeente, 2009

Rr: 1000 personen 65 jaar en ouder

- < 2
- 3 tot 4
- 5 tot 6
- > 7

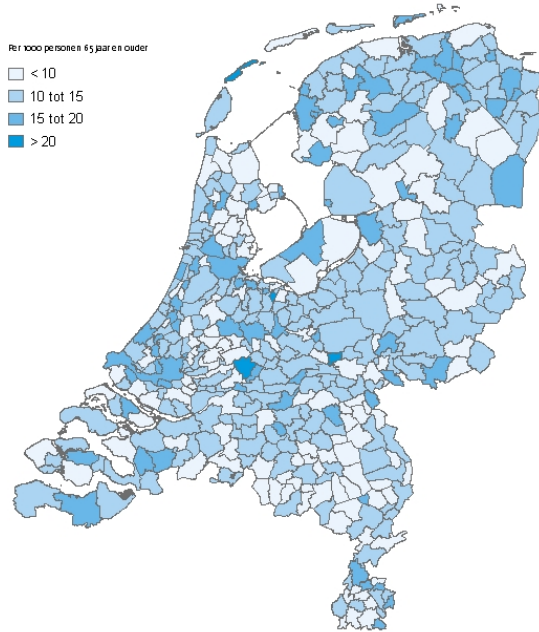


Planbureau voor de Leefomgeving

Sterftekansen aan overige doodsoorzaken per gemeente, 2009

Rr: 1000 personen 65 jaar en ouder

- < 10
- 10 tot 15
- 15 tot 20
- > 20

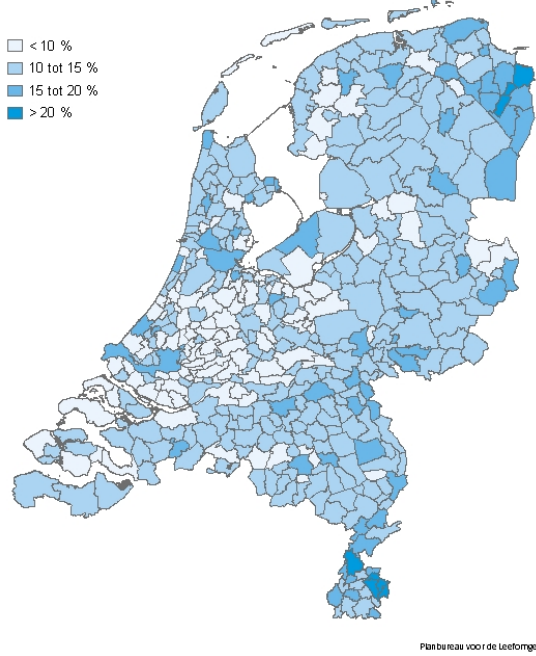


Planbureau voor de Leefomgeving

Van de sociaaleconomische variabelen werden voor mannen zowel de variabele gemiddeld besteedbaar inkomen als de variabele percentage personen met een uitkering geselecteerd; voor vrouwen bleek het inkomen echter geen significante invloed te hebben. Als gekeken wordt naar de Bèta waarde heeft de variabele percentage personen met een uitkering van alle opgenomen onafhankelijke variabelen de grootste verklarende waarde. Dit geldt zowel voor mannen als voor vrouwen. Uit *kaart 11* komt naar voren dat dit percentage regionaal uiteenloopt van 5% naar 25%. In de drie perifere provincies Groningen, Drenthe en Limburg lijken relatief veel gemeenten een hoog percentage personen met een WAO-, WW- of bijstandsuitkering te hebben. Ook in Amsterdam ligt het percentage personen met een uitkering hoog. Indien in een bepaalde gemeente het percentage personen met een uitkering 10 procentpunten hoger ligt dan in een andere gemeente, dan komt de levensverwachting zowel voor mannen als voor vrouwen rond 1 jaar lager uit. Voor mannen heeft het inkomen ook een significant effect. Blijkens *kaart 12* ligt de regionale bandbreedte van het gemiddeld inkomen (uitgedrukt in 1000 euro per jaar) tussen de klassen 0-15 en 25 en hoger. In gemeenten van de drie Randstad provincies lijkt dit inkomen beduidend hoger te liggen dan hierbuiten; met name in de drie noordelijke provincies ligt het veel lager. In gemeenten waar het gemiddeld inkomen 10 duizend euro hoger ligt, ligt de levensverwachting van mannen bijna een jaar hoger.

