Published in final edited form as:

Med Care. 2021 May 01; 59(5): 410–417. doi:10.1097/MLR.0000000000001526.

Subgroups of High-Risk VA Patients Based on Social Determinants of Health Predict Risk of Future Hospitalization

Dan V. Blalock, PhD[1,2], **Matthew L. Maciejewski, PhD**[1,3,4], **Donna M. Zulman, MD**[5,6], **Valerie A. Smith, DrPH**[1,3,4], **Janet Grubber, MSPH**[1], **Ann-Marie Rosland, MD MS**[7,8], **Hollis J. Weidenbacher, PhD**[1], **Liberty Greene, MS, Med**[5,6], **Leah L. Zullig, PhD**[1,3], **Heather E. Whitson, MD**[9,10,11], **Susan N. Hastings, MD**[1,3,9,10,11], **Anna Hung, PharmD, PhD**[1,3,12]

[1]Durham Center of Innovation to Accelerate Discovery and Practice Transformation (ADAPT), Durham Veterans Affairs Health Care System, Durham, NC

[2]Department of Psychiatry and Behavioral Sciences, Duke University, Durham, NC

[3]Department of Population Health Sciences, Duke University, Durham NC

[4]Division of General Internal Medicine, Department of Medicine, Duke University, Durham, NC

[5]Center for Innovation to Implementation, VA Palo Alto Health Care System, Menlo Park, CA

[6]Division of Primary Care and Population Health, Stanford University School of Medicine, Stanford CA

[7]VA Pittsburgh Center for Health Equity Research and Promotion, Pittsburgh PA

[8]Department of Medicine, University of Pittsburgh School of Medicine, Pittsburgh PA

[9]Geriatric Research, Education, and Clinical Center, Durham Veterans Affairs Health Care System, Durham, NC

[10]Center for the Study of Human Aging and Development, Duke University, Durham, NC

[11]Department of Medicine, Duke University School of Medicine, Durham, NC

[12]Duke Clinical Research Institute, Duke University, Durham, NC

Abstract

Objective.—Population segmentation has been recognized as a foundational step to help tailor interventions. Prior studies have predominantly identified subgroups based on diagnoses. In this study, we identify clinically coherent subgroups using social determinants of health (SDH) measures collected from Veterans at high risk of hospitalization or death.

Study Design and Setting.—SDH measures were obtained for 4,684 Veterans at high risk of hospitalization via mail survey. Eleven self-report measures known to impact hospitalization and amenable to intervention were chosen *a priori* by the study team to identify subgroups via latent

Corresponding author Dan V. Blalock, PhD, Durham Center of Innovation to Accelerate Discovery and Practice Transformation (ADAPT), Center for Health Services Research in Primary Care, Durham Veterans Affairs Health Care System, 411 West Chapel Hill Street, Suite 600, Durham, NC 27705, daniel.blalock@duke.edu.

CONFLICT OF INTEREST: All authors declare no conflicts of interest regarding this work.

class analysis (LCA). Associations between subgroups and Demographic and comorbidity characteristics were calculated via multinomial logistic regression. Odds of 180-day hospitalization were compared across subgroups via logistic regression.

Results.—Five subgroups of high-risk patients emerged - those with: minimal SDH vulnerabilities (8% hospitalized), poor/fair health with few SDH vulnerabilities (12% hospitalized), social isolation (10% hospitalized), multiple SDH vulnerabilities (12% hospitalized), and multiple SDH vulnerabilities without food or medication insecurity (10% hospitalized). In logistic regression, the ‘multiple SDH vulnerabilities’ subgroup had greater odds of 180-day hospitalization than did the “minimal SDH vulnerabilities’ reference subgroup (OR 1.53, 95% C.I: 1.09–2.14).

Conclusion.—Self-reported SDH measures can identify meaningful subgroups that may be used to offer tailored interventions to reduce their risk of hospitalization and other adverse events.

Keywords

high risk; social determinants; subgroup; segmentation; Veteran; hospitalization

Introduction

High-risk patients (i.e., those with a high likelihood of intensive services utilization, such as hospitalization, or severe clinical outcomes, such as death) incur a disproportionate share of health care expenditures, typically due to multimorbidity.[1,2] Improving care for these patients could potentially improve their health outcomes and prevent costly utilization, yielding population-level benefits.

A fundamental challenge in improving outcomes of high-risk patients is their heterogeneity of diagnosed conditions.[2–5] The multimorbidity contributing to their high-risk nature is comprised of diverse combinations of morbidities that lead to diverse outcomes and require diverse interventions.[6,7] Thus, it is important to identify clinically meaningful subgroups within the subset of the population that is high-risk in order to efficiently tailor or target interventions to these more homogeneous subgroups.[8,9] The clinical utility of subgroups and subsequent targeted interventions may shift depending on outcomes of interest, and utilization trends (e.g., hospitalizations) are an initial key outcome given the severity and marked increase of these outcomes observed in high-risk patients.[10]

Just as there are clinically actionable subgroups based on combinations of diagnosed chronic conditions, there may be meaningful subgroups of high-risk patients based on clusters of social determinants of health (SDH) measures. SDH measures broadly are any non-medical factors that may impact health outcomes. Few SDH measures are currently routinely collected in electronic health records (EHR),[11] but the Centers for Medicare and Medicaid have recently expanded coverage of supplemental benefits that may soon include a range of social programming for which SDH measurement would be highly beneficial.[12] In line with these expanding benefits, the US Preventative Task Force has begun developing expanded recommendations concerning SDH.[13,14]

One recent study in particular segmented high-risk patients in a regional healthcare system into subgroups defined by an increasing number of SDH measures.[15] This approach of segmenting high-risk patients could generate valuable insights for large national healthcare systems, like Veterans Affairs (VA), that offer clinical and social services for patients who may be more likely to have certain SDH vulnerabilities. In recent years, the VA has invested heavily in testing interventions for Veterans who are at high-risk for hospitalization;[16,17] identifying patients who are at particularly high risk due to social needs could improve patient selection and service priorities for these programs.

