



# Informatie voor gebiedsgericht werken in het sociale domein: schattingen van gezondheid in Amsterdamse buurten

Daan G. Uitenbroek · Henriette Dijkshoorn · Suzan van Dijken

© Bohn Stafleu van Loghum, onderdeel van Springer Media B.V. 2017

**Samenvatting** *Achtergrond:* Er is een toenemende vraag naar gegevens over de gezondheid, de zorgbehoefte en het gezondheidsgedrag op buurtniveau. Enquêtegegevens zijn op dat niveau niet of nauwelijks beschikbaar en bovendien kostbaar om te verzamelen. Geschatte gegevens op grond van kenmerken van een buurt kunnen in de behoefte voorzien. In dit artikel wordt aandacht besteed aan verschillende aspecten van het doen van schattingen op een laag geografisch niveau. Daarbij wordt in het bijzonder gekeken naar de methode die de GGD Amsterdam gebruikt voor het uitvoeren van gezondheidsschattingen op buurtniveau.

*Methoden algemeen:* Drie factoren zijn belangrijk bij het maken van een schatting: 1) de bevolkingssamenstelling van buurten wat betreft bijvoorbeeld leeftijd en geslacht, 2) buurtfactoren zoals verstedelijking, SES-score van de buurt en afstand naar voorzieningen, en 3) de regio waarin een buurt ligt. Multilevel-regressiemodellen toegepast op enquêtegegevens maken het mogelijk om met deze factoren rekening te houden bij het samenstellen van een schattingsformule. Deze formule kan vervolgens worden toegepast op gegevens over de inwoners en andere karakteristieken van buurten.

*Methoden Amsterdam:* In Amsterdam zijn de multilevelformules ontwikkeld op basis van persoonsgegevens afkomstig uit de gezondheidsenquête en

op grond van regiogegevens afkomstig uit de CBS-kerncijfers Wijken en Buurten. Als regio's zijn de 22 Amsterdamse beleidsgebieden genomen. Deze regio's zijn als random intercepts meegenomen in de multilevel-regressieanalyse. De multilevelformule is vervolgens toegepast op gegevens van 97 Amsterdamse buurten.

*Beschouwing:* Het is moeilijk om wijk- en buurtschattingen te valideren omdat een gouden standaard, een exacte vaststelling van de prevalentie van een bepaalde gezondheidsuitkomst voor een buurt, veelal niet bestaat. Berekeningen die de GGD Amsterdam heeft gedaan voor verschillende uitkomsten laten zien dat de meeste schattingen vallen binnen het betrouwbaarheidsinterval van vergelijkbare enquêtegegevens. Dan is niet aan te geven wat dichter bij de 'werkelijkheid' ligt, de multilevelschatting of de observatie met een enquête. Gesprekken met gebruikers uit de praktijk laten zien dat schattingen worden aanvaard, waarbij opvalt dat men goed beseft dat het gaat om een verwachting op grond van het profiel van de buurt, en dat deze verwachting kan afwijken van de realiteit.

**Trefwoorden** gezondheidsstatistiek · buurt · synthetische schattingen · zorgbehoefte

**Information for small area health and social work: estimation of health in Amsterdam neighbourhoods**

**Abstract** *Background:* There is increasing demand for health statistics at a low geographical level, such as the neighbourhood. There are little survey data available and collecting data is expensive. Estimates, formulating an expectation on the basis of knowledge of a neighbourhood, could be a solution. In this paper we discuss the methods of doing health estimations

---

D. G. Uitenbroek (✉) · H. Dijkshoorn  
afdeling Epidemiologie, Gezondheidsbevordering  
en Zorginnovatie, GGD Amsterdam, Nieuwe  
Achtergracht 100, Postbus 2200, 1000 CE Amsterdam,  
Nederland  
daanuitenbroek@ggd.amsterdam.nl

S. van Dijken  
afdeling Beleid & Onderzoek, GGD Flevoland, Flevoland,  
Nederland



for small geographical areas. Particular attention is given to the method used by the Amsterdam public health service.

*Methods in general:* Three factors are important in doing health estimations in neighbourhoods: population characteristics, neighbourhood characteristics, and the area in which the neighbourhood is located. Multilevel regression analysis allows for the development of a single estimation model which incorporates these three factors. This model can then be applied to neighbourhood level data to obtain the health estimate.

