


## RESEARCH ARTICLE

# Identifying subgroups of high-need, high-cost, chronically ill patients in primary care: A latent class analysis

Rowan G. M. Smeets <sup>\*</sup>, Arianne M. J. Elissen, Mariëlle E. A. L. Kroese, Niels Hameleers, Dirk Ruwaard

Department of Health Services Research, Maastricht University, Faculty of Health, Medicine and Life Sciences, Care and Public Health Research Institute (CAPHRI), Maastricht, The Netherlands

\* [rowan.smeets@maastrichtuniversity.nl](mailto:rowan.smeets@maastrichtuniversity.nl)



## Abstract

### Introduction

Segmentation of the high-need, high-cost (HNHC) population is required for reorganizing care to accommodate person-centered, integrated care delivery. Therefore, we aimed to identify and characterize relevant subgroups of the HNHC population in primary care by using demographic, biomedical, and socioeconomic patient characteristics.

### Methods

This was a retrospective cohort study within a Dutch primary care group, with a follow-up period from September 1, 2014 to August 31, 2017. Chronically ill patients were included in the HNHC population if they belonged to the top 10% of care utilizers and/or suffered from multimorbidity and had an above-average care utilization. In a latent class analysis, forty-one patient characteristics were initially used as potential indicators of heterogeneity in HNHC patients' needs.

### Results

Patient data from 12 602 HNHC patients was used. A 4-class model was considered statistically and clinically superior. The classes were named according to the characteristics that were most dominantly present and distinctive between the classes (i.e. mainly age, household position, and source of income). Class 1 ('older adults living with partner') included 39.3% of patients, class 2 ('older adults living alone') included 25.5% of patients, class 3 ('middle-aged, employed adults with family') included 23.3% of patients, and class 4 ('middle-aged adults with social welfare dependency') included 11.9% of patients. Diabetes was the most common condition in all classes; the second most prevalent condition differed between osteoarthritis in class 1 (21.7%) and 2 (23.8%), asthma in class 3 (25.3%), and mood disorders in class 4 (23.1%). Furthermore, while general practitioner (GP) care utilization increased during the follow-up period in the classes of older adults, it remained relatively stable in the middle-aged classes.

### OPEN ACCESS

**Citation:** Smeets RGM, Elissen AMJ, Kroese MEAL, Hameleers N, Ruwaard D (2020) Identifying subgroups of high-need, high-cost, chronically ill patients in primary care: A latent class analysis. PLoS ONE 15(1): e0228103. <https://doi.org/10.1371/journal.pone.0228103>

**Editor:** Bruno Pereira Nunes, Universidade Federal de Pelotas, BRAZIL

**Received:** September 13, 2019

**Accepted:** January 7, 2020

**Published:** January 29, 2020

**Peer Review History:** PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0228103>

**Copyright:** © 2020 Smeets et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Results are based on calculations by the department of Health Services Research, Maastricht University, the Netherlands, using non-public microdata from Statistics

Netherlands and claims data from health care information center 'Vektis'. Under certain conditions, these microdata are accessible for statistical and scientific research. For further information and access queries: ([microdata@cbs.nl](mailto:microdata@cbs.nl)). In addition, data from general practices connected to primary care group 'HZD' were used for the analysis. The data from the general practices are owned by the general practices and 'HZD'; requests for access to this data should be submitted to 'HZD' ([info@hzd.nu](mailto:info@hzd.nu))

**Funding:** The current study was funded by primary care group 'HZD', the Netherlands (<https://www.hzd.nu>). The funding agency had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** AE is currently serving as an Academic Editor for PLOS ONE. Furthermore, our funding source (i.e. primary care group 'HZD') is a commercial source. This does not alter our adherence to PLOS ONE policies on sharing data and materials.

## Conclusions

Although the HNHC population is heterogeneous, distinct subgroups with relatively homogeneous patterns of mainly demographic and socioeconomic characteristics can be identified. This calls for tailoring care and increased attention for social determinants of health.

## Introduction

Due to increasing numbers of chronically ill patients, in particular with multimorbidity, and rising health care costs, Western health care systems are faced with challenges to deliver high-quality, person-centered, and sustainable care [1–3]. In response to these developments, accountable care organizations (ACOs) were introduced in the United States several years ago [4–6]. Within ACOs, a value-based payment system is designed to incentivize providers to share accountability for the quality and cost of care for a defined population [6–8]. Likewise, more than a decade ago, 'care groups' were first introduced in Dutch primary care. In line with the ACOs, care groups unite providers, mostly general practitioners (GPs), with shared responsibility for all assigned patients receiving care for a specific chronic condition from a value-based bundled payment approach [5, 9]. These initiatives show that, similar to the US, the Netherlands aims to achieve more value-based care.

If health systems aim to increase the value of delivered care, it is crucial to focus on the population with the highest care use as they offer the largest potential for achieving improved value [10, 11]. This population with a disproportionately high care use is also referred to as the high-need, high-cost (HNHC) population [10, 12]. The identification of the HNHC population, as a subgroup of the total population, is embedded in the approach of population segmentation, which is defined as the division of a specific population into homogeneous subgroups with distinct needs and (health) characteristics [13–15]. A closely related concept in which principles of segmentation are applied, pertains to the concept of 'population (health) management' (PM), [16] as a way to promote 'population health' [17, 18]. Within population health, the focus is on the health outcomes of subgroups rather than individuals, by taking into account a large variety of determinants of health (i.e. physical, mental, social) [17, 18]. PM strategies generally aim to improve health needs of defined subgroups along 'the continuum of health and well-being', and aim to integrate services across multiple domains [16]. As such, PM strategies can be used to tailor interventions to the care needs of specific subgroups of patients, which is assumed to lead towards improving individual patients', and providers' experiences as well as population outcomes and cost (Quadruple Aim [19]).

With the growing availability of digital patient data, studies have identified common biomedical characteristics of the HNHC population, such as the high prevalence of (co-occurring) chronic conditions and mental illness [20, 21]. At the same time, studies have suggested that the HNHC population is diverse, not only in terms of patients' biomedical but also in their demographic and socioeconomic profiles [10, 20, 21]. These findings underline the importance of social determinants of health within the HNHC population. Yet, population segmentation studies have predominantly focused on specific populations, such as older adults [22–24] and Medicaid beneficiaries [25], and mainly characterized the identified patient subgroups by their biomedical characteristics (i.e., chronic diagnoses) [22–27]. Therefore, the main aim of this study was to identify and characterize, by means of latent class analysis (LCA), clinically relevant subgroups of the HNHC population in primary care, defined by demographic, biomedical, and socioeconomic patient characteristics as well as care utilization.

## Materials and methods

### Setting

This retrospective cohort study was conducted at a (primary) care group in the northern region of the Netherlands, covering 130 general practices. This care group was founded in 2009 and currently has bundled payment contracts with health insurers for the delivery of several disease management programs, including for patients with type 2 diabetes mellitus, COPD, and cardiovascular risks.

As this study used retrospective data and did not intervene into people's life or impose rules, no formal ethical approval was required (project number 164111), in line with the Dutch Medical Research (Human Subjects) Act.

### Data sources

All general practices connected to the care group were invited to extract and provide individual-level patient data from their electronic health records (EHRs). The EHR data covered 4.5 years: baseline was on September 1, 2014; the follow-up period covered three years (from September 1, 2014 to August 31, 2017). Furthermore, the EHR data were linked on the individual patient level to socioeconomic data (e.g., source of income) and health care claims data (e.g., pharmaceutical costs). Socioeconomic data were retrieved from Statistics Netherlands, which is involved in the collection, preparation, and publication of statistics on behalf of the Dutch government, science and commercial sector [28]. Claims data were retrieved from the health care information center 'Vektis', which collects and manages all claims under the Dutch Healthcare Insurance Act [29]. To ensure data confidentiality and safety, a third trusted party was involved in the provision of a pseudonymized version of the data set to the researchers.

