

Comparing the Health of Populations: Methods to Evaluate and Tailor Population Management Initiatives in the Netherlands

Roy J.P. Hendriks, MSc,¹ Hanneke W. Drewes, PhD,² Marieke Spreeuwenberg, PhD,^{3,4}
Dirk Ruwaard, PhD,³ and Caroline A. Baan, PhD^{1,2}

Abstract

Health care no longer focuses solely on patients and increasingly emphasizes regions and their populations. Strategies, such as population management (PM) initiatives, aim to improve population health and well-being by redesigning health care and community services. Hence, insight into population health is needed to tailor interventions and evaluate their effects. This study aims to assess whether population health differs between initiatives and to what extent demographic, personal, and lifestyle factors affect these differences. A population health survey that included the Short Form 12 version 2 (SF12, physical and mental health status), Patient Activation Measure 13 (PAM13), and demographic, personal, and lifestyle factors was administered in 9 Dutch PM initiatives. Potential confounders were determined by comparing these factors between PM initiatives using analyses of variance and chi-square tests. The influence of these potential confounders on the health outcomes was studied using multivariate linear regression. Age, education, origin, employment, body mass index, and smoking were identified as potential confounders for differences found between the 9 PM initiatives. Each had a noteworthy influence on all of the instruments' scores. Not all health differences between PM initiatives were explained, as the SF12 outcomes still differed between PM initiatives once corrected. For the PAM13, the differences were no longer significant. Demographic and lifestyle factors should be included in the evaluation of PM initiatives and population health differences found can be used to tailor initiatives. Other factors beyond health care (eg, air quality) should be considered to further refine the tailoring and evaluation of PM initiatives.

Keywords: population health, outcomes measurement, Triple Aim, evaluation

Background

TO ENSURE HEALTH CARE SYSTEMS are sustainable, current policies go beyond patients and focus on regions and (general) populations. These so-called population management (PM) initiatives, which aim to improve population health and quality of care while also reducing costs (Triple Aim), are increasingly being introduced.¹ The implementations of such

initiatives are a response to the increasing financial pressure exerted on health care systems by aging populations and new and expensive technologies.^{2,3} To transform health care systems, policy makers are introducing PM initiatives that address the full continuum of a defined population's health and well-being, and implementing interventions that integrate health care, prevention, and social services.⁴ Because of the surge of regional policy, PM, and the Triple Aim, the concept of population

¹Tranzo Scientific Center for Care and Welfare, Tilburg School of Social and Behavioral Sciences, Tilburg University, Tilburg, the Netherlands.

²Department for Quality of Care and Health Economics, Center for Nutrition, Prevention and Health Services, National Institute for Public Health and the Environment, Bilthoven, the Netherlands.

³Department of Health Services Research, Faculty of Health, Medicine and Life Sciences, CAPHRI School for Public Health and Primary Care, Maastricht University, Maastricht, the Netherlands.

⁴Research Centre for Technology in Care, Zuyd University of Applied Sciences, Heerlen, the Netherlands.

Prior Presentation: This study was presented as a poster presentation at the International Conference on Integrated Care 2017, held May 8–10, 2017 in Dublin, Ireland

health has gained traction among policy makers and researchers. A more in-depth understanding of population health has become essential to evaluate and tailor PM initiatives.⁵

Population health as a concept has been widely discussed^{6,7} but, despite the definition by Kindig and Stoddart,⁸ no consensus exists regarding the meaning of the term. For the purpose of this study, population health is considered to be the health of an entire population in 1 geographic area. Within this understanding of population health, many factors are thought to affect (population) health.^{9–11} A population's education level, for example, has an association with health, but also affects health indirectly through its impact on health behaviors and use of preventive services.¹² Other characteristics that can influence health include sex, age, and ethnicity.¹³ Those who wish to assess health in PM initiatives or evaluate the effectiveness of their interventions should consider such variables.^{4,14} Health assessments help design interventions for PM initiatives by identifying priority areas and, from an evaluation point of view, controlling for these variables will provide baseline measurements for population health that are more comparable between different initiatives (overt bias).¹⁵

