

Examining the Interfacility Variation of Social Determinants of Health in the Veterans Health Administration

Charlie M. Wray, DO, MS; Marzieh Vali, MS; Louise C. Walter, MD; Lee Christensen, MS; Samir Abdelrahman, PhD; Wendy Chapman, PhD; and Salomeh Keyhani, MD, MPH

Introduction: Recently, numerous studies have linked social determinants of health (SDoH) with clinical outcomes. While this association is well known, the interfacility variability of these risk factors within the Veterans Health Administration (VHA) is not known. Such information could be useful to the VHA for resource and funding allocation. The aim of this study is to explore the interfacility variability of 5 SDoH within the VHA.

Methods: In a cohort of patients (aged ≥ 65 years) hospitalized at VHA acute care facilities with either acute myocardial infarction (AMI), heart failure (HF), or pneumonia in 2012, we assessed (1) the proportion of patients with any of the following five documented SDoH: lives alone, marginal housing, alcohol use disorder, substance use disorder, and use of substance use services, using administrative diagnosis codes and clinic stop codes; and (2) the documented facility-level variability of these SDoH. To examine whether variability was due to regional coding differences, we assessed the variation

of living alone using a validated natural language processing (NLP) algorithm.

Results: The proportion of veterans admitted for AMI, HF, and pneumonia with SDoH was low. Across all 3 conditions, lives alone was the most common SDoH (2.2% [interquartile range (IQR), 0.7-4.7]), followed by substance use disorder (1.3% [IQR, 0.5-2.1]), and use of substance use services (1.2% [IQR, 0.6-1.8]). Using NLP, the proportion of hospitalized veterans with lives alone was higher for HF (14.4% vs 2.0%, $P < .01$), pneumonia (11% vs 1.9%, $P < .01$), and AMI (10.2% vs 1.4%, $P < .01$) compared with *International Classification of Diseases, Ninth Edition* codes. Interfacility variability was noted with both administrative and NLP extraction methods.

Conclusions: The presence of SDoH in administrative data among patients hospitalized for common medical issues is low and variable across VHA facilities. Significant facility-level variation of 5 SDoH was present regardless of extraction method.

Social determinants of health (SDoH) are social, economic, environmental, and occupational factors that are known to influence an individual's health care utilization and clinical outcomes.^{1,2} Because the Veterans Health Administration (VHA) is charged to address both the medical and nonmedical needs of the veteran population, it is increasingly interested in the impact SDoH have on veteran care.^{3,4} To combat the adverse impact of such factors, the VHA has implemented several large-scale programs across the US that focus on prevalent SDoH, such as homelessness, substance abuse, and alcohol use disorders.^{5,6} While such risk factors are generally universal in their distribution, variation across regions, between urban and rural spaces, and even within cities has been shown to exist in private settings.⁷ Understanding such variability potentially could be helpful to US Department of Veterans Affairs (VA) policymakers and leaders to better allocate funding and resources to address such issues.

Although previous work has highlighted regional and neighborhood-level variability of SDoH, no study has examined the facility-level variability of commonly encountered

social risk factors within the VHA.^{4,8} The aim of this study was to describe the interfacility variation of 5 common SDoH known to influence health and health outcomes among a national cohort of veterans hospitalized for common medical issues by using administrative data.

METHODS

We used a national cohort of veterans aged ≥ 65 years who were hospitalized at a VHA acute care facility with a primary discharge diagnosis of acute myocardial infarction (AMI), heart failure (HF), or pneumonia in 2012. These conditions were chosen because they are publicly reported and frequently used for interfacility comparison.

Using the *International Classification of Diseases—9th Revision* (ICD-9) and VHA clinical stop codes, we calculated the median documented proportion of patients with any of the following 5 SDoH: lived alone, marginal housing, alcohol use disorder, substance use disorder, and use of substance use services for patients presenting with HF, MI, and pneumonia (Table). These SDoH were chosen because they are intervenable risk factors for which the VHA has several programs (eg,

Author affiliations can be found at the end of the article.