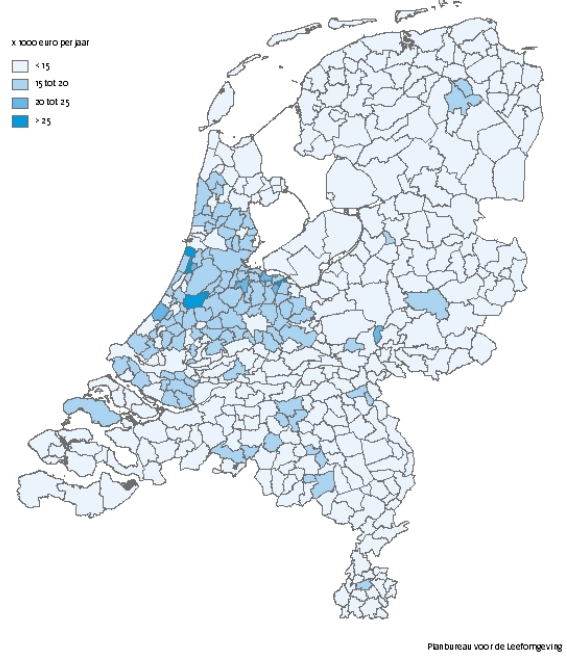
### Kaart 11

Percentage personen met een uitkering per gemeente, 2009



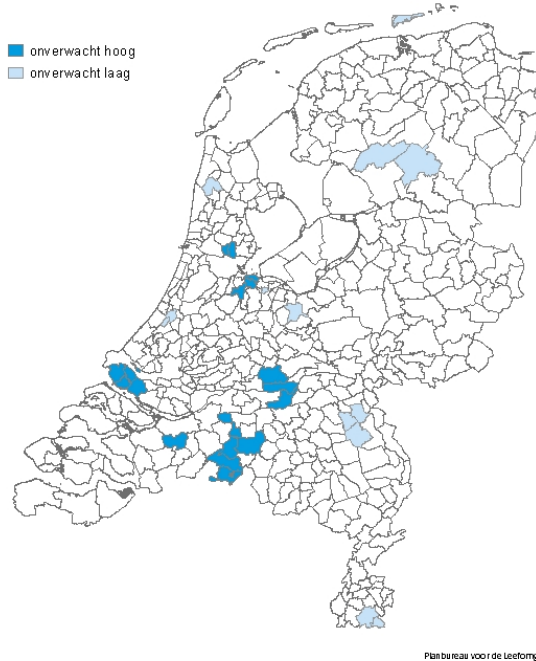
### Kaart 12

Gemiddeld besteedbaar inkomen per gemeente, 2009



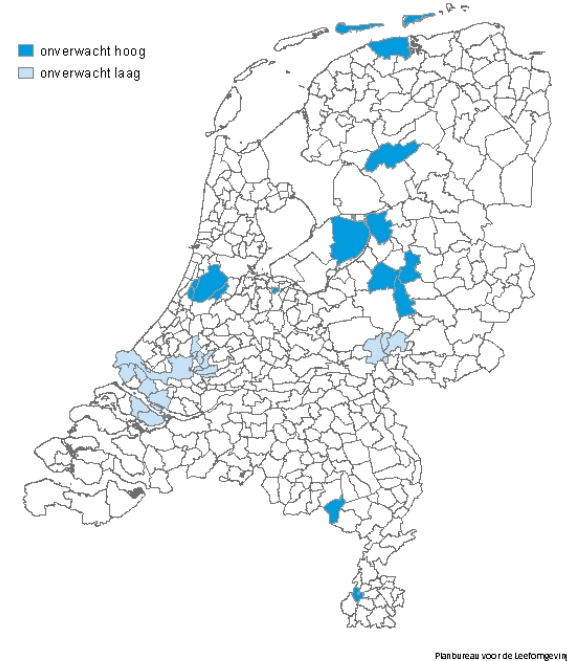
### Kaart 13

Gemeenten met een onverwacht hoge en lage levensverwachting, mannen



### Kaart 14

Gemeenten met een onverwacht hoge en lage levensverwachting, vrouwen





In het regressiemodel zijn ten slotte nog twee regionale dummies opgenomen: één voor gemeenten met een 'onverwacht' hoge levensverwachting en één voor gemeenten met een 'onverwacht' lage levensverwachting. Uit de *kaarten 13 en 14* komt naar voren dat het om een relatief gering aantal gemeenten gaat (zowel bij de mannen als de vrouwen zo'n twintigtal gemeenten met een lage levensverwachting tegen rond 15 gemeenten met een hoge levensverwachting). In deze gemeenten ligt de levensverwachting rond een jaar hoger dan wel rond een jaar lager dan verwacht (op basis van de andere verklarende variabelen).