In the Netherlands, the National Monitor Population Management (NMP) was created to gather knowledge regarding the experiences of stakeholders within 9 PM initiatives as well as their performance.¹⁶ These so-called pioneer sites each focus on a defined regional population and aim to achieve the Triple Aim. Even though their goals are similar, they are locally introduced initiatives that operate independent of each other. In most, local municipalities, hospitals, general practitioners, and insurance companies are involved on the board, supplemented by more intervention-specific stakeholders (eg, pharmacies, research institutes). Most of these interventions have the goal of improving prevention services and reducing secondary care use, but each of the pioneer sites has its own distinct set of interventions. More detailed information can be found in online Supplementary Data S1 (Supplementary Data are available online at www.liebertpub.com/pop) and elsewhere.¹⁶ To assess population health in these pioneer sites, different constructs were measured by the NMP, including health status (Short Form 12 version 2 [SF12]) and self-management (Patient Activation Measure 13 [PAM13]).¹⁷ In addition to these 2 constructs, demographic (eg, age, education) and other factors (eg, body mass index [BMI]) were included. However, the impact of these demographic, personal, and lifestyle factors on population health differences is unknown.

In order to gain a better understanding of how to evaluate and tailor PM initiatives, this study aims to assess whether population health differs between initiatives and to what extent this relationship is affected by demographic and lifestyle factors.

Methods

Ethics approval and consent to participate

The Psychological Ethics Committee of the Tilburg University (Psychological Ethical Testing Committee number: EC-2014.39) approved this study.

Study population

The study population consisted of citizens living in 9 Dutch pioneer sites. The NMP provided the baseline data used for analysis, which were gathered between December of

2014 and January of 2015. A survey was sent out in each site to a random sample of 600 insured adults (≥ 18 years old) who, within the previous year, had not received any surveys from insurance companies and remained registered with the same general physician. Participants were invited by mail to complete the survey, which could be done online or on paper. Participants who did not respond were sent 2 reminders.¹⁷

Survey instruments

Demographic, personal, and lifestyle factors. The survey included demographic characteristics (sex, age, educational level [high/low]), employment, and country of origin [native/non-native]), as well as personal and lifestyle characteristics. These included levels of disability (100% disabled), alcohol use (glasses per week), and smoking status (yes/no). Furthermore, BMI was calculated using the reported height and weight, and health literacy was assessed using Chew et al's Set of Brief Screening Questions.¹⁸

SF-12. The SF-12, a globally used instrument, assesses generic health status using 12 questions that produce a physical component score (PCS) and a mental component score (MCS).¹⁹ The Dutch version was used for this study.²⁰ Scoring was done using the proprietary software provided by Qualimetrics Inc. (Sacramento, CA), which associates set weights with questions' scores that are converted to a 0–100 scale for both the PCS and MCS. In this scale a higher score means better (physical or mental) health.¹⁹

PAM13. The PAM13 is a scale that is based on a developmental model of activation²¹ and can be used to assess self-reported knowledge, motivation, and health management skills.²² The questions' scores are summed and converted to a 0–100 scale; a higher score is positively associated with related health behaviors.²³ The Dutch version was used.²³

Analysis

All analyses were performed using SPSS 22 (IBM Corporation, Armonk, NY) and R Studio Version 0.99.441 (RStudio Inc., Boston, MA) for Windows.

First, responses to the SF12 and PAM13 were analyzed. Participants who provided a complete data set per instrument were compared regarding demographic, personal, and lifestyle factors with participants having 1 or more missing values for that instrument. If these groups differed, then it could be assumed that missing values were not Missing-Completely-At-Random and therefore would need to be imputed. MICE (multiple imputation by chained equation) would be used to impute the data²⁴ and further analyses would be performed on both complete cases and imputed data sets. The methodology is discussed in a previous and related article.²⁵

Second, an overview of demographic, personal, and lifestyle factors (independent variables; eg, age, education, sex, origin, employment, alcohol use) was created for all 9 pioneer sites. This overview also included the Pearson chi-square test for dichotomous variables and analysis of variance (ANOVA) for continuous variables to determine whether there were significant differences between the 9 pioneer sites. Factors with a P value ≤ 0.10 were considered potential confounders.²⁶

Third, separate univariate analyses were conducted for each outcome (SF12-PCS, SF12-MCS and PAM13) using