This study addressed two primary aims: 1) identify meaningful subgroups of high-risk VA patients based on patient-reported SDH, and 2) validate the utility of these subgroup classifications by examining differences in 180-day hospitalization rates across subgroups. Our overall objective was to identify discrete and clinically meaningful subgroups of high-risk Veterans that could help VA better tailor clinical and social services to distinct needs of these populations.[18]

Methods

Study Design and Population

In partnership with VA's Office of Primary Care, and due to the paucity of relevant data collected in standard care, a mail survey was fielded in 2018 to a nationally-representative stratified random sample of 10,000 Veterans who had at least one VA outpatient visit between 3/20/2017 and 3/18/2018, and were considered "high risk" (defined as 1-year risk of hospitalization or death in the ≥ 75^{th} percentile of VA's Care Assessment Need (CAN) score on 3/16/2018.[19] CAN score estimates probability of hospitalization or death within one year and is calculated based on demographics, medical conditions, vital signs, prior year Veteran Health Administration (VHA) health services utilization, medications dispensed, and laboratory results. Using the Dillman method,[20] the survey was sent with a cover letter, $2 bill incentive, prepaid return envelope, and 1–800 telephone number or return postcard to opt-out. The cover letter described the purpose of the survey, how Veterans were identified, and a statement that return of the mailed survey constituted informed consent. Veterans who did not opt out and who did not respond within a 6-week period were mailed a second survey with a prepaid envelope and cover letter.

Measure Selection

Survey measures were informed by the Cycle of Complexity Model,[21] which posits that patient complexity is a multi-factorial construct comprised of workload, acute shocks and medical events, capacity/resilience, and access/utilization. The domains in this theoretical framework provided a structured way to identify a preliminary list of measures by domain instead of choosing a set of measures without an underlying logic. The preliminary list was reduced to the one-third that the study team prioritized because they were of greatest interest, brief, and (when possible) validated.

While the initial paper from these data examined all 22 measures individually for prediction of 90-day and 180-day hospitalization,[22] the current investigation is limited to an *a priori* subset of 11 measures for creation of patient-level subgroups (described more below). These measures were chosen before data collection by five study team members representing multiple research and clinical perspectives based on two criteria: 1) variation among high-risk patients in the degree to which they endorse these measures from team members' experience and 2) anticipated relationship with 90-day hospitalization (the primary outcome for the initial manuscript).[22]

Reviewers independently gave each of the 22 survey measures a score of 0–2 based on whether they met neither, one or both criteria. Scores were summed across all members, and 12 measures met the pre-specified threshold of receiving a score of 5 or greater (life stressors; physical functioning; symptom burden; smoking status; loneliness; sleep; chaotic lifestyle; material needs; health status; social support; patient activation; and depression).

To determine if there were overlapping constructs, the five study team members were then asked to independently sort cards with each of the 12 highest scoring constructs to create groups of related constructs. After discussion, reviewers reached consensus on three general categories: 1) primarily medical challenges (e.g., physical functioning, symptom burden, smoking status, health status), 2) insufficient social support and/or loneliness (e.g., loneliness, social support); and 3) personal challenges (e.g., life stressors, sleep, chaotic lifestyle, material needs, patient activation, depression). Study team members then decided to drop two constructs (symptom burden and chaotic lifestyle) due to the high perceived correlation with other constructs, which can be problematic for maximizing the utility of LCA as subgroups may be limited when individual variables already naturally cluster together across an entire sample.

The final measures used to form subgroups were three health status constructs: global health status (SF-1),[23,24] functional status (Activities of Daily Living; ADLs),[25] smoking status,[26] and seven SDH constructs: loneliness (Three-Item Loneliness Scale),[27] sleep quality (PROMIS-Sleep),[28] material needs,[29] depression (Patient Health Questionnaire Screening Questions; PHQ-2),[30] social support (Medical Outcomes Study Social Support Survey; MOS-8),[31] patient activation (adapted from Consumer Health Activation Index; CHAI),[32] and recent life stressors.[33] Importantly, while 10 constructs were chosen, the material needs scale was thereafter broken out into its two subscales of medication insecurity and food insecurity questions, resulting in a final set of 11 constructs (8 SDH constructs). All-cause 180-day hospitalization was obtained from VA EHR and the timeframe was defined beginning at 7 days prior to the study team's receipt of each participant's completed survey. from the endpoint of the 1-year outpatient visit observation window (3/18/2018).

Latent Class Analysis

Each of the 11 survey measures was dichotomized based on existing evidence, expert clinical judgment from authors, and observed sample distributions with the goal of producing the maximally useful cut points to identify latent class membership (eTable 1). Smoking status, depression, and loneliness had strong evidence for *a priori* cut points. Global health status, functional status, food insecurity, medication insecurity, recent life stressors, and sleep quality had moderate evidence for cut points based on literature and expert clinical judgment. Patient activation and social support cut points were based solely on the empirical distribution from the analytic sample. All dichotomizations, and therefore probabilities of endorsement in the latent class solution, were coded such that higher values (or greater probabilities) represented outcomes associated with poorer health.

Latent classes were identified based on the 11 patient-reported, dichotomized, survey measures using the Latent Class Analysis (LCA) Stata Plugin (version 1.2).[34,35] Parameters are estimated by maximum likelihood using the expectation-maximization algorithm (FIML), allowing for missing data assuming missing at random.[34] These classes were also identified in R (version 3.3.6) using the poLCA package as a robustness check.[36] To determine the optimal class solution, we evaluated each class solution on: 1) Bayesian information criterion (BIC), Akaike's Information Criterion (AIC), consistent AIC (CAIC) to determine variance explained and parsimony, 2) scaled relative entropy to determine separation among classes, 3) size and interpretability of classes, and 4) bootstrapped likelihood ratio tests using the LCA Bootstrap Stata Function (version 1.0) for variance explained between solutions.[37,38] While no item probability thresholds are recommended, we opted to consider item probabilities ≥60% as representative for each class, based on prior studies.[39,40] If items were endorsed across all classes, we conducted sensitivity analyses to explore class solutions with those items removed.