#### 4. Conclusie

Voor de PBL/CBS regionale bevolkings- en huishoudensprognose is inzicht in verschillen in sterftekansen tussen regio's erg belangrijk. Hiertoe is in dit artikel, apart voor mannen en vrouwen, een verklarend model geschat dat gemeentelijke verschillen in de levensverwachting (als samenvattende indicator van sterftekansen) in verband brengt met diverse achtergrondvariabelen per gemeente. In een multivariaat regressiemodel is de gemeentelijke levensverwachting gerelateerd aan vier groepen van verklarende variabelen, te weten demografische, culturele/leefstijl, sociaaleconomische en geografische variabelen. Het verklarend model was zowel voor mannen als vrouwen in staat rond 70% van de regionale variatie in de levensverwachting te verklaren; in de Jong e.a. (2005) was dit rond 50%. Hierbij dient wel te worden opgemerkt dat bij de toepassing van het regressiemodel de variabelen gewogen zijn met het gemeentelijk aantal personen van 65 jaar en ouder (zie paragraaf 3.2).

Zowel bij mannen als vrouwen ging de belangrijkste verklarende kracht uit van de sociaaleconomische variabelen. Deze zijn geoperationaliseerd als het percentage personen met een WAO-, WW- of bijstandsuitkering per gemeente en het gemiddeld besteedbaar inkomen per gemeente. Deze twee variabelen geven een indicatie van de sociaaleconomische status, waarvan uit de literatuur is gebleken dat dit een belangrijk effect heeft op de levensverwachting. De uitkomsten van het regressiemodel bevestigen dit effect en geven aan dat het percentage personen met een uitkering een belangrijk deel van de regionale verschillen in de levensverwachting kan verklaren. Voor mannen had het gemiddeld besteedbaar inkomen ook een significant effect, maar voor vrouwen gold dit niet. Met name de drie perifere provincies Groningen, Drenthe en Limburg kennen relatief veel gemeenten met hoge percentages personen met een uitkering. Dit heeft een drukkend effect op de levensverwachting in deze gemeenten.

Ook de culturele en leefstijlvariabelen bleken, zoals verwacht volgens de literatuur, in deze studie belangrijk. Cultuur en leefstijl kan niet direct uit de beschikbare data worden gedistilleerd; het is enerzijds geoperationaliseerd via religie (door het percentage stemmen op de kleine christelijke partijen) en anderzijds via sterfte aan diverse doodsoorzaken. Deze doodsoorzaken verwijzen, zoals vermeld in paragraaf 2, naar aspecten van de leefstijl als roken, (verkeerd) eten en (alcohol) drinken. Zowel voor mannen als vrouwen blijken de doodsoorzaken ziekten van het hart- en vaatstelsel en ziekten van het ademhalingsstelsel een sterk negatief effect te hebben op de levensverwachting. Voor de vrouwen hadden tevens de doodsoorzaken nieuwvormingen en de overige doodsoorzaken nog een beduidend negatief effect op de levensverwachting. Net als in de Jong (2004), blijkt dus dat het belang van diverse doodsoorzaken voor de twee geslachten verschillend is. Hoewel het regionale patroon van sterfte aan diverse doodsoorzaken vrij moeilijk te duiden valt, kan met enige voorzichtigheid worden gesteld dat de sterftekansen aan ziekten van het hart- en vaatstelsel in diverse gemeenten in de drie noordelijke provincies en het zuiden van Limburg wat hoger liggen. Dit leidt dan tot een lagere levensverwachting in deze gemeenten. Voor de ziekten aan het ademhalingsstelsel is de regionale variatie wat minder sterk dan bij de ziekten van het hart- en vaatstelsel. In de 'intermediaire' provincies Gelderland, Overijssel en Noord-Brabant lijken relatief veel gemeenten voor te komen met een hogere incidentie van deze ziekten.