### Participants

We selected a cohort of chronically ill patients, limited to those with a full EHR registration over the 4.5-year research period. Patients were considered chronically ill if they had registered at least one GP consultation in the 1.5 years before baseline related to one of 28 conditions defined as chronic (see Table 1) [30, 31]. Chronically ill patients were included in the HNHC population if they belonged to the top 10% of care utilizers (over follow-up period) and/or suffered from multimorbidity and had an above-average care utilization (over follow-up period). The first criterion was applied as this is one of the commonly used thresholds for identifying HNHC patients according to previous studies [20, 32, 33]. The second criterion was applied because multimorbidity brings along a challenging complexity to the organization of care, especially in light of the current single-disease management programs for single chronic conditions [2, 3]. Furthermore, care utilization was measured as the total number of GP consultations weighted by the required time investment per type of consultation (i.e. 0.5 for telephone or e-mail consultation, 1.0 for regular consultation, 2.0 for extended regular consultation, 1.5 for home visit, 2.5 for extended home visit), determined by the Netherlands Institute for Health Services Research [34]. As the weighting factors based on time investment are related to costs [35], the patients selected for this study can be considered high-need, high-cost in primary care.

### Variables

Forty-one patient characteristics were initially used as potential indicators of heterogeneity in HNHC patients' needs in the LCA. These characteristics were included based on scientific studies describing these characteristics as relevant in relation to (high) care utilization [12, 36].

**Table 1. Baseline characteristics of the HNHC population (n = 12 602).**

Patient characteristics		
	n (%)	Missing, n (%)
Demographic characteristics		
Sex		0
Male	4495 (35.67)	
Female	8107 (64.33)	
Age, mean (SD) <sup>a</sup>	67.55 (14.80)	0
Household position		
Child living at home	141 (1.12)	
Single adult	3773 (29.94)	
Partner with children at home	1515 (12.02)	
Partner without children at home	6245 (49.56)	
Single parent	403 (3.20)	
Member of collective household	371 (2.94)	
Other	154 (1.22)	
Age of children living at parental home		
≤12	388 (3.08)	
>12	1752 (13.90)	
No children living at home	10 462 (83.02)	
Biomedical characteristics		
Type of chronic condition(s)		
Only physical	10 060 (79.83)	
Only mental	436 (3.46)	
Combination of both	2106 (16.71)	
Number of chronic conditions, mean (SD)	2.23 (0.93)	0
Prevalence of 28 chronic conditions		
Chronic alcohol abuse	163 (1.29)	
Endocardial conditions, valvular conditions	298 (2.36)	
Congenital cardiovascular anomaly	25 (0.20)	
HIV/AIDS	9 (0.07)	
Anxiety disorders	649 (5.15)	
Asthma	2142 (17.00)	
Stroke (including TIA)	986 (7.82)	
Chronic obstructive pulmonary disease (COPD)	2218 (17.60)	
Chronic back or neck disorder	2033 (16.13)	
Coronary heart diseases	1725 (13.69)	
Dementia including Alzheimer's	172 (1.36)	
Diabetes mellitus	4925 (39.08)	
Epilepsy	181 (1.44)	
Hearing disorders	679 (5.39)	
Visual disorders	1694 (13.44)	
Heart failure	659 (5.23)	
Heart arrhythmia	1446 (11.47)	
Cancer	2032 (16.12)	
Migraine	395 (3.13)	
Osteoporosis	737 (5.85)	
Burnout	452 (3.59)	
Osteoarthritis	2360 (18.73)	

(Continued)

Table 1. (Continued)

Patient characteristics		
	n (%)	Missing, n (%)
Personality disorders	120 (0.95)	
Rheumatoid arthritis	433 (3.44)	
Schizophrenia	53 (0.42)	
Mood disorders	1380 (10.95)	
Mental retardation	48 (0.38)	
Parkinson's disease	136 (1.08)	
Socioeconomic characteristics		
Housing situation		12 (0.10)
Owner-occupied	6777 (53.78)	
Rented <sup>b</sup>	5813 (46.13)	
Source of income		0
Paid work <sup>c</sup>	1974 (15.66)	
Social welfare or unemployment benefits	1838 (14.58)	
Pension benefits	8156 (64.72)	
Without income <sup>d</sup>	634 (5.03)	
Number of people in a household with an individual income		26 (0.21)
1	4594 (36.45)	
>1	7982 (63.34)	
Household dependence on social security payments, mean (SD)	11.63 (25.44)	346 (2.75)
Paid interest over debts, mean (SD)	48.89 (782.63)	20 (0.16)
Care utilization		
Pharmaceutical costs		16 (0.13)
≤€500	4773 (37.87)	
>€500 and ≤€1500	5122 (40.64)	
>€1500	2691 (21.35)	
GP care utilization before baseline, mean (SD)	29.97 (18.50)	0

<sup>a</sup> For continuous variables, mean (SD) is reported.

<sup>b</sup> Includes members of collective households.

<sup>c</sup> Includes employees, entrepreneurs, and managers.

<sup>d</sup> Includes students with and without individual income.

<https://doi.org/10.1371/journal.pone.0228103.t001>

Demographic characteristics were measured at baseline and included patients' sex, age (in years), household position (child living at home, single adult, partner with children at home, partner without children at home, single parent, member of a collective household, other), and age of children living at parental home (≤12 years, >12 years (i.e. the age that they generally leave elementary school), no children living at home). Biomedical characteristics were also measured at baseline and included patients' chronic disease diagnoses based on GP care use related to the chronic disease in the 1.5 years before baseline, type of chronic condition(s) (only physical, only mental, or a combination of both), and number of chronic conditions (1 to 28). All socioeconomic characteristics, except for source of income, were measured over the year 2014 and included patients' (household) housing situation (owner-occupied, rented), number of people in a household with an individual income (1, >1), household dependence on social security payments as proportion of gross household income (0% to 100%), and paid interest over debts (in euros, excluding mortgage or debts related to renovating personal property). Source of income (paid work, social welfare or unemployment benefits, pension benefits,

without income) was measured at baseline. Care utilization characteristics included GP care utilization on baseline (number of registred GP consultations) and patients' pharmaceutical costs ( $\leq\text{€}500$ ,  $>\text{€}500$  and  $\leq\text{€}1500$ ,  $>\text{€}1500$ ) which were measured over 2014.

## Data analysis

Data were validated and checked for outliers and missing values. We employed LCA, which is a sophisticated analysis technique to capture heterogeneity in the HNHC population's needs by the smallest number of unobserved homogeneous classes [37]. Furthermore, LCA is a person-oriented analysis technique [37] which aims to identify classes of individuals with similar patterns of, in the current study, (correlated) personal factors relevant to health care utilization. Initially, the LCA was run using all 41 patient characteristics (see Table 1) in order to explore the potential to identify clinically relevant subgroups. Furthermore, the analysis was conducted with a maximum likelihood estimator with robust standard errors (MLR). Missing values were handled by the default option in the Mplus software (version 8.1). To test whether the missing values were completely at random (MCAR), a MCAR Pearson-Chi Square and Likelihood Ratio Chi-Square test ( $P < .05$ ) was computed. Additionally, the number of random starts values was increased several times to prevent problems related to nonconvergence or local maxima.[38] By stepwise increasing the number of classes, starting with a 1-class model, and comparing various statistical indicators and clinical relevance, we decided on the final model. Statistical indicators for model fit included the Akaike Information Criterion (AIC), [39, 40] Bayesian Information Criterion (BIC), [41] bootstrapped likelihood ratio test (BLRT), [42] and entropy score. Lower values on AIC and BIC indicated better model fit; significant p-values on the BLRT showed dominance of the k class model, compared to the k-1 class model. The entropy score gave an indication of classification certainty, using a cutoff score of at least 0.8, indicating high classification certainty [38]. The BIC and BLRT were considered most important in deciding on the best model as these outperform other statistical indicators [43].