TABLE 1. CHARACTERISTICS OF THE SAMPLE POPULATION

Population	Blauwe zorg	Friesland Voorop	Goed Leven	Mijn Zorg	GZGR	PELGRIM	SSiZ	SmZ	Vitaal Vechtdal	ANOVA/chi-square	Study population
Total citizens ²⁸	176,055	646,910	106,270	273,500	183,920	417,780	273,340	516,500	112,655	-	-
Surveys sent	600	600	600	600	600	600	600	600	600	-	5400
Returned surveys	256	293	272	259	272	272	308	280	279	-	2491
Response rate (%)	42.7	48.8	45.3	43.2	45.3	45.3	51.3	46.7	46.5	-	46.1
Sex (% male)	49.8	43.8	48.5	49.4	47.0	44.4	43.0	45.4	44.3	0.675	46.1
Age (Standard deviation)	57.9 (16.3)	54.3 (16.6)	55.1 (15.8)	59.1 (14.0)	54.5 (17.3)	54.7 (15.3)	59.0 (15.7)	54.7 (16.9)	51.6 (15.1)	0.000	55.7 (16.1)
Education (% highly educated)	34.9	26.7	20.1	18.8	42.4	28.0	22.0	27.1	12.7	0.000	25.8
Origin (% native)	84.1	95.4	80.5	77.9	85.7	84.8	87.5	87.1	93.8	0.000	86.4
Employed (% paid job)	46.9	49.1	48.3	41.2	51.7	50.6	45.2	51.4	63.1	0.000	49.7
Disabled (%)	5.9	3.8	4.2	6.6	5.2	2.6	2.3	5.4	3.1	0.145	4.3
BMI	26.1	25.9	25.9	26.7	25.4	26.4	25.4	25.8	26.0	0.011	26.0
Alcohol use (glasses per week)	3.8	4.7	3.9	3.7	4.6	4.0	5.1	4.6	4.2	0.144	4.3
Smoking (% smokers)	16.5	16.2	20.1	20.1	17.0	14.9	13.7	21.2	23.6	0.048	18.1
Health Literacy (score Chew et al's Set of Brief Screening Questions)	3.4	3.5	3.3	3.4	3.4	3.3	3.4	3.3	3.3	0.230	3.4
SF12-PCS	49.7	50.9	48.7	46.5	50.0	50.2	50.1	49.7	51.4	0.000	49.7
SF12-MCS	49.0	51.9	49.6	48.4	48.5	50.0	50.0	49.5	51.3	0.000	49.8
PAM13	59.5	58.2	57.8	60.0	61.8	58.0	60.9	57.0	55.2	0.000	58.8

ANOVA, analysis of variance; BMI, body mass index; GZGR, Gezonde Zorg, Gezonde Regio; SF12-MCS, Short Form 12 version 2 Mental Component Score; SF12-PCS, Short Form 12 version 2 Physical Component Score; PAM13, Patient Activation Measure 13; SmZ, Slimmer met Zorg; SSiZ, Samen Sterk in Zorg.

regression analyses. The PM initiatives were coded as dummy variables, with one of the pioneer sites (Blauwe Zorg) as the reference group. Model fit was assessed using the F-ratio and a *P* value of ≤ 0.05 was considered significant in all regression analyses.²⁷ Then, the influence of each individual potential confounder was assessed using multivariate linear regression analyses. A variable was defined as a confounder when the average change in the dummy variables compared to the univariate model was more than 10% after adding that particular variable.²⁶ At each step, starting with the univariate model, the variable that produced the largest average change in the beta of dummy variables (and was more than 10%) was then added to the model. The remaining variables were then added step by step to this model. This process was repeated until there were no variables left. Finally, for comparison, a multiple regression analysis was executed with all potential confounders.

Results

The number of total citizens within the pioneer sites ranged from 106,270 to 646,910 (Table 1). A total of 5400 surveys were sent out, 2491 of which were completed and returned. The response analyses showed that groups with no missing values differed significantly from participants with 1 or more missing values on at least 1 demographic variable for all instruments (see online Supplementary Data S2). This indicates that missing values were not Missing-Completely-At-Random, warranting the use of multiple imputation. Subsequent analyses were therefore performed on both complete cases and imputed data sets.