Wederom gaat hiervan een negatief effect uit op de levensverwachting. Religie heeft maar een gering effect op de levensverwachting. Bij vrouwen ligt de levensverwachting iets hoger in gemeenten die gerekend kunnen worden tot de zogenaamde 'bible belt'.

Van de demografische variabelen bleek enkel voor mannen het percentage gehuwd een significant effect te hebben op de levensverwachting: een hoger percentage gehuwd gaat samen met een hogere levensverwachting. Vanuit de literatuur werd overigens al voor mannen een sterker effect van de gehuwde staat op de levensverwachting verwacht dan voor vrouwen. In de meeste grote steden en perifere regio's ligt het percentage gehuwde mannen lager dan gemiddeld, terwijl in gemeenten in de intermediaire zone het percentage juist wat hoger ligt.

Bovenop deze structurele variabelen zijn ten slotte in het regressiemodel nog een tweetal regionale dummies opgenomen. In de in deze dummies gespecificeerde gemeenten blijkt de levensverwachting hoger of lager te zijn dan op basis van de inhoudelijke variabelen verklaard kan worden.

Bij de interpretatie van de uitkomsten moet zorgvuldig rekening worden gehouden met het probleem van 'ecological fallacy' (zie Robinson, 1950). Dit wil zeggen dat verbanden die op geaggregeerd niveau (in dit geval de gemeenten) worden gevonden, niet noodzakelijkerwijs voortvloeien uit verbanden op individueel niveau. Volgens het model ligt bijvoorbeeld de levensverwachting in diverse gemeenten in de periferie lager dan elders in het land; dit wil echter nog niet zeggen dat het wonen in de periferie leidt tot een lage levensverwachting.

In de regionale prognose met het model PEARL (zie de Jong e.a., 2005) wordt eerst een inschatting gemaakt van de hoogte van de (uiteindelijk geselecteerde) verklarende variabelen in 2020. Vervolgens kan met behulp van het ontwikkelde regressiemodel de levensverwachting voor mannen en vrouwen op gemeentelijk niveau in 2020 worden geschat. Deze uitkomsten worden vervolgens weer doorvertaald naar leeftijdsspecifieke sterftetekansen die dan in het regionale prognosemodel als input worden gebruikt.

## **Literatuur**

Anselin, Luc, 1995, Local Indicators of Spatial Association—LISA, *Geographical Analysis* 27(2): 93–115.

Anson, J., 2003, Sex Differences in Mortality at the Local Level: An Analysis of Belgian Municipalities, *European Journal of Population* 19: 1-28.

Bailey, T., Gatrell, A., 1995, *Interactive spatial data analysis*, Harlow, England: Longman.

Caselli, G., Cerbara, L., Heinsq, F., Lipsi, R.M., 2003, What impact Do Contextual Variables Have on the Changing Geography of Mortality in Italy, *European Journal of Population* 19 (4): 339-373.

Chatterjee, S., Hadi, A., 2006, *Regression analysis by example*, Wiley.

Commission on Social Determinants of Health, 2008, *CDSH Final Report: Closing the gap in a generation: Health equity through action on the social determinants of health*, Geneva: World Health Organization. Online at: [http://www.who.int/social\\_determinants/thecommission/finalreport/en/index.html](http://www.who.int/social_determinants/thecommission/finalreport/en/index.html)

Dinkel, R.H., 1985, The seeming paradox of increasing mortality in a highly industrialized nation: the example of the Sovjet Union, *Population Studies* 39: 87-97.