Besides statistical indicators, clinical relevance of the model was a key factor, as the model should support daily clinical practice [15]. Also, the size of the classes within the model was taken into account (also referred to as substantiality) [15]. A model with classes including at least 10% of HNHC population was considered substantial to counterbalance efforts to tailor interventions in daily practice. Although we aimed to maintain the largest variety of patient characteristics, the model was made more parsimonious after identifying a clinically relevant model. Thus, we removed any variables that did not contribute to the division in clinically relevant classes, significantly deteriorated the model fit, and/or were regarded as being of less added value based on internal clinical insight. Patients in each class of the final model were described in terms of the probability of having a given patient characteristic. In line with previous studies using LCA, probabilities of 70% to 100% were considered high, probabilities of 40% to 69% moderate, and probabilities of less than 40% low [44, 45]. The continuous variables were described by their estimated mean (SE). Furthermore, each class was described in their top five of chronic conditions at baseline and mean GP care utilization (i.e. mean number of weighted GP consultations) over the follow-up period.

## Results

### Baseline characteristics

A total of 63 general practices (48.5%) participated. The complete data set included individual-level data from 58 551 chronically ill patients, of whom 12 602 patients (21.5%) met the inclusion criteria for the study (i.e., were considered HNHC). Baseline characteristics of the HNHC

**Table 2. Statistical indicators and relative class sizes for models with increasing numbers of latent classes.**

	1-class model	2-class model	3-class model	4-class model	5-class model
Loglikelihood	-183,726.630	-172,407.886	-164,350.740	-159,286.403	-154,427.535
AIC <sup>a</sup>	367,493.259	344,893.772	328,817.480	318,726.806	309,047.071
BIC <sup>b</sup>	367,642.092	345,183.995	329,249.094	319,299.810	309,761.466
Entropy	n/a	0.981	0.974	0.973	0.977
BLRT <sup>c</sup>	n/a	$P < .001$	$P < .001$	$P < .001$	$P < .001$
Relative class size	n/a	86.62/13.38	64.51/23.58/11.90	39.30/25.51/23.31/11.87	38.18/25.31/18.48/9.21/8.82

<sup>a</sup>AIC refers to Akaike Information Criterion

<sup>b</sup>BIC refers to Bayesian Information Criterion

<sup>c</sup>BLRT refers to bootstrapped likelihood ratio test

<https://doi.org/10.1371/journal.pone.0228103.t002>

population, including number (%) of missing values per characteristic, are shown in [Table 1](#). Patients' mean (SD) GP care utilization over the follow-up period was 66.9 contacts (33.3).

### Latent class analysis

A 4-class model was considered statistically and clinically superior. The 4-class model had a low value on BIC, a significant BLRT ( $P < .001$ ), high entropy score (0.973), and each class was sufficiently substantial by including at least 10% of the HNHC population (see [Table 2](#)). Although the 5-class model was statistically superior to the 4-class model, it included two classes with less than 10% of the HNHC population and resulted in less relevant and distinct classes compared to the 4-class model. More specifically, a 5-class model largely maintained three of the four classes of the 4-class model and subdivided the fourth and smallest class of the 4-class model into two smaller classes which were relatively indistinct from each other.

[Table 3](#) shows the final model, which includes nine of the initially used 41 patient characteristics and the probabilities of having each patient characteristic, given class membership (see also [Fig 1](#)). This means that the following variables were excluded in the final LCA due to less statistical relevance: age of children living at parental home, number of chronic conditions, prevalence of 28 chronic conditions, paid interest over debts, GP care utilization on baseline. The MCAR Pearson-Chi Square and Likelihood Ratio Chi-Square test showed that values were missing completely at random ( $P < 0.001$ ). As the entropy score was high, we report the final class counts and proportions for the latent classes that are based on their most likely latent class membership. Class 1 ( $n = 4953$ ; 39.3%) had a mean (SE) age of 74.5 years (0.10), had a high probability (0.91) of having a partner but no children at home, and a high probability (0.98) of receiving pension benefits. Based on these dominant characteristics, class 1 was named 'older adults living with partner'. Class 2 ( $n = 3215$ ; 25.5%) had a mean (SE) age of 78.8 years (0.15), had a high probability (0.92) of being single, and a high probability (0.99) of receiving pension benefits. Based on these dominant characteristics, class 2 was named 'older adults living alone'. Class 3 ( $n = 2938$ ; 23.3%) had a mean (SE) age of 51.0 years (0.24) and had a high probability of having a partner with or without children at home (0.82). In terms of socioeconomic status, members of class 3 had a moderate probability (0.62) of having paid work. Based on these dominant characteristics, class 3 was named 'middle-aged, employed adults with family'. Class 4 ( $n = 1496$ ; 11.9%) had a mean (SE) age of 52.2 years (0.32). With regard to household position, members of class 4 had a low probability (0.34) of being single and a low probability (0.33) of having a partner but no children at home. In terms of socioeconomic status, members of class 4 had a high probability (0.84) of receiving social welfare or unemployment benefits. Based on these dominant characteristics, class 4 was named 'middle-

**Table 3. Probabilities of having the (categorical) patient characteristic, given class membership, for each class within the final 4-class model.**

Patient characteristics	Probability (SE)			
	Class 1 (n = 4953)	Class 2 (n = 3215)	Class 3 (n = 2938)	Class 4 (n = 1496)
Demographic characteristics				
Sex				
Male	0.481 (0.01)	0.202 (0.01)	0.290 (0.01)	0.404 (0.01)
Female	0.519 (0.01)	0.798 (0.01)	0.710 (0.01)	0.596 (0.01)
Age, mean (SE) <sup>a</sup>	<b>74.47 (0.10)</b>	<b>78.78 (0.15)</b>	<b>51.01 (0.24)</b>	<b>52.22 (0.32)</b>
Household position				
Child living at home	0.001 (0.00)	0.000 (0.00)	0.034 (0.00)	0.027 (0.00)
Single adult	0.010 (0.00)	<b>0.917 (0.08)</b>	0.092 (0.01)	<b>0.341 (0.01)</b>
Partner with children at home	0.034 (0.00)	0.000 (0.00)	<b>0.391 (0.01)</b>	0.144 (0.01)
Partner without children at home	<b>0.905 (0.01)</b>	0.000 (0.00)	<b>0.424 (0.01)</b>	<b>0.332 (0.01)</b>
Single parent	0.022 (0.00)	0.002 (0.00)	0.044 (0.00)	0.105 (0.01)
Member of a collective household	0.010 (0.00)	0.080 (0.07)	0.002 (0.00)	0.039 (0.01)
Other	0.019 (0.00)	0.001 (0.00)	0.013 (0.00)	0.012 (0.00)
Biomedical characteristics				
Type of chronic condition				
Only physical	0.891 (0.00)	0.863 (0.01)	0.664 (0.01)	0.610 (0.01)
Only mental	0.008 (0.00)	0.014 (0.00)	0.077 (0.01)	0.085 (0.01)
Combination of both	0.101 (0.00)	0.123 (0.01)	0.259 (0.01)	0.306 (0.01)
Socioeconomic characteristics				
Housing situation				
Owner-occupied	0.637 (0.01)	0.343 (0.01)	0.723 (0.01)	0.272 (0.01)
Rented	0.363 (0.01)	0.657 (0.01)	0.277 (0.01)	0.728 (0.01)
Source of income				
Paid work	0.018 (0.00)	0.007 (0.00)	<b>0.621 (0.01)</b>	0.047 (0.01)
Social welfare or unemployment benefits	0.002 (0.00)	0.002 (0.00)	0.195 (0.01)	<b>0.844 (0.01)</b>
Pension benefits	<b>0.981 (0.00)</b>	<b>0.990 (0.00)</b>	0.000 (0.00)	0.042 (0.01)
Without income	0.000 (0.00)	0.001 (0.00)	0.184 (0.01)	0.068 (0.01)
Number of people with an individual income in a household				
1	0.026 (0.00)	0.969 (0.01)	0.218 (0.01)	0.483 (0.01)
>1	0.974 (0.00)	0.031 (0.01)	0.782 (0.01)	0.517 (0.01)
Household dependence				
on social security payments, mean (SE) <sup>a</sup>	1.28 (0.09)	0.35 (0.06)	9.28 (0.33)	75.81 (0.64)
Care utilization				
Pharmaceutical costs				
≤€500	0.353 (0.01)	0.318 (0.01)	0.513 (0.01)	0.340 (0.01)
>€500 and ≤€1500	0.439 (0.01)	0.423 (0.01)	0.349 (0.01)	0.378 (0.01)
>€1500	0.208 (0.01)	0.259 (0.01)	0.138 (0.01)	0.282 (0.01)