Statistical Analysis

Veteran characteristics compared across the final latent classes included age, sex, race/ethnicity, copay exemption, body mass index (BMI), residence in rural area, JEN frailty index[41] (an algorithm of 13 diagnostic domains representing long-term institutionalization risk), and Gagne comorbidity score (a comorbidity score that considers both Charlson and Elixhauser comorbidity scores with improved predictive capability).[42] These characteristics were compared among the latent classes using chi-squared tests for categorical variables and Kruskal-Wallis tests for continuous variables. A multinomial logistic regression model was estimated to examine whether these characteristics were associated with membership in each class, which was coded as a categorical dependent variable. For these analyses, class membership for each Veteran was assigned based on the highest probability among the five classes for each individual from the LCA.

Unadjusted 180-day VA hospitalization rates were then compared across classes using the LCA Distal Stata Function,[43,44] which is warranted when the identified latent classes are not well separated (i.e., lower entropy). This function uses a three-step method that first estimates parameters of the LCA model without the distal outcome, then uses posterior probabilities of class membership from this model to compute a weighting variable, and finally uses the weighting variable to calculate a weighted average of the distal outcome for each class (via logistic regression in the current analyses).[45] All analyses were conducted in Stata version 15, Alpha level was set to 0.05.

The survey for this study was administered in partnership with VHA's Office of Primary Care; the survey and this analysis were designated as non-research quality improvement by the VHA program office and the Durham VAMC Institutional Review Board.

Results

Of the 10,000 Veterans sent the mail survey, a total of 4,685 Veterans responded. Of those, one respondent was excluded because the Veteran did not answer any questions related to the 11 survey measures used for the latent class analysis, resulting in a final sample size of 4,684 (46.8% response rate). The full analysis of survey responders and non-responders is detailed elsewhere.[22] Responders and nonresponders were similar on most demographic and clinical information, with the primary differences being that responders were older (86.3% ≥ 60 years old vs. 69.4% ≥ 60 years old), more likely to be white (71.8% vs. 61.3%), and more likely to have diagnosed hypertension (73.1% vs. 63.3%). Additional missing data were sparse and assumed to be missing at random, supporting the use of LCA's FIML as noted above. A total of 468 (10.0%) hospitalizations and 87 (1.9%) deaths were observed in the 180-day period.

Latent Class Analysis

Although the BIC and AIC were lowest for the solution with six classes (eTable 2, eFigure 1), CAIC was lowest for the solution with five classes. Entropy decreased from the two-class solution through the five-class solution, before increasing slightly again in the six-class solution. The six-class solution identified through Stata was not chosen because one of the classes included <4% of the sample and could not be identified in R. Based on interpretability of classes, and a bootstrapped likelihood ratio test p-value of less than 0.01 when comparing 4-class and 5-class solutions, we retained the 5-class solution.

Across the five classes, endorsement of at least one stressful life event was highly probable in each class (all classes ≥ 70%). A sensitivity analysis removing stressful life events produced an almost identical five-class solution. Therefore, we retained the *a priori* inclusion of stressful life events in our models.

Based on examination of the probabilities of survey measure endorsement across classes (eFigure 2), the five classes were named: 1) poor/fair health with few SDH vulnerabilities, 2) social isolation, 3) minimal SDH vulnerabilities, 4) multiple SDH vulnerabilities, and 5) multiple SDH vulnerabilities without food or medication insecurity (Table 1). The 'minimal SDH vulnerabilities' subgroup comprised the largest proportion of the sample (31%) and was chosen as the reference group for statistical comparisons of 180-day hospitalization rates across classes.

Demographic, Comorbidity and Hospitalization Differences between Classes

We calculated descriptive statistics for Veteran characteristics after assigning the highest probability class to each Veteran (see Table 2). There was significant variation across classes in each of the descriptive variables. Odds ratios and 95% confidence intervals (CIs) are presented in Table 3. In adjusted analyses, Veterans who were younger (as opposed to older), male (as opposed to female), Black Non-Hispanic (as opposed to White Non-Hispanic), copay-exempt (as opposed to copay non-exempt), and have a higher (as opposed to lower) Jen Frailty Index were each more likely to be in the two subgroups with 'multiple SDH vulnerabilities' (Classes 4 and 5) than the 'minimal SDH vulnerabilities' subgroup (Table 3).

Compared to the ‘minimal SDH vulnerabilities’ subgroup, those in the ‘poor/fair health with few SDH vulnerabilities’ subgroup were more likely to be younger, copay-exempt, reside in a rural area, and have a higher Gagne score. Those in the ‘social isolation’ subgroup were more likely to be younger, copay-exempt, have a lower BMI, and have a lower Gagne score than the ‘minimal SDH vulnerabilities’ subgroup. All subgroups were more likely to have higher CAN scores than the ‘minimal SDH vulnerabilities’ subgroup.

The 180-day VA hospitalization rate was 8% for Veterans in the ‘minimal SDH vulnerabilities’ subgroup, 10% of Veterans in each of the ‘social isolation’ and ‘multiple SDH vulnerabilities without food or medication insecurity’ subgroups, and 12% of Veterans in the ‘multiple SDH vulnerabilities’ and ‘poor/fair health with few SDH vulnerabilities’ subgroups (Figure 1). In logistic regression via the LCA Distal Function, odds of 180-day hospitalization mirrored the hospitalization rates (Table 4). Compared to the ‘minimal SDH vulnerabilities’ subgroup, Veterans in the ‘multiple SDH vulnerabilities’ subgroup had higher odds (OR=1.53, 95% CI: 1.09, 2.14) of hospitalization. The other three subgroups also had higher odds of 180-day hospitalization compared to the ‘minimal SDH vulnerabilities’ subgroup, but all 95% confidence intervals included 1.00.