Correspondence:

Charlie M. Wray
(charlie.wray@ucsf.edu)

Fed Pract. 2021;38(1):15-19.
doi:10.12788/fp.0080

TABLE Patients With Social Determinants of Health

Variables	Individual Conditions			All Conditions
	HF	AMI	Pneumonia	
Hospitals, No. ^a	82	35	85	91
Patients, No.	9,853	2,776	9,362	21,991
Patient Demographics				
Age, mean (SD)	77 (8.9)	75 (8.5)	77 (9.1)	77 (9.0)
Male, No. (%)	9,699 (98.4)	2,731 (98.4)	9,170 (97.9)	21,578 (98.2)
Proportion of Patients With Social Determinants of Health, median (IQR)				
Lives alone (NLP)	14.4 (9.1-19.0) ^b	10.2 (4.3-17.1) ^b	11.0 (5.9-15.2) ^b	12.3 (8.4-17.2) ^b
Lives alone (ICD-9)	2.0 (1.0-5.2)	1.4 (0-3.4)	1.9 (0.7-5.4)	2.2 (0.7-4.7)
Marginal housing	0 (0-0.8)	0 (0-1.7)	1.2 (0-2.0)	0.6 (0.3-1.4)
Alcohol use disorder	0 (0-0.7)	0 (0-0)	0 (0-1.3)	0.5 (0-0.8)
Substance use disorder	1.2 (0-2.2)	1.6 (0-3.0)	1.3 (0-2.2)	1.3 (0.5-2.1)
Use of substance use services	0.9 (0-1.6)	1.0 (0-1.7)	1.6 (0-2.2)	1.2 (0.6-1.8)

Abbreviations: AMI, acute myocardial infarction; HF, heart failure, ICD-9, *International Classification of Diseases, Ninth Edition*; IQR, interquartile range; NLP, natural language processing.

^aFacility required ≥ 50 mentions of a social determinants of health to be included in analysis.

^bSignifies $P < .01$ when compared with 'lives alone' based on ICD-9 codes.

Administrative codes: 'lives alone' (V60.3, V60.4), 'marginal housing' (V60.0, V60.1), 'alcohol use disorder' (291.0, 303.xx, 305.0, 571.x), 'substance use disorder' (965.00, 965.01, 965.02, 965.09, E850.0, E850.1, E850.2, E935.0, E935.1, E935.2, 304.xx, 305.x, 965.0x, 967.0, 969.4, 969.6, 969.7, E850.0, E850.1, E850.2, E854.1, E854.2, E854.3, V58.6), 'use of substance use services' (US Department of Veterans Affairs clinic description codes: 130, 131)

homeless outreach, substance abuse, and tobacco cessation). To examine the variability of these SDoH across VHA facilities, we determined the number of hospitals that had a sufficient number of admissions (≥ 50) to be included in the analyses. We then examined the administratively documented, facility-level variation in the proportion of individuals with any of the 5 SDoH administrative codes and examined the distribution of their use across all qualifying facilities.

Because variability may be due to regional coding differences, we examined the difference in the estimated prevalence of the risk factor lives alone by using a previously developed natural language processing (NLP) program.⁹ The NLP program is a rule-based system designed to automatically extract information that requires inferencing from clinical notes (eg, discharge summaries and nursing, social work, emergency department physician, primary care, and hospital admission notes). For instance, the program identifies whether there was direct or indirect evidence that the patient did or did not live

alone. In addition to extracting data on lives alone, the NLP program has the capacity to extract information on lack of social support and living alone—2 characteristics without VHA interventions, which were not examined here. The NLP program was developed and evaluated using at least 1 year of notes prior to index hospitalization. Because this program was developed and validated on a 2012 data set, we were limited to using a cohort from this year as well.

All analyses were conducted using SAS Version 9.4. The San Francisco VA Medical Center Institutional Review Board approved this study.