<sup>a</sup> For continuous variables, mean (SE) is reported

<https://doi.org/10.1371/journal.pone.0228103.t003>

aged adults with social welfare dependency'. See also [S1 File](#) for a description of typical qualitative personas who characterize the four classes.

In terms of the top five chronic conditions per class at baseline (see [Fig 2](#)), diabetes mellitus was most common in each of the four classes, with prevalence ranging from 30.5% in class 3 to 43.4% in class 1. The second most prevalent condition differed between osteoarthritis in class 1 (21.7%) and 2 (23.8%), asthma in class 3 (25.3%), and mood disorders in class 4 (23.1%).



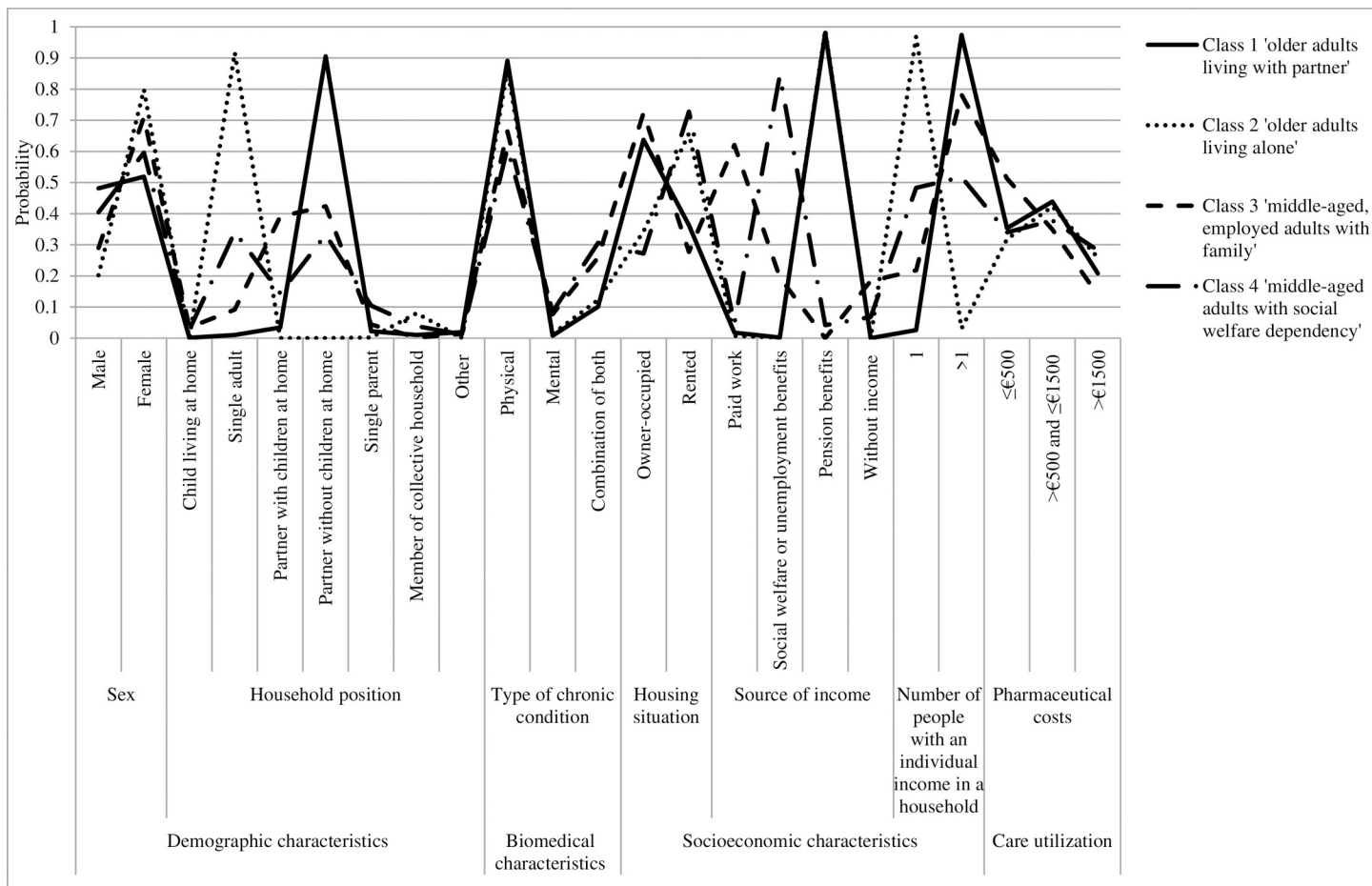


Fig 1. Probabilities of having the (categorical) patient characteristic, given class membership, for each class within the final 4-class model.

<https://doi.org/10.1371/journal.pone.0228103.g001>

With regard to GP care utilization of the classes over the follow-up period (see Fig 3), class 2 showed the highest mean care utilization. Both classes with the older adults showed the largest mean (SD) increase in care utilization over time—from 9.8 (6.9) in the first to 11.7 (8.7) in the sixth half year and from 11.5 (8.3) in the first to 14.0 (10.5) in the sixth half year—while the classes with the middle-aged adults were more stable over time—from 10.1 (7.1) in the first to 10.7 (8.2) in the sixth half year and from 11.3 (8.0) in the first to 12.1 (9.5) in the sixth half year.