Discussion

There is increasing recognition that health outcomes are greatly impacted by social and environmental factors,[11,13,14] but VA investigations into the impact of SDH on utilization outcomes have thus far focused on individual SDH predictions rather than the role of multiple SDH factors.[46] Segmentation of populations into clinically coherent subgroups is recognized as one strategy to help tailor interventions to individuals with common needs,[18] but most prior risk segmentation analyses have identified clinical subgroups on the basis of diagnosed conditions.[4,47] Thus, this study is novel as the only VA investigation, largest investigation, and one of only two investigations to the authors’ knowledge, to subgroup high-risk patients based on patient-reported SDH measures rarely captured in standard EHRs. This study may therefore offer insights into innovative interventions to reduce hospitalizations and improve health outcomes.

On the basis of 11 self-reported patient-level measures, we identified five distinct subgroups from latent class analysis within the subset of the Veteran population that is high risk for hospitalization or death in the upcoming year. These findings suggest that it may be difficult to capture these subgroups with demographic and comorbidity measures found in EHRs.

These distinct subgroups had 180-day VA admission rates that differed somewhat in expected ways. The ‘minimal SDH vulnerabilities’ subgroup had the lowest admission rate (8%) and Veterans in the ‘multiple SDH vulnerabilities’ subgroup and the ‘poor/fair health with few SDH vulnerabilities’ subgroup had a 50% higher admission rate (12%). Odds ratios were only significant for the ‘multiple SDH vulnerabilities’ subgroup coefficient, suggesting that endorsement of more self-reported SDH vulnerabilities increased risk of hospitalization in a way that may not be clear when any individual or small combination of SDH measures is used to predict hospitalization. These findings suggest that there may be benefit to assessing SDH factors in patients who are at high-risk for hospitalization, and referring individuals with multiple SDH vulnerabilities to appropriate clinical and social services. Future work should examine whether other outcomes differ across these five subgroups of high-risk patients.

At least two of the SDH subgroups that emerged in this study ('minimal SDH vulnerabilities' and 'multiple SDH vulnerabilities') are conceptually similar to the patient-reported SDH subgroups from the recent study of patients in a regional Kaiser Permanente health system, and appear to have a similar magnitude of difference with regard to risk of hospitalization/inpatient utilization.[15] These similarities across VA and Kaiser Permanente suggest that these two groups may be found in multiple health systems, and do not represent a data artifact. Our remaining three SDH subgroups, 'poor/fair health with few SDH vulnerabilities', 'social isolation', and 'multiple SDH vulnerabilities without food and medication insecurity' appear to identify more distinct SDH patterns. These differences may be especially important when investigating *which* SDH to measure for clinical impact and intervention tailoring, rather than *how many* SDH to measure. These differences may also be a function of the Veteran population sampled, as their SDH vulnerabilities likely differ from those of their non-Veteran peers.

Food and medication insecurity, which have been associated with poorer diabetes control,[48] appear to meaningfully differentiate among these subgroups of Veterans who have multiple SDH vulnerabilities. Other investigations have also noted the importance of distinguishing among the different profiles of complexity in patients with complex medical vulnerabilities, rather than stopping at the distinction of "complex".[47] In the current study, however, SDH complexity (represented best in the 'multiple SDH vulnerabilities' subgroup) and medical complexity (which may be best represented in the 'poor/fair health with few SDH vulnerabilities' subgroup) appear associated with the highest odds of VA hospitalization within 180 days – our primary outcome. While the identification of SDH complexity with numerous measures (8 of 11) in the current sample suggests robust measurement of SDH may be necessary to identify the patients most at-risk, it also suggests this identification may pay dividends in future risk prediction and therefore generate more efficient targeting and tailoring of interventions.

It is important to consider potential adverse consequences of grouping patients based on SDH characteristics. A primary concern may be the ethical challenges to identifying vulnerabilities that a system is not prepared to address, and privacy and security issues arise when asking about these issues at point of care and incorporating them into a patient's health record.[49,50] The VA is in a strong position to address these challenges (as evidenced by the system-wide protocols in place for patients who screen positive for Military Sexual Trauma, PTSD, and alcohol use). Nevertheless, systems that implement SDH assessments should ensure that clinicians can offer resources or referrals for any vulnerabilities exposed through this process. In addition, excessive reliance on population subgroups could result in a patient's individual needs being missed if those needs fall outside of common needs for the group. It is therefore important that systems use the groups as a starting point for considering a patient's needs and appropriateness for programs, but conduct a more comprehensive assessment before offering specific services.

There are several limitations that must be acknowledged. First, these results are not generalizable beyond Veterans identified as high-risk from the VA CAN score in 2018. However, this group represented a high-probability sample for observing an important but low-probability event of 180-day hospitalization. Second, latent class analysis is only as good as the variables used. We chose measures in an iterative group process, but the final five-class solution had entropy that was lower than ideal. This five-class solution would benefit from replication in future work. It is possible that there are more ideal measures we did not incorporate that would have better distinguished among these subgroups, such as resilience and life chaos which were found to be predictive of 180-day hospitalization in a prior paper.[22] Nevertheless, these subgroups, albeit imperfect, were associated with meaningful differences in the impactful and observable outcome of 180-day hospitalization. Finally, this mail survey likely under-represents Veterans with unstable housing or homelessness.

In conclusion, study findings suggest that patient-reported SDH measures can facilitate population segmentation into meaningful subgroups that might not emerge if limited to current EHR data. Future research should seek to further refine these subgroups, as well as examine whether they may benefit from distinct, tailored interventions to reduce their risk of hospital admission and other adverse events. With additional refinement, widespread efforts to implement SDH screening and integrate this information into the EHR may present an opportunity for health systems to incorporate this information into population segmentation efforts to inform patient care and resource allocations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

FUNDING AND ACKNOWLEDGEMENTS

This study was funded by the VHA Office of Primary Care (XVA 21–159) (PIs Maciejewski and Zulman). This work was supported by the Office of Research and Development, Health Services Research and Development Service, Department of Veterans Affairs and the Durham Center of Innovation to Accelerate Discovery and Practice Transformation (ADAPT), (CIN 13–410) at the Durham VA Health Care System. Dr. Blalock was supported by Grant No. TPH 21–000 from the Department of Veterans Affairs Office of Academic Affiliations. Dr. Maciejewski was supported by a Research Career Scientist award from the Department of Veterans Affairs (RCS 10–391). Dr. Whitson's contributions to this work were supported by the National Institutes of Health (UL1TR002553, UH3 AG056925, R33AG057806, and P30AG028716).