*Methods used in Amsterdam:* For the estimation of loneliness a formula was developed which considered age and gender of persons, and the value of real estate and the proportion of people in the area which received benefits. We made estimations for 97 Amsterdam neighbourhoods and aggregated these to the level of 22 areas. Next, we compared the aggregated estimates with observations from a health survey. The estimates mostly were within the confidence intervals of the observations.

*Discussion:* It is difficult to determine the validity of estimates because golden standards with which to compare estimates are rare in public health. The health indicators analysed by the Public Health Service Amsterdam show that the estimates are mostly within the confidence intervals of observations. It is impossible to judge which is nearer to 'reality', multilevel or survey estimate. Discussion with users shows a high acceptance of estimated health indicators. There is an understanding that it concerns an expectation on the basis of the neighbourhoods profile, and that there can be a divergence between expectation and 'reality'.

**Keywords** synthetic estimation · small area statistics · neighbourhood

### Kernpunten

- Geschatte gegevens op grond van kenmerken van een buurt kunnen in de behoefte aan buurt- en wijkstatistiek voorzien.
- Het is moeilijk om buurtschattingen te valideren, maar vergelijkingen met enquêtegegevens laten zien dat de meeste schattingen vallen binnen het betrouwbaarheidsinterval van geobserveerde gegevens.
- Gesprekken met gebruikers laten zien dat schattingen worden aanvaard en dat men goed beseft dat het gaat om een verwachting op grond van het profiel van de buurt.

### Inleiding

Met de toename van het wijk- en buurtgericht werken in het sociale domein en vanwege de steeds grotere omvang van gemeenten bestaat er een groeiende behoefte aan informatie op kleinschalig geografisch niveau, bijvoorbeeld in de vorm van wijk- of buurtanalyses [1]. Enquête-onderzoeken in geografisch kleine gebieden zoals buurten is kostbaar. Een alternatief zijn schattingen op buurtniveau, dat wil zeggen het genereren van verwachtingen over de toestand in een buurt op grond van het profiel van de buurt. In Nederland zijn verscheidene onderzoekers bezig met het ontwikkelen van statistische modellen voor buurtschattingen, zoals de Vraag Aanbod Analyse Monitor eerste lijn (VAAM) van het Nivel [2-4]. In dit artikel bespreken wij een aantal praktische aspecten van de toepassing van schattingen en gaan we in op de manier waarop de GGD Amsterdam modellen voor buurtschattingen ontwikkelt en toepast.

### Uitgangsidee

In principe is alle informatie die men nodig heeft voor het ontwikkelen van het beleid te schatten. In de VAAM wordt informatie geschat om zorgverleners inzicht te geven in de omvang van de vraag naar verschillende typen zorg in de wijk [2-4]. Andere onderzoekers doen schattingen naar aan gezondheid gerelateerde gedragingen, [5] bijvoorbeeld hoeveel mensen in een wijk roken of problematisch alcoholgebruiker zijn, [6-8] het aantal zwangere vrouwen in een buurt dat rookt, [9] de mate van maatschappelijke participatie [10] of de opkomst voor borstkankerscreening [11]. Een bijzondere toepassing is het schatten van noodzakelijke zorg voor een buurt [12]. Zo kan men het aantal uren noodzakelijke huisartsenzorg schatten op grond van een profiel van een buurt, en dit vervolgens vergelijken met de beschikbare zorg, om gebieden waar mogelijk een tekort aan zorg is in kaart te brengen [3].

Schattingsprocedures werken door middel van het toepassen van een schattingsformule op gegevens over een buurt. Er zijn drie soorten gegevens die belangrijk zijn bij schattingen. Ten eerste persoonsfactoren, bijvoorbeeld de leeftijd en het geslacht van personen woonachtig in een buurt, en andere individuele kenmerken die voor de aan- of afwezigheid van een bepaalde uitkomst relevant zijn. Dan zijn er omgevingsfactoren, bijvoorbeeld of de buurt waar men uitspraken over wil doen veel groen heeft, of er veel mensen werkloos zijn, of er goede voorzieningen zijn op het gebied van de gezondheidszorg of anderszins. Dezelfde factor kan zowel op persoons- als op buurtniveau belangrijk zijn. De persoon heeft bijvoorbeeld zelf een werkloosheidsstatus, maar hij of zij zal ook wonen in een buurt met veel of weinig werklozen, onafhankelijk van de persoonlijke status. Ten slotte is er nog het effect van de regio. Personen en omstandigheden in buurten in Friesland en Lim-

burg kunnen bijvoorbeeld vergelijkbaar zijn, maar het gaat wel om twee heel aparte delen van Nederland met elk een eigen cultuur.