The study sample was 46.1% male, consisted of mostly Dutch natives, and was on average 55.7 years old. Between PM initiatives, there were significant differences in age, education, origin, employment, BMI, and smoking, identifying these factors as potential confounders. Overall, Mijn Zorg was the only initiative that showed consistently below-average scores on all variables. Most PM initiatives scored better on one and worse on other variables. For example, Gezonde Zorg, Gezonde Regio had the highest score on education level, but also had a higher than average unemployment rate. As expected, descriptive results from the imputed data were comparable (see online Supplementary Data S3).

An overall ANOVA showed that outcomes of the SF12-PCS, SF12-MCS, and PAM13 differed significantly between PM initiatives before adding confounders to the model. The more detailed univariate regression model showed that for each instrument, there were 1 or 2 PM initiatives that differed significantly from the reference group (Table 2). For the SF12-PCS, Mijn Zorg was the only significantly lower deviation, while Friesland Voorop and Vitaal Vechtdal scored significantly higher for the SF12-MCS. Vitaal Vechtdal scored significantly lower on the PAM13 as well, which also was the case with the imputed data (see online Supplementary Data S3).

Per instrument, the aforementioned potential confounders were entered in the regression model individually. Extended results of these regression analyses can be seen in an additional file (Supplementary Data S4). Each characteristic had a noteworthy (>10%) influence on each of the instruments and was included in the final analyses. For the SF12-PCS, the largest confounders were age, BMI, and employment. Age also was an important influencer in the SF12-MCS, in addition to origin and smoking status. PAM13 was influenced mainly by education, employment, and smoking status. Somewhat different results were found in the imputed data for PAM13 (Supplementary Data S4). Here, education and employment still had the largest effect on PAM13 scores, but smoking status was no longer a confounder. Origin also was not a confounder, although age and BMI still influenced outcomes.

Table 2 shows the results of the (multivariate) linear regression models including all confounders, each showing a significant F-ratio indicating sufficient model fit. For the SF12-PCS, the Table shows that all PM initiatives had noteworthy changes ($\geq 10\%$), but no differences changed from significant to not significant or vice versa. Mijn Zorg remained the only significant deviation. In mental health (SF12-MCS), Vitaal Vechtdal initially scored better. However, this difference was nuanced when results were controlled for specific population factors. Friesland Voorop, like Vitaal Vechtdal, also scored better on the SF12-MCS, but the difference remained roughly the same after controlling for confounders. For the PAM13, Vitaal Vechtdal was the only significantly different score before controlling for confounders. After controlling for confounders, the difference was no longer

TABLE 2. RESULTS OF REGRESSION ANALYSES

<i>Instrument</i>	<i>F-test</i>	<i>Blauwe zorg(Reference)</i>	<i>Friesland Voorop</i>	<i>GoedLeven</i>	<i>Mijn Zorg</i>	<i>GZGR</i>	<i>PELGRIM</i>	<i>SSiZ</i>	<i>SmZ</i>	<i>Vitaal Vechtdal</i>
SF12-PCS										
Crude	4.7*	49.748	1.193	-1.025	-3.174*	0.254	0.460	0.364	-0.057	1.703
Adjusted	34.3*	63.507	0.505	-1.266	-2.288*	-0.938	0.330	0.530	-1.293	0.550
Adjusted mean	-	63.507	64.012	62.241	61.219	62.569	63.837	64.037	62.214	64.057
SF12-MCS										
Crude	3.7*	48.965	2.926*	0.631	-0.594	-0.457	0.993	1.024	0.577	2.360*
Adjusted	7.7*	44.647	2.998*	0.622	-0.869	-0.826	1.166	1.264	0.033	1.958*
Adjusted mean	-	44.647	47.645	45.269	43.778	43.821	45.813	45.911	44.68	46.605
PAM13										
Crude	2.7*	59.463	-1.252	-1.625	0.506	2.371	-1.428	1.454	-2.434	-4.303*
Adjusted	6.2*	59.764	-0.452	-0.827	1.627	1.268	-1.190	2.218	-1.682	1.672
Adjusted mean	-	59.764	59.312	58.937	61.391	61.032	58.574	61.982	58.082	61.436

GZGR, Gezonde Zorg, Gezonde Regio; SF12-PCS, Short Form 12 – Physical Component Score; SF12-MCS, Short Form 12 – Mental Component Score; PAM13, Patient Activation Measure 13; SmZ, Slimmer met Zorg; SSiZ, Samen Sterk in Zorg.