References

1. Yoon J, Zulman D, Scott JY, Maciejewski ML. Costs associated with multimorbidity among VA patients. Med Care. 2014;52 Suppl 3:31. 10.1097/MLR.0000000000000061 [doi].
2. Goodman RA, Ling SM, Briss PA, Parrish RG, Salive ME, Finke BS. Multimorbidity patterns in the united states: Implications for research and clinical practice. J Gerontol A Biol Sci Med Sci. 2016;71(2):215–220. 10.1093/gerona/glv199 [doi]. [PubMed: 26714567]
3. Kronick RG, bella M, gilmer TP, somers SA. faces of medicaid II: Recognizing the care needs of people with multiple chronic conditions. https://www.chcs.org/resource/the-faces-of-medicaid-ii-recognizing-the-care-needs-of-people-with-multiple-chronic-conditions/. Accessed may 6, 2020.
4. Joynt KE, Figueroa JF, Beaulieu N, Wild RC, Orav EJ, Jha AK. Segmenting high-cost medicare patients into potentially actionable cohorts. Healthc (Amst). 2017;5(1–2):62–67. doi: S2213-0764(16)30228-7 [pii]. [PubMed: 27914968]

5. Figueroa JF, Zhou X, Jha AK. Characteristics and spending patterns of persistently high-cost medicare patients. Health Aff (Millwood). 2019;38(1):107–114. 10.1377/hlthaff.2018.05160 [doi]. [PubMed: 30615516]
6. Wolff JL, Starfield B, Anderson G. Prevalence, expenditures, and complications of multiple chronic conditions in the elderly. Arch Intern Med. 2002;162(20):2269–2276. doi: 20083 [pii]. [PubMed: 12418941]
7. Pham HH, Schrag D, O'Malley AS, Wu B, Bach PB. Care patterns in medicare and their implications for pay for performance. N Engl J Med. 2007;356(11):1130–1139. doi: 356/11/1130 [pii]. [PubMed: 17360991]
8. Ward BW, Schiller JS. Prevalence of multiple chronic conditions among US adults: Estimates from the national health interview survey, 2010. Prev Chronic Dis. 2013;10:E65. 10.5888/pcd10.120203 [doi]. [PubMed: 23618545]
9. Machlin SR, Soni A. Health care expenditures for adults with multiple treated chronic conditions: Estimates from the medical expenditure panel survey, 2009. Prev Chronic Dis. 2013;10:E63. 10.5888/pcd10.120172 [doi]. [PubMed: 23618543]
10. Zulman DM, Pal Chee C, Wagner TH, et al. Multimorbidity and healthcare utilisation among high-cost patients in the US veterans affairs health care system. BMJ Open. 2015;5(4):e007771–007771. 10.1136/bmjopen-2015-007771 [doi].
11. Adler NE, Stead WW. Patients in context--EHR capture of social and behavioral determinants of health. N Engl J Med. 2015;372(8):698–701. 10.1056/NEJMp1413945 [doi]. [PubMed: 25693009]
12. Kouzoukas D, Lazio J. Advance notice of methodological changes for calendar year (CY) 2020 for Medicare Advantage (MA) capitation rates, Part C and Part D payment policies and 2020 draft call letter. https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Downloads/Advance2020Part2.pdf. Updated 2019. Accessed October 13, 2020.
13. Krist AH, Davidson KW, Ngo-Metzger Q, Mills J. Social determinants as a preventive service: U.S. preventive services task force methods considerations for research. Am J Prev Med. 2019;57(6 Suppl 1):S6–S12. doi: S0749-3797(19)30321-6 [pii]. [PubMed: 31753280]
14. Fichtenberg CM, Alley DE, Mistry KB. Improving social needs intervention research: Key questions for advancing the field. Am J Prev Med. 2019;57(6 Suppl 1):S47–S54. doi: S0749-3797(19)30326-5 [pii]. [PubMed: 31753279]
15. Rogers A, Hu YR, Schickedanz A, Gottlieb L, Sharp A. Understanding high-utilizing patients based on social risk profiles: A latent class analysis within an integrated health system. J Gen Intern Med. 2020. 10.1007/s11606-019-05510-9 [doi].
16. Yoon J, Chang E, Rubenstein LV, et al. Impact of primary care intensive management on high-risk veterans' costs and utilization: A randomized quality improvement trial. Ann Intern Med. 2018;168(12):846–854. 10.7326/M17-3039 [doi]. [PubMed: 29868706]
17. Chang ET, Zulman DM, Asch SM, et al. An operations-partnered evaluation of care redesign for high-risk patients in the veterans health administration (VHA): Study protocol for the PACT intensive management (PIM) randomized quality improvement evaluation. Contemp Clin Trials. 2018;69:65–75. doi: S1551-7144(18)30080-6 [pii]. [PubMed: 29698772]
18. 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. 10.1377/hlthaff.2015.1311 [doi]. [PubMed: 27140981]
19. Wang L, Porter B, Maynard C, et al. Predicting risk of hospitalization or death among patients receiving primary care in the veterans health administration. Med Care. 2013;51(4):368–373. 10.1097/MLR.0b013e31827da95a [doi]. [PubMed: 23269113]
20. Dillman DA. mail and telephone surveys: The total design method. New York, NY: John Wiley; 1978.
21. Zullig LL, Whitson HE, Hastings SN, et al. A systematic review of conceptual frameworks of medical complexity and new model development. J Gen Intern Med. 2016;31(3):329–337. 10.1007/s11606-015-3512-2 [doi]. [PubMed: 26423992]
22. Zulman DM, maciejewski ML, grubber J, et al. patient-reported social and behavioral determinants of health and predicted hospitalization in high-risk VA patients. JAMA Netw Open. 2020;3(10):e2021457. [PubMed: 33079198]