De volgende formule geeft dit model weer:

$$\text{Uitkomst } y_{ij} = \text{intercept} + \sum \alpha_{ij} x_{ij} + \sum \beta_j z_j + \mu_{jklm} + \varepsilon_{ij}$$

Waarbij

|                           |  |
|---------------------------|--|
| $y_{ij}$                  | eigenschap van persoon $i$ die woont in regio $j$  |
| $\sum \alpha_{ij} x_{ij}$ | verklaring door persoonsfactoren   |
| $\sum \beta_j z_j$        | verklaring door omgevingsfactoren  |
| $\mu_{jklm}$              | de regio-effecten; meestal neemt men maar één hoger niveau, de meeromvattende regio of wijk waarin een buurt ligt, 'j', maar regio kan verschillende (geneste/gestapelde) niveaus 'jklm' betreffen, het aantal bepaalt de gebruiker; is de random intercept in een multilevelanalyse |
| $\varepsilon_{ij}$        | restvariantie op persoonsniveau, wat de persoon uniek maakt na rekening te hebben gehouden met de andere factoren uit het model  |

De uitkomst is dan de schatting voor de buurt van het aantal mensen met een eigenschap of aandoening. Daarvan wordt vervolgens een gemiddelde, percentage, totaal of een andere functie genomen voor de uiteindelijke buurtschatting.

Wanneer het gaat om het onderzoeken en voorspellen van het gedrag van personen wordt bij onderzoek meestal alleen rekening gehouden met de persoonsfactoren. De VAAM doet buurtvoorspellingen op basis van alleen omgevingsfactoren. Multilevelanalyse maakt het mogelijk om in een enkele vergelijking met alle factoren uit de formule rekening te houden. Daarbij is het vooral ook belangrijk dat de waarden van de  $\mu$ -parameter kunnen worden bepaald, met verschillende waarden voor buurten afhankelijk van de regio waarin zij liggen. Dit is een additionele optie in gangbare multilevelpakketten [13, 14].

Er bestaat een essentieel onderscheid tussen schatting en observatie, zoals een vragenlijstonderzoek. Bij schattingen gaat het over een verwachting op grond van de factoren die men heeft meegenomen bij het samenstellen van de schattingsformule. Factoren die men niet heeft meegenomen kunnen tot afwijkingen leiden. Een voorbeeld: men verwacht op grond van het sociaal-economisch profiel van de buurt een verhoogd aantal rokers. Wanneer er een actieve huisarts in de buurt praktijk houdt, kan dat er echter toe leiden dat het feitelijke aantal rokers lager ligt. Zoiets is niet te modelleren. Anderzijds, bij observaties is het probleem dat men niet altijd de juiste mensen onderzoekt. Er kunnen in de steekproef bijvoorbeeld te weinig patiënten van de actieve huisarts meedoen doordat bepaalde groepen bewoners slecht responderen. Dan zou men ook op grond van een observatie tot een foute uitspraak kunnen komen. Dit probleem speelt sterker naarmate het aantal observaties kleiner is.

Het is lastig om aan te geven wanneer een schatting de voorkeur verdient boven een observatie. Een schattingsformule die op basis van een grote gegevensset met voldoende variabelen ontwikkeld is en toegepast is in een buurt die niet al te veel afwijkt wat betreft de aan- en afwezigheid van de factoren die mee zijn genomen in de schattingsformule, kan zeer valide schattingen geven. Schattingen zullen dan de voorkeur verdienen boven een kleine steekproef of een onderzoek met een lage respons. Er bestaat echter geen vuistregel aan de hand waarvan we kunnen bepalen wanneer de ene methode de voorkeur verdient boven de andere.