### Discussion

The present study suggests that the HNHC population in primary care is a heterogeneous population, which can be divided into four subgroups with distinct patterns of particularly demographic and socioeconomic characteristics. Main differences between the subgroups were found in demographic and socioeconomic factors (i.e. age, household position, and source of income). In terms of chronic conditions, the subgroups with older adults most frequently suffered from physical and age-related conditions (e.g. osteoarthritis, cancer), while the middle-aged subgroups most frequently had conditions more typically found in relatively younger people (i.e., asthma and mood disorders). Furthermore, while the subgroups with older adults showed an increase in mean care utilization over time, the middle-aged subgroups showed a

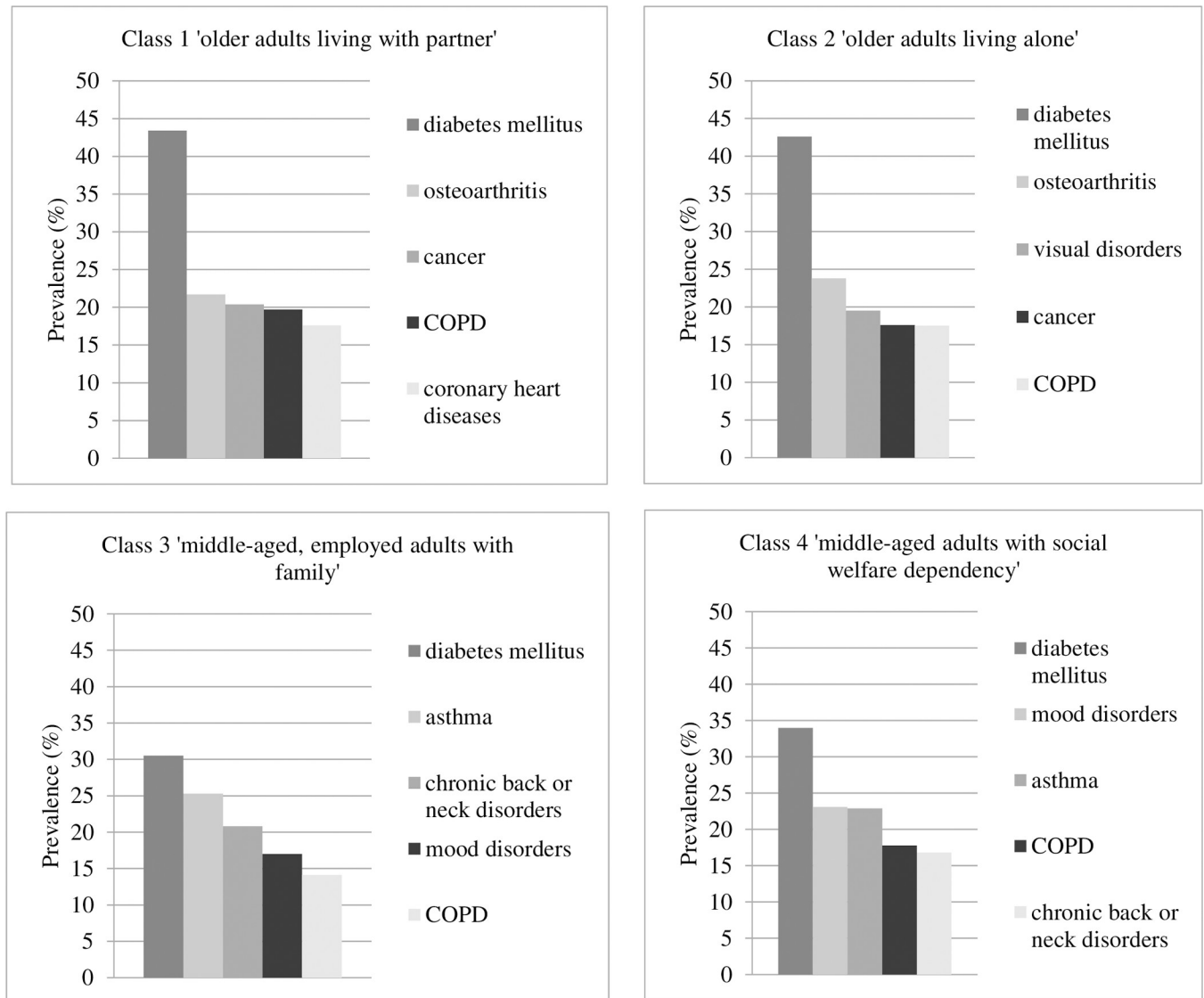
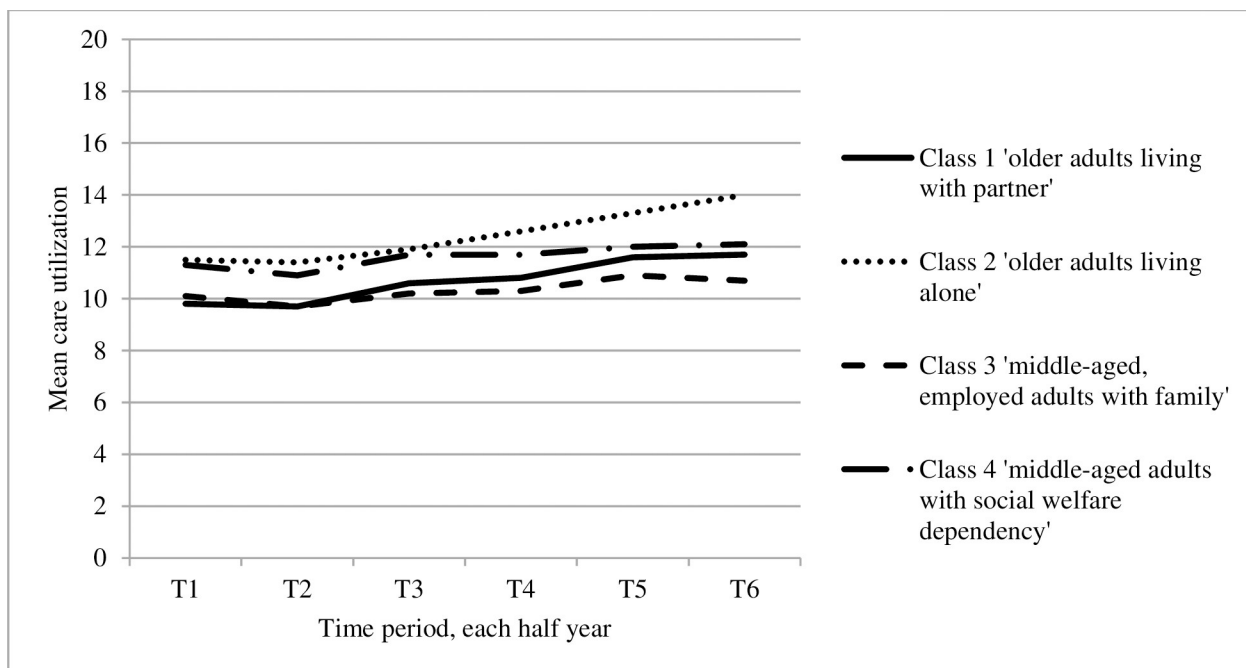


Fig 2. Top five of chronic conditions (%) per class within the final 4-class model.

<https://doi.org/10.1371/journal.pone.0228103.g002>

more stable pattern over time. In addition, class 2 ('older adults living alone') showed the highest mean care utilization over time. This finding corresponds with a study of Dreyer, Steventon [36] who showed that living alone is associated with higher care utilization in older adults.

The current study indicates that the sex distribution within the HNHC population, as well as in three of the four identified subgroups, is unbalanced: more than 64% of the HNHC population is female. In the current person-oriented analysis, unlike in a variable-oriented analysis, there is no assessment of relations between variables including corrections for confounders. Rather, the current analysis has focused on identifying subgroups based on patterns of variables within individual patients. One possible explanation for the unbalanced population in terms of sex is that women typically get older and, as a result, are overrepresented among the older aged HNHC patients compared to men. In addition, scientific studies have found that women have significantly higher consultation rates compared to men, but particularly during working years [46, 47].