23. Zajacova A, Dowd JB. Reliability of self-rated health in US adults. Am J Epidemiol. 2011;174(8):977–983. 10.1093/aje/kwr204 [doi]. [PubMed: 21890836]

24. DeSalvo KB, Bloser N, Reynolds K, He J, Muntner P. Mortality prediction with a single general self-rated health question. A meta-analysis. J Gen Intern Med. 2006;21(3):267–275. doi: JGI291 [pii]. [PubMed: 16336622]

25. Fillenbaum GG, Smyer MA. The development, validity, and reliability of the OARS multidimensional functional assessment questionnaire. J Gerontol. 1981;36(4):428–434. 10.1093/geronj/36.4.428 [doi]. [PubMed: 7252074]

26. Global adult tobacco survey collaborative group. tobacco questions for surveys: A subset of key questions from the global adult tobacco survey (GATS), 2nd edition. Atlanta, GA: Centers for disease control and prevention 2011.

27. Hughes ME, Waite LJ, Hawkley LC, Cacioppo JT. A short scale for measuring loneliness in large surveys: Results from two population-based studies. Res Aging. 2004;26(6):655–672. 10.1177/0164027504268574 [doi]. [PubMed: 18504506]

28. Yu L, Buysse DJ, Germain A, et al. Development of short forms from the PROMIS sleep disturbance and sleep-related impairment item banks. Behav Sleep Med. 2011;10(1):6–24. 10.1080/15402002.2012.636266 [doi]. [PubMed: 22250775]

29. Berkowitz SA, Seligman HK, Choudhry NK. Treat or eat: Food insecurity, cost-related medication underuse, and unmet needs. Am J Med. 2014;127(4):303–310.10.1016/j.amjmed.2014.01.002 [doi]. [PubMed: 24440543]

30. Kroenke K, Spitzer RL, Williams JB. The patient health questionnaire-2: Validity of a two-item depression screener. Med Care. 2003;41(11):1284–1292. 10.1097/01.MLR.0000093487.78664.3C [doi]. [PubMed: 14583691]

31. Moser A, Stuck AE, Silliman RA, Ganz PA, Clough-Gorr KM. The eight-item modified medical outcomes study social support survey: Psychometric evaluation showed excellent performance. J Clin Epidemiol. 2012;65(10):1107–1116. 10.1016/j.jclinepi.2012.04.007 [doi]. [PubMed: 22818947]

32. Wolf MS, Smith SG, Pandit AU, et al. Development and validation of the consumer health activation index. Med Decis Making. 2018;38(3):334–343. 10.1177/0272989X17753392 [doi]. [PubMed: 29436308]

33. Cornoni-Huntley J, Ostfeld AM, Taylor JO, et al. Established populations for epidemiologic studies of the elderly: Study design and methodology. Aging (Milano). 1993;5(1):27–37. 10.1007/BF03324123 [doi]. [PubMed: 8481423]

34. LCA stata plugin (version 1.2) [software]. (2015). University park: The methodology center, Penn State. Retrieved from methodology.psu.edu.

35. Lanza ST, dziak JJ, huang L, wagner AT, & collins LM (2015). LCA stata plugin users' guide (version 1.2). University park: The methodology center, Penn State. Retrieved from methodology.psu.edu.

36. linzer Drew A., lewis jeffrey B. (2011). poLCA: An R package for polytomous variable latent class analysis. journal of statistical software, 42(10), 1–29. http://www.jstatsoft.org/v42/i10/.

37. LCA bootstrap stata function (version 1.0) [software]. (2016). University park: The methodology center, Penn State. Retrieved from methodology.psu.edu.

38. Huang L, dziak JJ, wagner AT, & lanza ST (2016). LCA bootstrap stata function users' guide (version 1.0). University park: The methodology center, Penn State. Retrieved from methodology.psu.edu.

39. Rinehart DJ, Oronce C, Durfee MJ, et al. Identifying subgroups of adult superutilizers in an urban safety-net system using latent class analysis: Implications for clinical practice. Med Care. 2018;56(1):e1–e9. 10.1097/MLR.0000000000000628 [doi]. [PubMed: 27632768]

40. Goldsmith ES, MacLehose RF, Jensen AC, et al. Complementary, integrative, and nondrug therapy use for pain among US military veterans on long-term opioids. Med Care. 2020;58 Suppl 2 9S(Suppl 2 9 Suppl):S116–S124. 10.1097/MLR.0000000000001333 [doi]. [PubMed: 32826781]

41. Kinosian B, Wieland D, Gu X, Stallard E, Phibbs CS, Intrator O. Validation of the JEN frailty index in the national long-term care survey community population: Identifying functionally

impaired older adults from claims data. BMC Health Serv Res. 2018;18(1):908–2. 10.1186/s12913-018-3689-2 [doi]. [PubMed: 30497450]