### Gebruikte gegevens en opstellen van de formule

De GGD Amsterdam doet de buurtschattingen op basis van multilevel-regressiemodellen zoals boven omschreven, die zijn gevalideerd door Twigg et al. [6]. De schattingsformule wordt ontwikkeld voor 22 gebieden – regio's in het bovenstaande model – in Amsterdam waarvoor voldoende betrouwbare enquêtegegevens aanwezig zijn. Gegevens over de gezondheidsuitkomsten waarvoor een schattingsformule moet worden gemaakt zijn ook uit de enquête afkomstig en worden als onafhankelijke variabele in het model opgenomen. De resulterende formule wordt toegepast op gegevens van 97 Amsterdamse buurten zoals gedefinieerd in de CBS-kerncijfers Wijken en Buurten [15] om gezondheidsuitkomsten voor deze buurten te schatten. Er zijn enquêtegegevens van 7.218 respondenten van de Amsterdamse Gezondheidsmonitor (AGM 2012) gebruikt (respons: 38%). De AGM 2012 was onderdeel van de Gezondheidsmonitor Nederland [16]. De gegevens van de 22 gebieden zijn verrijkt met 29 omgevingsfactoren uit de CBS-kerncijfers Wijken en Buurten. Omdat de CBS-omgevingsfactoren sterk gecorreleerd zijn is hiervoor een selectie gemaakt uit de meer dan honderd omgevingsfactoren die beschikbaar zijn. Gebruikt zijn onder andere de bevolkingssamenstelling van het gebied naar leeftijd, etniciteit, sociaal-economische kenmerken, huishoudsamenstelling, mate van verstedelijking en afstand naar diverse voorzieningen. De multilevel logistische regressieanalyse is uitgevoerd in R, pakket lme4. De 22 gebieden zijn in het model als random intercept meegenomen.

Een stapsgewijze procedure is gebruikt om een combinatie van variabelen te vinden die de uitkomstvariabele het beste verklaart. Als persoonsfactoren zijn standaard leeftijd en geslacht meegenomen, andere factoren zijn niet bekeken. In het uiteindelijke multilevel logistische regressiemodel zijn op regioniveau alleen variabelen meegenomen die statistisch significant samenhangen met de uitkomstvariabele. Het multilevel logistische regressiemodel is gebruikt voor het opstellen van een formule om de prevalentie van de uitkomstmaat te schatten voor de 97 buurten.

**Tabel 1** Observatie en schatting van het percentage eenzamen in 22 Amsterdamse gebieden

| Gebied                                     | n<br>AGM | Geobserveerd<br>(%) | 95 %-<br>(%) | 95 %+<br>(%) | Geschat<br>(%) | Sign. |
|--|----------|---------------------|--------------|--------------|----------------|-------|
| 1 Centrum, West                            | 354      | 6,20                | 3,60         | 10,30        | 6,20           |       |
| 2 Centrum, Oost                            | 346      | 7,20                | 4,30         | 12,00        | 6,50           |       |
| 3 Westerpark                               | 193      | 13,20               | 8,20         | 20,60        | 9,30           |       |
| 4 Bos en Lommer                            | 217      | 7,00                | 4,00         | 11,80        | 10,90          |       |
| 5 Oud-West en de Baarsjes                  | 453      | 10,90               | 7,70         | 15,10        | 8,40           |       |
| 6 Geuzenveld en Sloterveer                 | 220      | 22,70               | 13,60        | 35,50        | 11,80          | *     |
| 7 Osdorp                                   | 227      | 12,90               | 8,40         | 19,40        | 10,70          |       |
| 8 De Aker en Sloten                        | 178      | 7,30                | 3,60         | 14,20        | 7,10           |       |
| 9 Slotervaart                              | 202      | 8,50                | 5,30         | 13,30        | 10,40          |       |
| 10 Zuid                                    | 244      | 8,30                | 4,80         | 13,90        | 5,20           |       |
| 11 Buitenveldert en Zuidas                 | 143      | 5,60                | 3,00         | 10,10        | 7,40           |       |
| 12 De Pijp en Rivierenbuurt                | 311      | 8,60                | 5,60         | 13,00        | 7,70           |       |
| 13 Oud-Oost                                | 242      | 11,10               | 6,60         | 18,00        | 9,60           |       |
| 14 Oostelijk havengebied en Indische buurt | 396      | 8,40                | 6,00         | 11,60        | 8,90           |       |
| 15 Watergraafsmeer                         | 202      | 8,60                | 4,60         | 15,20        | 7,70           |       |
| 16 Zeeburgereiland en IJburg               | 80       | 3,20                | 0,90         | 10,50        | 6,10           |       |
| 17 Noord-West                              | 995      | 12,10               | 9,80         | 14,80        | 10,50          |       |
| 18 Oud-Noord                               | 792      | 14,90               | 12,20        | 18,00        | 11,80          | *     |
| 19 Noord-Oost                              | 672      | 12,00               | 9,20         | 15,50        | 11,00          |       |
| 20 Bullewijk en Bijlmer-Centrum            | 113      | 19,20               | 11,80        | 29,70        | 14,00          |       |
| 21 Bijlmer-Oost                            | 358      | 11,50               | 8,10         | 16,10        | 13,70          |       |
| 22 Gaasperdam en Driemond                  | 280      | 15,80               | 10,80        | 22,60        | 12,10          |       |