* *P* < 0.05

significant, resulting in no significant differences between PM initiatives on PAM13. This result also was seen in the imputed data (see online Supplementary Data S3).

Discussion

This study compared health within 9 Dutch PM initiatives using the SF12-PCS, SF12-MCS, and PAM13 instruments as outcomes and examined to what extent these are affected by demographic, personal, and lifestyle factors. The intention was to provide the needed insight to improve evaluation and tailor the interventions of these initiatives. The included physical (SF12-PCS), mental (SF12-MCS), and self-management (PAM13) constructs showed differences between initiatives before controlling for any population factors. This was mainly because of 1 or 2 outlying PM initiatives. After controlling for confounding factors, differences between PM initiatives were nuanced, but the SF12-PCS and the SF12-MCS still showed significant differences. For PAM13, the only significant difference from the reference group became insignificant. Age and origin as well as education, employment, and smoking had a large influence on differences between initiatives.

The effects of demographic, personal, and lifestyle factors were in line with expectations, as this effect has been seen for various instruments at the individual level,^{23,29,30} as well as the population level.^{31,32} Nonetheless, although adjusted results provide a clearer image for evaluation purposes, as improving health should be the end goal, unadjusted differences should not be disregarded. The impact variables have on differences in health can be used to tailor interventions for specific populations.³³ The characteristics with the largest impact on outcomes were age, origin, education, employment, and smoking. A number of these factors, including health literacy, BMI, smoking, and alcohol use, can be addressed by (health) interventions. Smoking behavior, for example, can be addressed by implementing smoking bans in schools.³⁴ In this manner, PM initiatives can focus interventions on variables that are shown to affect the health of their population.

Some reservations must be considered when interpreting these results. Data were collected at a single point in time and it would be of interest to compare changes over time to see the impact of interventions in a particular region. Furthermore, the provided recommended proprietary software was used to calculate the SF12-MCS and SF12-PCS, ensuring the correct calculation of both SF12 component scores. Unfortunately, even though the use of this software is recommended, it prevented the use of imputed data to calculate the sum scores. Results from PAM13 imputed data showed that differences, albeit small, could occur. Response rates in most pioneer sites were ~45%, which sufficed and is comparable to other surveys,³⁵ but because of limited information regarding the sample population, it was not possible to assess selection bias in depth.

Finally, even though the confounders studied did explain some differences, not all of the variation in population health between PM initiatives was explained. For example, Friesland Voorop and Vitaal Vechtdal still showed significantly higher scores for the SF12-MCS than the reference group after controlling for confounders. Whether these differences are clinically relevant is difficult to establish. For example, for the SF12, studies consider a range of relevant differences,

ranging from 2.5 points and higher.^{36,37} This would mean that the differences in this study are not clinically relevant, but it is debatable whether such a hard value can be set for populations. From an evaluation perspective, a baseline measurement should be as equal as possible, as differences after controlling for confounders might indicate the presence of hidden bias.³⁸ Efforts, such as the research into the “Limburg-factor,”³⁹ should be made to seek out hidden bias. This can be done, for instance, by examining the literature more deeply, integrating qualitative analyses, or by looking for influencers elsewhere (eg, environmental factors such as air quality and the availability of green space)^{40,41} However, many of these factors are currently seen as beyond the scope of responsibility of health care, while having a considerable impact on population health. This emphasizes the importance of wider integration of care domains by PM initiatives to ensure complete coverage of health.

Conclusion

At first glance, the 9 populations compared in this study differed in physical health (SF12-PCS), mental health (SF12-MCS), and self-management capacity (PAM13). However, once the identified confounders were included in the model, these differences became smaller and were no longer significant for self-management capacity. The impact the current confounders have on results can be used to guide future evaluations and tailoring of PM initiatives. Unexplained differences in health between PM initiatives require further investigation.