- Gaag, N. van der, Jong, A.H. de, 1997, Population scenarios for the European Union: regional scenarios, *Maandstatistiek van de bevolking*, Jaargang 45, december, blz. 17-31.
- Griffith, Daniel, 1987, *Spatial Autocorrelation: A Primer*, Resource Publications in Geography, Association of American Geographers.
- Jong, A. de, 2002, Gehuwden leven het langst, *Maandstatistiek van de bevolking*, Jaargang 50, juni, blz. 4-8.
- Jong, A. de, 2004, Analyse van doodsoorzaken op basis van overlevingstafeltechnieken, 1970-2003, *Bevolkingstrends*, 52(4): 23-36.
- Jong, A. de, Alders, M., Feijten, P., Visser, P., Deerenburg, I., Huis, M. van, Leering, D., 2005, *Achtergronden en veronderstellingen bij het model PEARL, Naar een nieuwe bevolkings- en allochtonenprognose*, Ruimtelijk Planbureau/Centraal Bureau voor de Statistiek, NAI Uitgevers, Den Haag/Rotterdam.
- Loke, R.E., 2011, Automated quality measures for spatial time series of demographic data, *Interne notitie*.
- Lucht, F. van der, Picavet, H., 2006, Sociaal-economische verschillen in leefstijl. Nationaal Kompas Volksgezondheid: [http://www.rivm.nl/vtv/object\\_document/o3193n21022.html](http://www.rivm.nl/vtv/object_document/o3193n21022.html).
- Mackenbach, J.P., Kunst, A.E., Looman, C.W.N., Beeck, E.F. van, 1991, Regionale sterfteverschillen in Nederland, In: *Sterfte en gezondheid nu en straks*, Nederlandse Vereniging voor Demografie, 's-Gravenhage.
- Mackenbach, J., 1992, Socio-economic health differences in the Netherlands: a review of recent empirical findings, *Social Science & Medicine* 34 (3): 213-226.
- Mackenbach, J., Stirbu, I., Roskam, A., Schaap, M., Menvielle, G., Leinsalu, M., 2008, Socioeconomic inequalities in health in 22 European countries, *The New England Journal of Medicine*, 358: 2468-2481.
- Marmot Review, 2010, *Fair Society, Healthy Lives*, Strategic Review of Health Priorities in England post-2010, Online at: <http://www.ucl.ac.uk/marmotreview>.
- McPherson, K., Marsh, T., Brown, M., 2009, *Tackling Obesities: Future Choices – Modelling Future Trends in Obesity and the Impact on Health*, Government Office for Science, 2nd Edition, Online at: <http://www.foresight.gov.uk/OurWork/ActiveProjects/Obesity/KeyInfo/Index.asp>
- Minkler, M., Fuller-Thomson, E., Guralnik, J., 2006, Socioeconomic spectrum in the United States, *The New England Journal of Medicine*, 355: 695-703.
- Mitchell, Andy, 2005, *The ESRI Guide to GIS Analysis, Volume 2*, ESRI Press.
- Pos, S., Bouwens, J., 2004, Preventie van hart- en vaatziekten bij jong volwassenen van 20 tot 40 jaar. Aanknopingspunten voor prioritering en beleidskeuzes, NIGZ.
- Prospective Studies Consortium, 2009, Body-mass index and cause-specific mortality in 900,000 adults: collaborative analyses of 57 prospective studies, *The Lancet*, <http://www.thelancet.com>, Published online March 18, 2009, DOI:10.1016/S0140-6736(09)60318-4.



- Robinson, W.S., 1950, Ecological Correlations and the Behavior of Individuals. *American Sociological Review* 15 (3): 351–357.
- Rogers, R.G., 1995, Marriage, sex and mortality, *Journal of Marriage and the Family* 57: 515-526.
- Smith, G.D., Carroll, D., Rankin, D.S., Rowan, D., 1992, Socioeconomic differentials in mortality: Evidence from Glasgow graveyards, *British Medical Journal* 305: 1554-1557.
- Smits J., Keij, I., Westert, G., 2001, Effecten van sociaal-economische status van kleine, middelgrote en grote geografische eenheden op de sterfte, *Maandstatistiek van de bevolking*, Jaargang 49, november, blz. 4-10.
- Strutz, T., 2010, *Data Fitting and Uncertainty - A practical introduction to weighted least squares and beyond*, Vieweg+Teubner, Wiesbaden.
- Valkonen, T., Poppel, F. van, 1997, The contribution of smoking to sex differences in life expectancy: four Nordic countries and The Netherlands 1970-1989, *European Journal of Public Health* 7: 302-310.
- Waller, L.A., Gotway, C.A., 2004, *Applied spatial statistics for public health data*, Wiley, New Jersey.
- World Health Organization, 2007, *International Statistical Classification of Diseases and Related Health Problems, 10th Revision*, Online at: <http://www.who.int/classifications/icd/en>.
- Wilkinson, R., Pickett, K., 2009, *The Spirit Level: why more equal societies almost always do better*, Allen Lane, London.
- Zick, C.D., Smith, K.R., 1991, Marital transitions, poverty and gender differences in mortality, *Journal of Marriage and the Family* 53: 327-336.