42. Gagne JJ, Glynn RJ, Avorn J, Levin R, Schneeweiss S. A combined comorbidity score predicted mortality in elderly patients better than existing scores. J Clin Epidemiol. 2011;64(7):749–759. 10.1016/j.jclinepi.2010.10.004 [doi]. [PubMed: 21208778]
43. LCA distal BCH stata function (version 1.1) [software]. (2017). University park: The methodology center, Penn State. Retrieved from methodology.psu.edu.
44. Huang L, dziak JJ, bray BC, & wagner AT (2017). LCA distal stata function users' guide (version 1.1). University park: The methodology center, Penn State. Retrieved from methodology.psu.edu.
45. Bolck A, croon M, hagenaars J. estimating latent structure models with categorical variables: One-step versus three-step estimators. political analysis. 2004;12(1):3–27.
46. Davis CI, Montgomery AE, Dichter ME, Taylor LD, Blosnich JR. Social determinants and emergency department utilization: Findings from the veterans health administration. Am J Emerg Med. 2020;38(9):1904–1909. doi: S0735-6757(20)30441-1 [pii]. [PubMed: 32739860]
47. Prenovost KM, Fihn SD, Maciejewski ML, Nelson K, Vijan S, Rosland AM. Using item response theory with health system data to identify latent groups of patients with multiple health conditions. PLoS One. 2018;13(11):e0206915. 10.1371/journal.pone.0206915 [doi]. [PubMed: 30475823]
48. Berkowitz SA, Meigs JB, DeWalt D, et al. Material need insecurities, control of diabetes mellitus, and use of health care resources: Results of the measuring economic insecurity in diabetes study. JAMA Intern Med. 2015;175(2):257–265. 10.1001/jamainternmed.2014.6888 [doi]. [PubMed: 25545780]
49. Lynn J Downsides of incorporating behavioral and social data into an EHR. http://www.emrandhipaa.com/emr-and-hipaa/2015/06/19/downsides-of-incorporating-behavioral-and-social-data-into-an-ehr/. Updated 2015. Accessed October 13, 2020.
50. Garg A, Boynton-Jarrett R, Dworkin PH. Avoiding the unintended consequences of screening for social determinants of health. JAMA. 2016;316(8):813–814. 10.1001/jama.2016.9282 [doi]. [PubMed: 27367226]

Author Manuscript

Author Manuscript

Author Manuscript

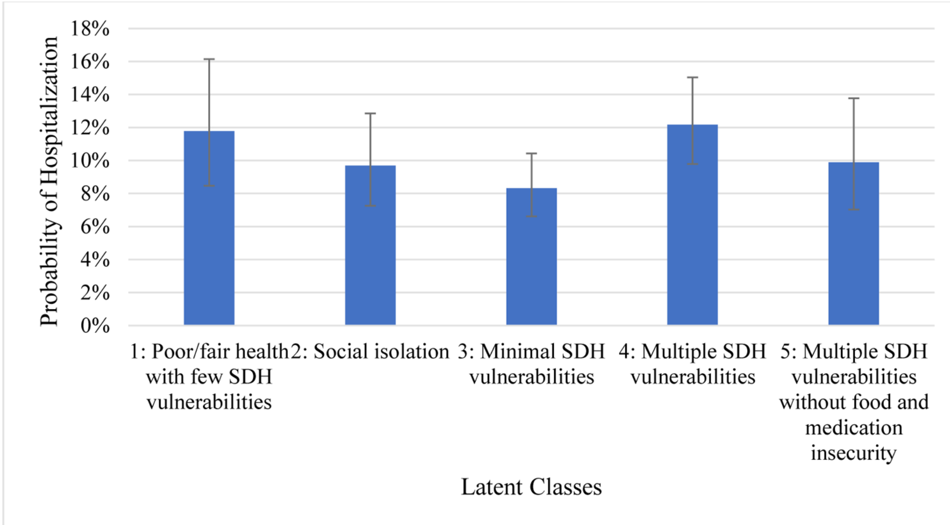

Figure 1.

Probability of 180-day Hospitalization by Class

Table 1.

Parameter Estimates for 5-class Model

		Latent Class				
		1	2	3	4	5
		Poor/fair health with few SDH vulnerabilities	**Social isolation**	**Minimal SDH vulnerabilities**	**Multiple SDH vulnerabilities**	**Multiple SDH vulnerabilities without food and medication insecurity**
Class Membership Probabilities (n=4,684)		14%	22%	31%	19%	14%
Survey-based Indicator	Overall Proportion	Probability of Endorsing Each Indicator				
SDH Construct						
Depressed	0.27	0.28	0.13	0.02	0.71	0.59
Lonely	0.53	0.20	0.78	0.14	0.94	0.99
NO social support	0.53	0.24	0.89	0.25	0.87	0.69
NO patient activation	0.53	0.55	0.54	0.37	0.73	0.71
Sleep disturbed	0.29	0.35	0.21	0.09	0.61	0.45
Food insecure	0.21	0.10	0.20	0.03	0.80	0.00
Meds insecure	0.12	0.10	0.10	0.03	0.43	0.01
Stress	0.78	0.82	0.82	0.70	0.93	0.91
Health Construct						
Smokes	0.21	0.18	0.25	0.14	0.40	0.19
Poor/fair health	0.51	0.89	0.29	0.21	0.80	0.87
Needs help with ADL	0.31	0.58	0.14	0.12	0.53	0.57

ADL = activities of daily living; 1 survey respondent was missing values for all 11 indicator variables.

Red Text = Item Response Probabilities ≥ .60. Blue Text = Meds Insecure Response Probability <.60 used to define class due to the comparably high probability of endorsement when compared to all other classes.

Table 2.