Author Disclosure Statement

The authors declare that there are no conflicts of interest. The authors received the following financial support: This study was funded under SPR project S/133002 of The National Institute of Public Health and the Environment in the Netherlands. The funder had no role in the design of the study, collection, analysis and interpretation of the data, and writing of the manuscript.

References

1. Steenkamer BM, Drewes HW, Heijink R, Baan CA, Struijs JN. Defining population health management: a scoping review of the literature. *Popul Health Manag* 2017;20:74–85.
2. OECD. Health Care Systems: efficiency and policy settings. Paris: Organisation for Economic Co-operation and Development, 2010.
3. OECD. Health Care Systems: getting more value for money. Paris: Organisation for Economic Co-operation and Development, 2010.
4. Struijs JN, Drewes HW, Heijink R, Baan CA. How to evaluate population management? Transforming the Care Continuum Alliance population health guide into a broadly applicable analytical framework. *Health Policy* 2015;119:522–529.
5. Tolonen H, Koponen P, Mindell JS, et al. Under-estimation of obesity, hypertension and high cholesterol by self-reported data: comparison of self-reported information and objective measures from health examination surveys. *Eur J Public Health* 2014;24:941–948.
6. Etches V, Frank J, Di Ruggiero E, Manuel D. Measuring population health: a review of indicators. *Ann Rev Public Health* 2006;27:29–55.