Voor prevalentie van eenzaamheid in de buurt, een van de speerpunten van het Amsterdamse gezondheidsbeleid, was de formule bijvoorbeeld:

$$\begin{aligned} \%_{\text{eenzamen}} = & \text{intercept} + \alpha X_{\text{geslacht}} + \alpha X_{\text{leeftijd}} \\ & + \beta Z_{\text{woningwaarde}} + \beta Z_{\%_{\text{wet\_werk\_en\_bijstand}}} \\ & + \mu_{\text{regio},j} = 1..22 \end{aligned}$$

### Doen van de schattingen en validering

Voor ieder van de 97 buurten is een spreadsheet in Excel gemaakt, met in de rijen de opbouw naar leeftijd en geslacht in de buurten en in de kolommen de omgevingsfactoren per buurt. De omgevingsfactoren zijn constant voor de gehele buurt. Aan het Excel-bestand is voor iedere buurt het random intercept voor de regio toegevoegd als een laatste kolom ( $\mu_j$  in de formule), om te corrigeren voor het effect van de regio waarin de buurt ligt. Op ieder van de 97 spreadsheets is het multilevel logistische regressiemodel uit de vorige fase toegepast voor het schatten van de prevalentie van de uitkomstmaat in de buurt.

Om de geschatte resultaten te valideren zijn de uitkomsten voor de 97 buurten geaggregeerd naar de 22 gebieden. Dit is gedaan voor verschillende uitkomstvariabelen. Voor de 22 gebieden zijn redelijk betrouwbare observaties beschikbaar uit de AGM 2012.

In tab. 1 staat het resultaat voor eenzaamheid.

De betrouwbaarheidsintervallen in tab. 1 zijn berekend volgens de Wilson-methode [17], de standaardmethode voor percentages in SPSS. Deze intervallen zijn breder dan normaal omdat de geobserveerde gegevens vanwege steekproefstratificatie en non-respons zijn gewogen naar de bevolkingssamenstelling. Hiervoor is gecorrigeerd met SPSS-complex samples.

Tab. 1 laat zien dat de geschatte prevalentie van eenzaamheid in twee van de 22 gebieden buiten het betrouwbaarheidsinterval van de geobserveerde prevalentie valt (verwachting 5%). In 12 van de 22 gebieden ligt de schatting hoger dan de observatie (verwachting 50%), de correlatie tussen schatting en observatie is 0,90. Dit alles geeft aan dat het verschil tussen schatting en observatie binnen de kansverwachting van de observatie valt.

### Beschouwing

In dit artikel is een methode besproken voor het doen van schattingen van gezondheidsindicatoren op buurtniveau, op basis van kenmerken van de buurt. GGD Amsterdam gebruikt hiervoor multilevel logistische regressiemodellen, zoals beschreven door Twigg et al. [6, 7]. Deze methode lijkt schattingen op te leveren die goed overeenkomen met onze enquêtegegevens. In het model zijn drie factoren meegenomen. Ten eerste individuele factoren, afkomstig uit de Amsterdamse Gezondheidsmonitor 2012. Ten tweede omgevingsfactoren, afkomstig uit de CBS-kerncijfers

Wijken en Buurten. Ten derde is het effect van de regio meegenomen, door het schatten van de afwijking van de intercept voor verschillende regio's. Als voorbeeld is een model voor het voorspellen van het vóórkomen van eenzaamheid ontwikkeld op basis van gegevens over 7.218 personen woonachtig in 22 Amsterdamse regio's. Dit model is in een volgende fase gebruikt om schattingen te maken voor de prevalentie van eenzaamheid in 97 Amsterdamse buurten.