RESULTS

In total, 21,991 patients with either HF (9,853), pneumonia (9,362), or AMI (2,776) were identified across 91 VHA facilities. The majority were male (98%) and had a median (SD) age of 77.0 (9.0) years. The median facility-level proportion of veterans who had any of the SDoH risk factors extracted through administrative codes was

low across all conditions, ranging from 0.5 to 2.2%. The most prevalent factors among patients admitted for HF, AMI, and pneumonia were lives alone (2.0% [Interquartile range (IQR), 1.0-5.2], 1.4% [IQR, 0-3.4], and 1.9% [IQR, 0.7-5.4]), substance use disorder (1.2% [IQR, 0-2.2], 1.6% [IQR: 0-3.0], and 1.3% [IQR, 0-2.2] and use of substance use services (0.9% [IQR, 0-1.6%], 1.0% [IQR, 0-1.7%], and 1.6% [IQR, 0-2.2%], respectively [Table]).

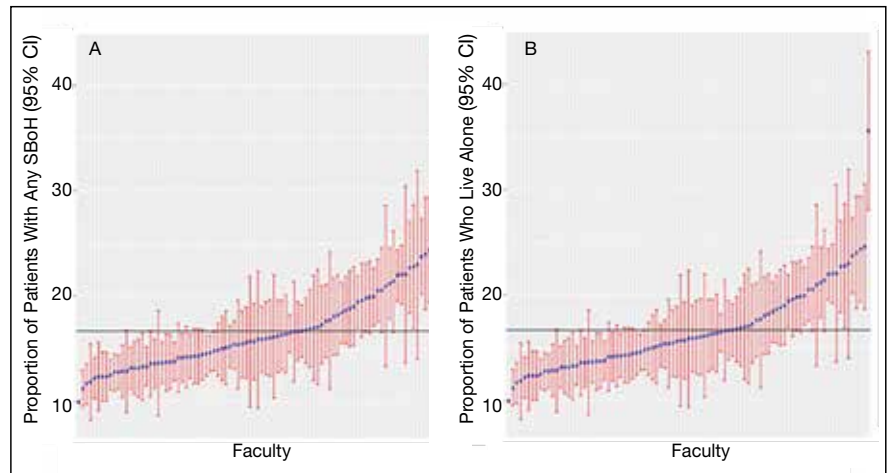
When utilizing the NLP algorithm, the documented prevalence of lives alone in the free text of the medical record was higher than administrative coding across all conditions (12.3% vs. 2.2%; $P < .01$). Among each of the 3 assessed conditions, HF (14.4% vs 2.0%, $P < .01$) had higher levels of lives alone compared with pneumonia (11% vs 1.9%, $P < .01$), and AMI (10.2% vs 1.4%, $P < .01$) when using the NLP algorithm. When we examined the documented facility-level variation in the proportion of individuals with any of the 5 SDoH administrative codes or NLP, we found large variability across all facilities—regardless of extraction method (Figure).

DISCUSSION

While SDoH are known to impact health outcomes, the presence of these risk factors in administrative data among individuals hospitalized for common medical issues is low and variable across VHA facilities. Understanding the documented, facility-level variability of these measures may assist the VHA in determining how it invests time and resources—as different facilities may disproportionately serve a higher number of vulnerable individuals. Beyond the VHA, these findings have generalizable lessons for the US health care system, which has come to recognize how these risk factors impact patients' health.¹⁰

Although the proportion of individuals with any of the assessed SDoH identified by administrative data was low, our findings are in line with recent studies that showed other risk factors such as social isolation (0.65%), housing issues (0.19%), and financial strain

FIGURE Facility-Level Variation of Social Risk Factors in VA Acute Care Facilities



Abbreviations: ICD-9, International Classification of Diseases–9th Revision; SDoH, social determinant of health. Point estimates with 95% CI. A, Proportion of patients with any SDoH extracted using ICD-9 and clinic description codes; (b) Proportion of patients with any SDoH who lives alone, extracted using natural language processing.

(0.07%) had similarly low prevalence.^{8,11} Although the exact prevalence of such factors remains unclear, these findings highlight that SDoH do not appear to be well documented in administrative data. Low coding rates are likely due to the fact that SDoH administrative codes are not tied to financial reimbursement—thus not incentivizing their use by clinicians or hospital systems.

In 2014, an Institute of Medicine report suggested that collection of SDoH in electronic health data as a means to better empower clinicians and health care systems to address social disparities and further support research in SDoH.¹² Since then, data collection using SDoH screening tools has become more common across settings, but is not consistently translated to standardized data due to lack of industry consensus and technical barriers.¹³ To improve this process, the Centers for Medicare and Medicaid Services created “z-codes” for the ICD-10 classification system—a subset of codes that are meant to better capture patients' underlying social risk.¹⁴ It remains to be seen if such administrative codes have improved the documentation of SDoH.