7. Vuik S, Siegel S, Darzi A. How should we measure the distribution of health in a population? 2017. <http://healthaffairs.org/blog/2017/03/17/how-should-we-measure-the-distribution-of-health-in-a-population/> Accessed March 22, 2017.
8. Kindig D, Stoddart G. What is population health? *Am J Public Health* 2003;93:380–383.
9. Parrish RG, McDonnell SM, Remington PL. Surveillance for determinants of population health. In: Lee LM, Teutsch SM, Thacker SB, St Louis ME, eds. *Principles and practice of public health surveillance*, 3rd edition. Oxford: Oxford University Press, 2010:275–305.
10. Zavras D, Tsiantou V, Pavi E, Mylona K, Kyriopoulos J. Impact of economic crisis and other demographic and socioeconomic factors on self-rated health in Greece. *Eur J Public Health* 2013;23:206–210.
11. Molarius A, Berglund K, Eriksson C, et al. Socioeconomic conditions, lifestyle factors, and self-rated health among men and women in Sweden. *Eur J Public Health* 2007;17:125–133.
12. Feinstein L, Sabates R, Anderson TM, Sorhaindo A, Hammond C. What are the effects of education on health? Measuring the Effects of Education on Health and Civic Engagement: proceedings of the Copenhagen Symposium. Copenhagen: Organisation for Economic Co-operation and Development, 2006:171–354.
13. Dahlgren G, Whitehead M. Policies and strategies to promote social equity in health: background document to WHO—Strategy paper for Europe. Copenhagen: World Health Organization Regional Offices for Europe, 1991.
14. Linden A, Adams JL, Roberts N. Strengthening the case for disease management effectiveness: un-hiding the hidden bias. *J Eval Clin Pract* 2005;12:140–147.
15. Imbos T, Janssen MPE, Berger MPF. *Methodologie en Statistiek 1*, vol. 4. Maastricht, Netherlands: Universitaire Pers Maastricht, 2008.
16. Drewes HW, Struijs JN, Baan CA. How the Netherlands is integrating health and community services. 2016. <http://catalyst.nejm.org/netherlands-integrating-health-community-services/> Accessed November 14, 2016.
17. Drewes HW, Heijink R, Struijs JN, Baan CA. Samen werken aan duurzame zorg. Landelijke monitor Proeftuinen. Bilthoven, Netherlands: National Institute for Public Health and the Environment, 2015.
18. Chew LD, Bradley KA, Bokyo EJ. Brief questions to identify patients with inadequate health literacy. *Fam Med* 2004;36:588–594.
19. Ware JE, Kosinski MA, Turner-Bowker DM, Gandek B. User's manual for the SF-12v2[®] health survey with a supplement documenting SF-12[®] health survey. Lincoln, RI: QualityMetric Incorporated, 2002.
20. Aaronson NK, Muller M, Cohen PDA, et al. Translation, validation, and norming of the Dutch language version of the SF-36 Health Survey in community and chronic disease populations. *J Clin Epidemiol* 1998;51:1055–1068.
21. Hibbard JH, Stockard J, Mahoney ER, Tusler M. Development of the Patient Activation Measure (PAM): conceptualizing and measuring activation in patients and consumers. *Health Serv Res* 2004;39:1005–1026.
22. Hung M, Carter M, Hayden C, et al. Psychometric assessment of the patient activation measure short form (PAM-13) in rural settings. *Qual Life Res* 2013;22:521–529.
23. Rademakers J, Nijman J, van der Hoek L, Heijmans M, Rijken M. Measuring patient activation in The Netherlands: translation and validation of the American short form Patient Activation Measure (PAM13). *BMC Public Health* 2012;12:577.
24. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. *J Stat Softw* 2011;45:1–67.
25. Hendriks RJP, Spreeuwenberg M, Drewes HW, Ruwaard D, Baan CA. How to measure population health: an exploration towards an integration of valid and reliable instruments. *Popul Health Manage* (Submitted).
26. Twisk JWR. *Inleiding in de toegepaste biostatistiek*, vol. 4. Amsterdam: Reed Business Education, 2014.
27. Field A. *Discovering statistics using IBM SPSS statistics*, 4th edition. London: Sage, 2014.
28. CBS; Statistics Netherlands. Statline. 2015. <http://statline.cbs.nl/Statweb/> Accessed June 9, 2017.
29. Fleishman JA, Lawrence WF. Demographic variation in SF-12 scores: true differences or differential item functioning? *Med Care* 2003;41:75–86.
30. Krynen AM, Osborne D, Duck IM, Houkema CA, Sibley CG. Measuring psychological distress in New Zealand: item response properties and demographic differences in the Kessler-6 screening measure. *N Z J Psychol* 2013;42:95–109.
31. Audureau E, Rican S, Coste J. From deindustrialization to individual health-related quality of life: multilevel evidence of contextual predictors, mediators and modulators across French regions, 2003. *Health Place* 2013;22:140–152.
32. Stallman HM. Psychological distress in university students: a comparison with general population data. *Aust Psychol* 2010;45:249–257.
33. Berkman LF. Unintended consequences of social and economic policies for population health: towards a more intentional approach. *Eur J Public Health* 2011;21:547–548.
34. Piontek D, Buehler A, Rudolph U, et al. Social contexts in adolescent smoking: does school policy matter? *Health Educ Res* 2008;23:1029–1038.
35. Nulty DD. The adequacy of response rates to online and paper surveys: what can be done? *Assess Eval High Educ* 2008;33:301–314.
36. Clement ND, MacDonald D, Simpson AH. The minimal clinically important difference in the Oxford knee score and Short Form 12 score after total knee arthroplasty. *Knee Surg Sports Traumatol Arthrosc* 2014;22:1933–1939.
37. Singh J, Sloan JA, Johanson NA. Challenges with health-related quality of life assessment in arthroplasty patients: problems and solutions. *J Am Acad Orthop Surg* 2010;18:72–82.
38. Rosenbaum PR. Observational study. In: Everitt BS, Howell DC, eds. *Encyclopedia of statistics in behavioral science*, vol. 3. Chichester: John Wiley & Sons, 2005:1451–1462.
39. Jansen M, Kuppens E. *Op zoek naar de Limburg-factor*. Geleen: GGD Zuid Limburg, 2015.
40. Kampa M, Castanas E. Human health effects of air pollution. *Environ Pollut* 2008;151:362–367.
41. Maas J, Verheij RA, Vries Sd, Spreeuwenberg P, Schellevis FG, Groenewegen PP. Morbidity is related to a green living environment. *J Epidemiol Community Health* 2009;63:967–973.

Address correspondence to:

Roy J.P. Hendriks, MSc

Department for Quality of Care and Health Economics

Center for Nutrition, Prevention and Health Services

National Institute for Public Health and the Environment

PO Box 1

Bilthoven 3720 BA

The Netherlands

E-mail: roy.hendriks@rivm.nl