**Fig 3. GP care utilization measured over the follow-up period for each class within the final 4-class model.**

<https://doi.org/10.1371/journal.pone.0228103.g003>

Our findings show that the HNHC population is a demographically and socioeconomically diverse population and includes not only older adults but also many middle-aged people. To date, studies have predominantly focused on (biomedical) segmentation in populations of older adults: an example is the recent Embrace study, [48] which identified three risk profiles for older adults. In line with the demographic heterogeneity found in our study, a study by Wammes, Tanke [12] found that many high-cost patients (in the Dutch curative health system) are not older than 65 years of age. Supporting our approach, the authors [12] emphasized the need for studying the general population with extensive data and targeting interventions toward high-cost patients of various ages. Furthermore, our findings suggest that middle-aged HNHC patients are generally characterized by more socioeconomic vulnerability (e.g., dependence on social welfare) and a higher prevalence of mental conditions (e.g., mood disorders) than are older HNHC patients. These findings add to an increasing awareness about the importance of social and context-related determinants of health [25, 49, 50]. First, Shadmi [51] suggests broadening the understanding and measurement of multimorbidity by including a large variety of health and health-related aspects (e.g., social, cultural, and economic background of populations) that correlate with multimorbidity. In addition, corresponding to our finding that current segmentation often lacks inclusion of relevant demographic and socioeconomic characteristics, the study by Chin-Yee, Subramanian [52] and Khoury, Iademarco [53] also argued that adding environmental and social characteristics (a rather “population perspective”) to the genetic profiling in precision medicine can be of added value to public health.

With the growing recognition of the effectiveness of segmentation for patient-centered interventions, [54] the segmentation conducted in the present study can guide clinical practice toward more integrated and person-centered care. By gathering insight into demographic characteristics other than age and gender (e.g., household position) as well as the socioeconomic context of patients (e.g., main source of income), clinical practice in primary care can be attuned to a more holistic view of patients. This view can suggest potentially relevant goals,

interventions, and professionals (within primary care and in cooperation with other disciplines), which can be further discussed in a shared decision-making process with the patient. Such an approach can be inspired by the 'Bridges to Health' model, [55] which aims to systematically connect priority concerns, major components of health care, and goals for health care within identified population segments [56]. Thus, while older adults living alone might benefit from increased social support, middle-aged adults with social welfare dependency might rather benefit from financial and mental support. As such, this segmentation approach can serve as a starting point for more biopsychosocial attention and can inform the discussion of tailored interventions with the patient [56]. However, the individual consultation is still key to assess personal needs and preferences with a patient during a consultation, and agree on an individual treatment course.

Further research, in particular qualitative inquiry, is necessary to identify the most important concerns and components of health care per HNHC subgroup. In addition, the current study has focused on HNHC patients in primary care, which is widely considered the most suitable medical home for chronically ill patients [57]. Although as a result, our findings are mainly useful for improvement of primary care management, there is some evidence that patients with a disproportionately high use of primary care resources also account for significantly high(er) costs in specialist care [58, 59]. For policy making, the subgroups can also help to give insight into the distribution of the patient population over the identified subgroups within certain geographical areas and help to efficiently target resources. In more urban areas, for example, the middle-aged subgroups might be larger than in rural areas.

One of the most important strengths of the current study is the relatively large set of individual-level patient data, with a variety of patient characteristics. A second strength is the use of the model-based analysis technique LCA, which offers a large set of statistical indicators to decide on the best-fitting model and ways to cope with issues of local maxima and nonconvergence [38]. The study also has some limitations. First, individual level data of the non-participating practices were not available in this study. This hampered a direct comparison of participating practices ( $n = 63$ ; 48.5% of care group) with non-participating practices in order to assess representativeness of the sample. However, particular patient characteristics (i.e. sex, age, household position, and source of income) of the sample were compared to the patient characteristics of the general population in the northern region of the Netherlands that is covered by the primary care group. This comparison showed that the sample is largely similar in patient characteristics to the general population. For example, 50.8% of the sample is female; 50.5% of the general population is female, 20.1% of the sample receives pension benefits; 22.1% of the general population receives pension benefits. Second, EHRs typically include incomplete registrations and may have limited data quality. Nevertheless, the quality of registrations was checked and validated, and the (categorical) missing values were found to be MCAR. Third, the data set included patients who can be considered dependent, as they belonged to the same household. A sensitivity analysis with only completely independent observations showed the same division among classes, implying a negligible effect of the dependent observations on the identification of subgroups. Fourth, only patients with a full EHR registration over the research period were included. This has excluded specific types of patients, such as patients who died before the end of the follow-up period. It is possible that the excluded patients would have been identified as a separate 'near end of life' HNHC subgroup, as identified by some previous population segmentation studies as well [24, 55, 60]. Nevertheless, specific payments arrangements are already in place in Dutch primary care for this patient population who is near the end of life and needs (expensive) palliative care. Fifth, generalizability of the subgroups may be limited, as the data set was retrieved from a specific Dutch region with limited ethnic/cultural diversity and a relatively aged population, compared

to the Dutch average. In future research, the generalizability of the subgroups needs to be determined.

## Conclusions

Despite the heterogeneity of the HNHC population, distinct subgroups with relatively homogeneous patterns of particularly demographic and socioeconomic characteristics can be identified. This study adds to the increasing awareness of the demographic and socioeconomic heterogeneity of the HNHC population, in addition to biomedical diversity. To accommodate person-centered, integrated care delivery, the identified classes need to be connected to tailored care (i.e. concerns, components, goals). This connection can be inspired by the proposed strategies within the 'Bridges to Health' model [55].

## Supporting information

**S1 File. Description of typical qualitative personas who characterize the four identified classes.**

(DOCX)

## Acknowledgments

We thank primary care group 'HZD' and connected general practices for facilitating this study by providing individual-level patient data. Furthermore, we thank all organizations that provided support with either provision, preparation, pseudonymization of data and/or connection of EHR data to other datasets (i.e. 'Calculus', 'ZorgTTP', Statistics Netherlands ('CBS'), 'Vektis'). We also thank researchers Dorijn Hertroijs, PhD; Niels Janssen, MSc, and Sebastian Köhler, PhD, for providing additional assistance in data analysis.

## Author Contributions

**Conceptualization:** Rowan G. M. Smeets, Arianne M. J. Elissen, Mariëlle E. A. L. Kroese, Niels Hameleers, Dirk Ruwaard.

**Data curation:** Rowan G. M. Smeets, Niels Hameleers.

**Formal analysis:** Rowan G. M. Smeets, Niels Hameleers.

**Investigation:** Rowan G. M. Smeets, Niels Hameleers.

**Methodology:** Rowan G. M. Smeets, Arianne M. J. Elissen, Niels Hameleers.

**Supervision:** Arianne M. J. Elissen, Mariëlle E. A. L. Kroese, Dirk Ruwaard.

**Writing – original draft:** Rowan G. M. Smeets, Arianne M. J. Elissen.

**Writing – review & editing:** Rowan G. M. Smeets, Arianne M. J. Elissen, Mariëlle E. A. L. Kroese, Niels Hameleers, Dirk Ruwaard.