### Beperkingen van de methode

Beschikbaarheid van gegevens is de belangrijkste beperking. Bij de individuele kenmerken moeten de gegevens die wordt gebruikt voor het opstellen van het predictiemodel in detail aanwezig zijn in de vorm van enquêtegegevens én beschikbaar zijn voor het betreffende geografische niveau waarvoor de uiteindelijke schattingen worden gedaan. Voor deze individuele kenmerken is men daarom meestal beperkt tot gegevens uit het bevolkingsregister, zoals leeftijd en geslacht. Voor de omgevingsfactoren kan men gebruikmaken van de meer dan honderd indicatoren van de CBS-kerncijfers Wijken en Buurten. Deze factoren zijn onderling echter sterk gecorreleerd en er moet een selectie worden gemaakt om problemen met multicollineariteit te vermijden.

### Vergelijking met andere schattingsmethoden

De multilevel-regressiemethode die hier is gebruikt, is uitgebreider dan de methode die in de VAAM wordt gebruikt, een methode die overeenkomt met de Engelse ONS-methode [5]. Bij deze methode wordt alleen rekening gehouden met verschillen tussen gebieden wat betreft de omgevingsfactoren, terwijl de schattingsformule op het buurt- en niet op het persoonsniveau wordt ontwikkeld. Een vergelijking tussen de multilevel-regressiemethode en de ONS-methode laat zien dat de multilevel-regressiemethode niet aantoonbaar beter presteert, terwijl deze methode ingewikkelder is om uit te voeren [5].

In Amsterdam geven we toch de voorkeur aan de multilevelmethode, omdat deze methode het meest plausibel lijkt. Bovendien blijkt dat de gebruiker deze methode acceptabeler vindt. De regioparameter draagt hiertoe bij, een bestuurder wil nu eenmaal graag dat er rekening wordt gehouden met de eigenheid van 'zijn/haar' regio; dat ondanks alle overeenkomsten Noord toch echt verschilt van West.

### Validatie en acceptatie

Het valideren van schattingen is moeilijk. Van de meeste indicatoren in de publieke gezondheid bestaat geen gouden standaard, we weten het aantal rokers of eenzamen in een buurt niet exact. Het is dan ook niet mogelijk om vast te stellen in hoeverre een schatting van een buurtprevalentie op basis van een multilevel-

regressiemodel met de 'ware' buurtprevalentie overeenkomt. Daar komt bij dat de schattingen die er zijn meestal binnen het betrouwbaarheidsinterval van de observaties liggen. Dit is zo voor eenzaamheid, maar ook voor vijf andere gezondheidsindicatoren die door ons zijn onderzocht. Als schattingen binnen de betrouwbaarheidsmarge van de observaties liggen kan niet worden getoetst of een uitgebreider model met misschien wel veel slimmere methoden ook echt tot een betere beschrijving van de werkelijke waarden leidt. Bovendien is niet vast te stellen of het de observatie of de schatting is die dichterbij de 'werkelijkheid' ligt.

Voor de 97 Amsterdamse buurten zijn voor zes gezondheidsindicatoren buurtschattingen gemaakt. We zijn nu aan het onderzoeken in hoeverre de methode voor de gebruikers relevant en acceptabel is. Hiervoor hebben wij een aantal presentaties gedaan voor sociale wijkteams in Amsterdam. Het concept dat het hier gaat om geschatte informatie, om een verwachting op grond van wat bekend is over de wijk en de inwoners van de wijk en dat dit niet noodzakelijkerwijs met de realiteit overeenkomt, lijkt aanvaardbaar. Dit is vooral zo omdat de uitkomst van de schattingen over het algemeen overeenkomt met de praktische ervaring met buurtverschillen van de medewerkers van de sociale wijkteams en omdat men het gevoel heeft dat met relevante factoren rekening is gehouden. Men is daarbij vaak op zoek naar het gevoel dat de buurt 'het beter doet' dan op grond van de factoren uit een schatting zou worden verwacht, als indicator van het eigen succes. Het wordt dus niet noodzakelijk gevonden dat de schatting overeenkomt met de 'ware' werkelijkheid, maar dat de schatting overeenkomt met een, vaak complexe, *verwachting* van de werkelijkheid. Op deze manier kunnen schattingen in het gebruik een heel andere betekenis hebben dan observaties.