Descriptive Statistics of EHR-based Patient Characteristics by Class

Characteristic	n	Class 1: Poor/fair health with few SDH vulnerabilities	Class 2: Social isolation	Class 3: Minimal SDH vulnerabilities	Class 4: Multiple SDH vulnerabilities	Class 5: Multiple SDH vulnerabilities without food and medication insecurity	p-value[a]
Age (n, (%))	4684						
<60		42 (8%)	165 (17%)	93 (6%)	236 (30%)	106 (14%)	<0.001
60–80		347 (68%)	668 (67%)	1123 (70%)	494 (63%)	530 (68%)	
>80		124 (24%)	167 (17%)	389 (24%)	58 (7%)	142 (18%)	
Male (n, (%))	4684	499 (97%)	923 (92%)	1547 (96%)	701 (89%)	720 (93%)	<0.001
VA Copay status (n, (%))	3571						
Exempt		319 (85%)	653 (84%)	969 (77%)	556 (92%)	490 (89%)	<0.001
Non-exempt		57 (15%)	126 (16%)	292 (23%)	49 (8%)	60 (11%)	
Race/Ethnicity (n, (%))	4412						
White non-Hispanic		391 (80%)	729 (77%)	1216 (81%)	491 (66%)	538 (74%)	<0.001
Black non-Hispanic		62 (13%)	157 (17%)	196 (13%)	189 (25%)	137 (19%)	
Hispanic		24 (5%)	43 (5%)	58 (4%)	42 (6%)	36 (5%)	
Other		8 (2%)	20 (2%)	33 (2%)	22 (3%)	20 (3%)	
BMI (n, (%))	4507						
<25		109 (22%)	226 (23%)	322 (21%)	141 (19%)	153 (21%)	<0.001
25-<30		146 (29%)	339 (35%)	533 (34%)	220 (29%)	225 (31%)	
30–<35		124 (25%)	255 (26%)	423 (27%)	199 (26%)	193 (26%)	
≥ 35		117 (24%)	153 (16%)	271 (17%)	195 (26%)	163 (22%)	
Reside in rural area (n, (%))	4681	222 (45%)	338 (35%)	588 (38%)	280 (37%)	303 (41%)	0.004
JEN Frailty Index (mean (sd))	4650	4.59 (1.87)	4.09 (1.67)	4.19 (1.73)	4.36 (1.75)	4.66 (1.85)	<0.001
Gagne (mean (sd))	4684	2.17 (2.22)	1.34 (1.68)	1.66 (1.91)	1.38 (1.76)	1.91 (2.10)	<0.001
CAN score	4684	86.60 (7.70)	84.83 (7.15)	84.30 (7.36)	85.44 (7.61)	86.21 (7.85)	<0.001

[a] p-values based on chi-squared tests for categorical variables and Kruskal-Wallis tests for continuous variables

Table 3.

Comparison of Characteristics Between Latent Classes and Minimal SDH Vulnerabilities Class from Multinomial Logistic Regression*

	Adjusted Odds Ratio (95% CI) (Ref = Minimal SDH vulnerabilities)			
	Class 1	**Class 2**	**Class 4**	**Class 5**
Characteristic	**Poor/fair health with few SDH vulnerabilities**	**Social isolation**	**Multiple SDH vulnerabilities**	**Multiple SDH vulnerabilities without food and medication insecurity**
Age				
<60	Ref	Ref	Ref	Ref
60–80	0.64 (0.42–0.97)[b]	0.39 (0.29–0.52)[a]	0.21 (0.16–0.28)[a]	0.45 (0.33–0.63)[a]
>80	0.66 (0.42–1.05)	0.27 (0.19–0.38)[a]	0.07 (0.05–0.11)[a]	0.35 (0.24–0.52)[a]
Male	1.35 (0.70–2.60)	0.68 (0.46–1.01)	0.56 (0.38–0.84)[a]	0.58 (0.38–0.89)[b]
VA Copay status				
Non-exempt	Ref	Ref	Ref	Ref
Exempt	1.79 (1.26–2.52)[a]	1.30 (1.01–1.67)[b]	2.74 (1.90–3.95)[a]	2.18 (1.57–3.02)[a]
Missing	2.13 (1.46–3.12)[a]	1.26 (0.95–1.68)	2.35 (1.57–3.50)[a]	2.73 (1.92–3.90)[a]
Race/Ethnicity				
White non-Hispanic	Ref	Ref	Ref	Ref
Black non-Hispanic	1.00 (0.72–1.38)	1.17 (0.92–1.49)	1.91 (1.49–2.44)[a]	1.46 (1.13–1.89)[a]
Hispanic	1.25 (0.75–2.10)	1.08 (0.71–1.65)	1.60 (1.03–2.49)[b]	1.25 (0.79–1.98)
Other	0.60 (0.26–1.39)	0.74 (0.41–1.36)	1.30 (0.72–2.34)	1.07 (0.58–1.95)
BMI				
<25	Ref	Ref	Ref	Ref
25-<30	0.86 (0.64–1.17)	0.88 (0.70–1.10)	0.99 (0.75–1.31)	0.94 (0.73–1.23)
30-<35	0.91 (0.67–1.25)	0.77 (0.60–0.99)[b]	0.93 (0.70–1.24)	0.91 (0.69–1.20)
≥ 35	1.26 (0.90–1.76)	0.70 (0.53–0.93)[b]	1.30 (0.96–1.75)	1.10 (0.82–1.48)
Reside in rural area	1.44 (1.15–1.79)[a]	0.92 (0.77–1.10)	1.12 (0.92–1.38)	1.22 (1.00–1.49)[b]
JEN Frailty Index	1.00 (0.93–1.08)	0.96 (0.91–1.02)	1.04 (0.98–1.11)	1.09 (1.02–1.16)[b]
Gagne	1.09 (1.02–1.15)[a]	0.90 (0.86–0.95)[a]	0.93 (0.88–0.99)[b]	1.03 (0.98–1.08)
CAN score	1.03 (1.01–1.05)[a]	1.03 (1.01–1.04)[a]	1.02 (1.01–1.04)[a]	1.02 (1.01–1.04)[a]

[a] p<0.01,

[b] p<0.05

[*] n=4,216 after excluding observations with missing covariate data (excluding copay status due to large proportion missing)

Table 4.

Association between Class Membership and 180-day Hospital Admission (n=4,684)

	Odds Ratio (95% CI)	**Probability of 180-Day Hospitalization (95% CI)**
Class 1 (Poor/fair health with few SDH vulnerabilities)	1.47 (0.89–2.42)	12% (8–16%)
Class 2 (Social isolation)	1.18 (0.75–1.86)	10% (7–13%)
Class 3 (Minimal SDH vulnerabilities)	Ref	8% (7–10%)
Class 4 (Multiple SDH vulnerabilities)	1.53 (1.09–2.14)	12% (10–15%)
Class 5 (Multiple SDH vulnerabilities without food and medication insecurity)	1.21 (0.78–1.87)	10% (7–14%)