## References

1. Salisbury C, Johnson L, Purdy S, Valderas JM, Montgomery AA. Epidemiology and impact of multimorbidity in primary care: a retrospective cohort study. *Br J Gen Pract.* 2011; 61(582):e12–e21. <https://doi.org/10.3399/bjgp11X548929> PMID: 21401985
2. van Weel C, Schellevis FG. Comorbidity and guidelines: conflicting interests. *The Lancet.* 2006; 367(9510):550–551. [https://doi.org/10.1016/S0140-6736\(06\)68198-1](https://doi.org/10.1016/S0140-6736(06)68198-1)
3. Søndergaard E, Willadsen TG, Guassora AD, Vestergaard M, Tomasdottir MO, Borgquist L, et al. Problems and challenges in relation to the treatment of patients with multimorbidity: general practitioners'

- views and attitudes. *Scand J Prim Health Care*. 2015; 33(2):121–126. <https://doi.org/10.3109/02813432.2015.1041828> PMID: 26158584
4. Fisher ES, Shortell SM. Accountable care organizations: accountable for what, to whom, and how. *JAMA*. 2010; 304(15):1715–1716. <https://doi.org/10.1001/jama.2010.1513> PMID: 20959584
  5. Struijs JN, Baan CA. Integrating care through bundled payments—lessons from the Netherlands. *N Engl J Med*. 2011; 364(11):990–991. <https://doi.org/10.1056/NEJMp1011849> PMID: 21410368
  6. McClellan M, McKethan AN, Lewis JL, Roski J, Fisher ES. A national strategy to put accountable care into practice. *Health Aff (Millwood)*. 2010; 29(5):982–990. <https://doi.org/10.1377/hlthaff.2010.0194> PMID: 20439895
  7. Fisher ES, McClellan MB, Bertko J, Lieberman SM, Lee JJ, Lewis JL, et al. Fostering accountable health care: moving forward in Medicare. *Health Aff (Millwood)*. 2009; 28(2):w219–w231. <https://doi.org/10.1377/hlthaff.28.2.w219> PMID: 19174383
  8. Ganguli I, Ferris TG. Accountable care at the frontlines of a health system: bridging aspiration and reality. *JAMA*. 2018; 319(7):655–656. <https://doi.org/10.1001/jama.2017.18995> PMID: 29228143
  9. de Bakker DH, Struijs JN, Baan CA, Raams J, de Wildt JE, Vrijhoef HJM, et al. Early results from adoption of bundled payment for diabetes care in the Netherlands show improvement in care coordination. *Health Aff (Millwood)*. 2012; 31(2):426–433. <https://doi.org/10.1377/hlthaff.2011.0912> PMID: 22323174
  10. Blumenthal D, Chernof B, Fulmer T, Lumpkin J, Selberg J. Caring for high-need, high-cost patients—an urgent priority. *N Engl J Med*. 2016; 375(10):909–911. <https://doi.org/10.1056/NEJMp1608511> PMID: 27602661
  11. Hayes SL, Salzberg CA, McCarthy D, Radley DC, Abrams MK, Shah T, et al. High-need, high-cost patients: who are they and how do they use health care? A population-based comparison of demographics, health care use, expenditures Issue Brief (Commonw Fund). 2016; 26:1–14.
  12. Wammes JJG, Tanke M, Jonkers W, Westert GP, van der Wees P, Jeurissen PPT. Characteristics and healthcare utilisation patterns of high-cost beneficiaries in the Netherlands: a cross-sectional claims database study. *BMJ Open*. 2017; 7(11):e017775. <https://doi.org/10.1136/bmjopen-2017-017775> PMID: 29133323
  13. Cooil B, Aksoy L, Keiningham TL. Approaches to customer segmentation. *J Relationsh Mark*. 2008; 6(3–4):9–39. [https://doi.org/10.1300/J366v06n03\\_02](https://doi.org/10.1300/J366v06n03_02).
  14. Vuik SI, Mayer EK, Darzi A. Patient segmentation analysis offers significant benefits for integrated care and support. *Health Aff (Millwood)*. 2016; 35(5):769–775. <https://doi.org/10.1377/hlthaff.2015.1311> PMID: 27140981
  15. Yan S, Kwan YH, Tan CS, Thumboo J, Low LL. A systematic review of the clinical application of data-driven population segmentation analysis. *BMC Med Res Methodol*. 2018; 18(1):121. <https://doi.org/10.1186/s12874-018-0584-9> PMID: 30390641
  16. Struijs JN, Drewes HW, Heijink R, Baan CA. How to evaluate population management? Transforming the Care Continuum Alliance population health guide toward a broadly applicable analytical framework. *Health Policy*. 2015; 119(4):522–529. <https://doi.org/10.1016/j.healthpol.2014.12.003> PMID: 25516015
  17. Kindig DA. Understanding population health terminology. *Milbank Q*. 2007; 85(1):139–161. <https://doi.org/10.1111/j.1468-0009.2007.00479.x> PMID: 17319809
  18. Krewski D, Hogan V, Turner MC, Zeman PL, McDowell I, Edwards N, et al. An integrated framework for risk management and population health. *Hum Ecol Risk Assess*. 2007; 13(6):1288–1312. <https://doi.org/10.1080/10807030701655798>
  19. Bodenheimer T, Sinsky C. From Triple to Quadruple Aim: care of the patient requires care of the provider. *Ann Fam Med*. 2014; 12(6):573–576. <https://doi.org/10.1370/afm.1713> PMID: 25384822
  20. Wammes JJG, van der Wees PJ, Tanke MAC, Westert GP, Jeurissen PPT. Systematic review of high-cost patients' characteristics and healthcare utilisation. *BMJ Open*. 2018; 8(9):e023113. <https://doi.org/10.1136/bmjopen-2018-023113> PMID: 30196269
  21. Zulman DM, Chee CP, Wagner TH, Yoon J, Cohen DM, Holmes TH, et al. Multimorbidity and health-care utilisation among high-cost patients in the US Veterans Affairs Health Care System. *BMJ Open*. 2015; 5(4):e007771. <https://doi.org/10.1136/bmjopen-2015-007771> PMID: 25882486
  22. Lafortune L, Béland F, Bergman H, Ankri J. Health state profiles and service utilization in community-living elderly. *Med Care*. 2009; 47(3):286–294. <https://doi.org/10.1097/MLR.0b013e3181894293> PMID: 19165121
  23. Liu LF, Tian WH, Yao HP. Utilization of health care services by elderly people with National Health Insurance in Taiwan: the heterogeneous health profile approach. *Health Policy*. 2012; 108(2–3):246–255. <https://doi.org/10.1016/j.healthpol.2012.08.022> PMID: 23022286
  24. Zhou YY, Wong W, Li H. Improving care for older adults: a model to segment the senior population. *Perm J*. 2014; 18(3):18–21. <https://doi.org/10.7812/TPP/14-005> PMID: 24937151