### Literatuur

1. Molleman G, Pieters M. Reflectie vanuit het GGD-perspectief. Tijdschr Gezondheidswet. 2014;92:8.
2. Zwaanswijk M, Quak ES, Bouwman J, et al. Validiteit van de Vraag Aanbod Analyse Monitor (VAAM): overeenstemming met populatiegegevens uit lokale gezondheidspeilingen. Tijdschr Gezondheidswet. 2009;87:272-8.
3. Ruizendaal-de Graaf WA, Kenens R, Bakker DH de. The construction of a decision tool to analyse local demand and local supply for GP care using a synthetic estimation model. Hum Resour Health. 2013;55:1-9.
4. Ruizendaal-de Graaf WA, Brink CL van den, Hoek L van den, et al. Geschatte lokale cijfers over gezondheid en gezondheidsdeterminanten ter ondersteuning van het gezondheidsbeleid. Tijdschr Gezondheidswet. 2015;93:274-80.
5. Pickering K, Scholes S, Bajekal M. Synthetic estimation of healthy lifestyles indicators: stage 2 report. London: Nat-Cen Social Research; 2004. [www.publichealth.ie/files/file/Synthetic\\_Estimation\\_Stage\\_2\\_Report.pdf](http://www.publichealth.ie/files/file/Synthetic_Estimation_Stage_2_Report.pdf). Geraadpleegd op: 8 augustus 2017.

6. Twigg L, Moon G. Predicting small area health-related behaviour: a comparison of multilevel synthetic estimation and local survey data. *Soc Sci Med*. 2002;54:931–7.
7. Twigg L, Moon G, Jones K. Predicting small-area health-related behaviour: a comparison of smoking and drinking indicators. *Soc Sci Med*. 2000;50:1109–20.
8. Cui Y, Baldwin SB, Lightstone AS, et al. Small area estimates reveal high cigarette smoking prevalence in low-income cities of Los Angeles County. *J Urban Health*. 2012;893:397–406.
9. Szatkowski L, Fahy SJ, Coleman T, et al. Small area synthetic estimates of smoking prevalence during pregnancy in England. *Popul Health Metr*. 2015;13:34.
10. Mohan J, Twigg L, Barnard S, et al. Social capital, geography and health: a small-area analysis for England. *Soc Sci Med*. 2005;60:1267–83.
11. Schneider KL, Lapane KL, Clark MA, et al. Using small-area estimation to describe county-level disparities in mammography. *Prev Chronic Dis*. 2009;64:A125.
12. Wright J, Williams R, Wilkinson JR. Development and importance of health needs assessment. *BMJ*. 1998;316(7140):1310–3.
13. R. Extract lme Random Effects. <https://stat.ethz.ch/R-manual/R-devel/library/nlme/html/ranef.lme.html>. Geraadpleegd op: 8 augustus 2017.
14. MIWin. Parameters and parameter estimates. <http://www.bristol.ac.uk/cmm/software/support/support-faqs/parameters.html>. Geraadpleegd op: 8 augustus 2017.
15. Centraal Bureau voor de Statistiek (CBS). Kerncijfers Wijken en Buurten. Den Haag: CBS; 2016. [www.cbs.nl/nl-NL/menu/methoden/dataverzameling/kerncijfers-wijkbuurt-kob.htm](http://www.cbs.nl/nl-NL/menu/methoden/dataverzameling/kerncijfers-wijkbuurt-kob.htm). Geraadpleegd op: 8 augustus 2017.
16. Brink CL van den, Savelkoul M. Gezondheidsmonitor GGD'en, CBS en RIVM. In: *Volksgezondheid Toekomst Verkenning, Nationaal Kompas Volksgezondheid*. Bilthoven: RIVM; 2014. <https://www.volksgezondheidenzorg.info/>. Geraadpleegd op: 8 augustus 2017.
17. Newcombe RG. Two sided confidence intervals for the single proportion: comparison of seven methods. *Stat Med*. 1998;17:857–72.