25. Rinehart DJ, Oronce C, Durfee MJ, Ranby KW, Batal H, Hanratty R, et al. Identifying subgroups of adult super-utilizers in an urban safety-net system using latent class analysis: implications for clinical practice. *Med Care*. 2018; 56(1):e1–e9. <https://doi.org/10.1097/MLR.0000000000000628> PMID: 27632768
26. Xu J, Williams-Livingston A, Gaglioti A, McAllister C, Rust G. A practical risk stratification approach for implementing a primary care chronic disease management program in an underserved community. *J Health Care Poor Underserved*. 2018; 29(1):202–213. <https://doi.org/10.1353/hpu.2018.0014> PMID: 29503295
27. Buja A, Claus M, Perin L, Rivera M, Corti MC, Avossa F, et al. Multimorbidity patterns in high-need, high-cost elderly patients. *PLoS ONE*. 2018; 13(12):e0208875. <https://doi.org/10.1371/journal.pone.0208875> PMID: 30557384
28. Statistics Netherlands (CBS). Organisation n.d. [cited 2019 January 23]. Available from: <https://www.cbs.nl/en-gb/about-us/organisation>.
29. Vektis. Over Vektis n.d. [cited 2019 January 28]. Available from: <https://www.vektis.nl/over-vektis>.
30. O'Halloran J, Miller GC, Britt H. Defining chronic conditions for primary care with ICPC-2. *Family Pract*. 2004; 21(4):381–386. <https://doi.org/10.1093/fampra/cmh407> PMID: 15249526
31. van Oostrom SH, Picavet HSJ, van Gelder BM, Lemmens LC, Hoeymans N, van Dijk CE, et al. Multimorbidity and comorbidity in the Dutch population – data from general practices. *BMC Public Health*. 2012; 12(1):715. <https://doi.org/10.1186/1471-2458-12-715> PMID: 22935268
32. Joynt KE, Gawande AA, Orav EJ, Jha AK. Contribution of preventable acute care spending to total spending for high-cost Medicare patients. *JAMA*. 2013; 309(24):2572–2578. <https://doi.org/10.1001/jama.2013.7103> PMID: 23797716
33. Wodchis WP, Austin PC, Henry DA. A 3-year study of high-cost users of health care. *Can Med Assoc J*. 2016; 188(3):182–188. <https://doi.org/10.1503/cmaj.150064> PMID: 26755672
34. Flinterman LE, de Bakker DH, Verheij RA. Differentiatie inschrijftarief huisartsenzorg. Deel 1: evaluatie van de huidige differentiatie van het inschrijftarief. Utrecht: NIVEL; 2015.
35. Nederlands Zorgautoriteit (NZa). Prestatie- en tariefbeschikking huisartsenzorg en multidisciplinaire zorg 2018. 2018.
36. Dreyer K, Steventon A, Fisher R, Deeny SR. The association between living alone and health care utilisation in older adults: a retrospective cohort study of electronic health records from a London general practice. *BMC Geriatr*. 2018; 18(1):269. <https://doi.org/10.1186/s12877-018-0939-4> PMID: 30514225
37. Collins LM, Lanza ST. Latent class and latent transition analysis: with applications in the social, behavioral, and health sciences. Hoboken, NJ: Wiley; 2010.
38. Jung T, Wickrama KAS. An introduction to latent class growth analysis and growth mixture modeling. *Soc Personal Psychol Compass*. 2008; 2(1):302–317. <https://doi.org/10.1111/j.1751-9004.2007.00054.x>
39. Bozdogan H. Model selection and Akaike's Information Criterion (AIC): the general theory and its analytical extensions. *Psychometrika*. 1987; 52(3):345–370. <https://doi.org/10.1007/bf02294361>
40. Akaike H. Statistical predictor identification. *Ann Inst Stat Math*. 1970; 22(1):203–217. <https://doi.org/10.1007/bf02506337>
41. Schwarz G. Estimating the dimension of a model. *Ann Stat*. 1978; 6(2):461–464. <https://doi.org/10.1214/aos/1176344136>
42. Asparouhov T, Muthén B. Using Mplus TECH11 and TECH14 to test the number of latent classes. Mplus Web Notes [Internet]. 2012; 14:[22 p.]. Available from: <http://statmodel.com/examples/webnotes/webnote14.pdf>.
43. Nylund KL, Asparouhov T, Muthén BO. Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. *Struct Equ Modeling*. 2007; 14(4):535–569. <https://doi.org/10.1080/10705510701575396>
44. Beeber AS, Thorpe JM, Clipp EC. Community-based service use by elders with dementia and their caregivers: a latent class analysis. *Nurs Res*. 2008; 57(5):312–321. <https://doi.org/10.1097/01.NNR.0000313500.07475.eb> PMID: 18794715
45. Ryan CJ, DeVon HA, Horne R, King KB, Milner K, Moser DK, et al. Symptom clusters in acute myocardial infarction: a secondary data analysis. *Nurs Res*. 2007; 56(2):72–81. <https://doi.org/10.1097/01.NNR.0000263968.01254.d6> PMID: 17356437
46. Nabalamba A, Millar WJ. Going to the doctor. *Health Rep*. 2007; 18(1):23–35. PMID: 17441441
47. Wang Y, Hunt K, Nazareth I, Freemantle N, Petersen I. Do men consult less than women? An analysis of routinely collected UK general practice data. *BMJ Open*. 2013; 3(8):e003320. <https://doi.org/10.1136/bmjopen-2013-003320> PMID: 23959757

48. Spoorenberg SLW, Wynia K, Uittenbroek RJ, Kremer HPH, Reijneveld SA. Effects of a population-based, person-centred and integrated care service on health, wellbeing and self-management of community-living older adults: a randomised controlled trial on Embrace. *PLoS ONE*. 2018; 13(1): e0190751. <https://doi.org/10.1371/journal.pone.0190751> PMID: 29351295
49. Adler NE, Cutler DM, Fielding JE, Galea S, Glymour M, Koh HK, et al. Addressing social determinants of health and health disparities. *Vital directions for health and health care: an initiative of the National Academy of Medicine* Washington DC. 2016.
50. Braveman P, Gottlieb L. The social determinants of health: it's time to consider the causes of the causes. *Public Health Rep*. 2014; 129(1\_suppl2):19–31. <https://doi.org/10.1177/003335491412900105>
51. Shadmi E. Disparities in multiple chronic conditions within populations. *J Comorb*. 2013; 3(2):45–50. <https://doi.org/10.15256/joc.2013.3.24> PMID: 29090147
52. Chin-Yee B, Subramanian SV, Verma AA, Laupacis A, Razak F. Emerging trends in clinical research with implications for population health and health policy. *Milbank Q*. 2018; 96(2):369–401. <https://doi.org/10.1111/1468-0009.12328> PMID: 29870114
53. Khoury MJ, Iademarco MF, Riley WT. Precision public health for the era of precision medicine. *Am J Prev Med*. 2016; 50(3):398–401. <https://doi.org/10.1016/j.amepre.2015.08.031> PMID: 26547538
54. Soto-Gordoa M, de Manuel E, Fullaondo A, Merino M, Arrospide A, Igartua JI, et al. Impact of stratification on the effectiveness of a comprehensive patient-centered strategy for multimorbid patients. *Health Serv Res*. 2019; 54(2):466–473. <https://doi.org/10.1111/1475-6773.13094> PMID: 30467846
55. Lynn J, Straube BM, Bell KM, Jencks SF, Kambic RT. Using population segmentation to provide better health care for all: the “Bridges to Health” model. *Milbank Q*. 2007; 85(2):185–208. <https://doi.org/10.1111/j.1468-0009.2007.00483.x> PMID: 17517112
56. National Academy of Medicine. *Effective care for high-need patients: opportunities for improving outcomes, value, and health*. 2017.
57. Reynolds R, Dennis S, Hasan I, Slewa J, Chen W, Tian D, et al. A systematic review of chronic disease management interventions in primary care. *BMC Fam Pract*. 2018; 19:1–13. <https://doi.org/10.1186/s12875-017-0677-2>
58. Smits FT, Brouwer HJ, Zwinderman AH, Mohrs J, Smeets HM, Bosmans JE, et al. Morbidity and doctor characteristics only partly explain the substantial healthcare expenditures of frequent attenders: a record linkage study between patient data and reimbursements data. *BMC Fam Pract*. 2013; 14(1):1–8. <https://doi.org/10.1186/1471-2296-14-138> PMID: 24044374
59. Vedsted P, Fink P, Sørensen HT, Olesen F. Physical, mental and social factors associated with frequent attendance in Danish general practice. A population-based cross-sectional study. *Soc Sci Med*. 2004; 59(4):813–823. <https://doi.org/10.1016/j.socscimed.2003.11.027> PMID: 15177837
60. Low LL, Kwan YH, Liu N, Jing X, Low ECT, Thumboo J. Evaluation of a practical expert defined approach to patient population segmentation: a case study in Singapore. *BMC Health Serv Res*. 2017; 17(1):771. <https://doi.org/10.1186/s12913-017-2736-8> PMID: 29169359