As health care systems have grown to understand the impact of SDoH on health outcomes, other means of collecting these data have evolved.^{1,10} For example, NLP-based extraction methods and electronic screening

tools have been proposed and utilized as alternative for obtaining this information. Our findings suggest that some of these measures (eg, lives alone) often may be documented as part of routine care in the electronic health record, thus highlighting NLP as a tool to obtain such data. However, other studies using NLP technology to extract SDoH have shown this technology is often complicated by quality issues (ie, missing data), complex methods, and poor integration with current information technology infrastructures—thus limiting its use in health care delivery.¹⁵⁻¹⁸

While variance among SDoH across a national health care system is natural, it remains an important systems-level characteristic that health care leaders and policymakers should appreciate. As health care systems disperse financial resources and initiate quality improvement initiatives to address SDoH, knowing that not all facilities are equally affected by SDoH should impact allocation of such resources and energies. Although previous work has highlighted regional and neighborhood levels of variation within the VHA and other health care systems, to our knowledge, this is the first study to examine variability at the facility-level within the VHA.^{2,4,13,19}

Limitations

There are several limitations to this study. First, though our findings are in line with previous data in other health care systems, generalizability beyond the VA, which primarily cares for older, male patients, may be limited.⁸ Though, as the nation's largest health care system, lessons from the VHA can still be useful for other health care systems as they consider SDoH variation. Second, among the many SDoH previously identified to impact health, our analysis only focused on 5 such variables. Administrative and medical record documentation of other SDoH may be more common and less variable across institutions. Third, while our data suggests facility-level variation in these measures, this may be in part related to variation in coding across facilities. However, the single SDoH variable extracted using NLP also varied at the facility-level, suggesting that coding may not entirely drive the variation observed.

CONCLUSIONS

As US health care systems continue to address SDoH, our findings highlight the various challenges in obtaining accurate data on a patient's social risk. Moreover, these findings highlight the large variability that exists among institutions in a national integrated health care system. Future work should explore the prevalence and variance of other SDoH as a means to help guide resource allocation and prioritize spending to better address SDoH where it is most needed.

Author Affiliations

Charlie Wray is an Internist in the Division of Hospital Medicine; **Marzieh Vali** is a Statistician in the Northern California Institute for Research and Education; **Louise Walter** is a Geriatrician in the Division of Geriatrics; and **Salomeh Keyhani** is an Internist in the Division of General Internal Medicine; all at the San Francisco Veterans Affairs Medical Center. **Lee Christensen** is a Project Manager and **Samir Abdelrahman** is an Assistant Professor, both in the Department of Biomedical Informatics, University of Utah in Salt Lake City. **Wendy Chapman** is the Associate Dean of Digital Health and Informatics in the Centre for Digital Transformation of Health, University of Melbourne, Victoria, Australia. Charlie Wray is an Assistant Professor of Medicine, Louise Walter and Salomeh Keyhani are Professors of Medicine; all in the Department of Medicine, University of California, San Francisco.

Acknowledgments

This work was supported by NHLBI R01 RO1 HL116522-01A1. Support for VA/CMS data is provided by the US Department of Veterans Affairs, Veterans Health Administration, Office of Research and Development, Health Services Research and Development, VA Information Resource Center (Project Numbers SDR 02-237 and 98-004).

Author disclosures

The authors report no actual or potential conflicts of interest with regard to this article.

Disclaimer

The opinions expressed herein are those of the authors and do not necessarily reflect those of *Federal Practitioner*, Frontline Medical Communications Inc., the US Government, or any of its agencies.

References

1. Social determinants of health (SDOH). <https://catalyst.nejm.org/doi/full/10.1056/CAT.17.0312>. Published December 1, 2017. Accessed December 8, 2020.
2. Hafez E, Searle KM, Predmore Z, et al. The Impact of Social Determinants of Health on hospitalization in the Veterans Health Administration. *Am J Prev Med*. 2019;56(6):811-818. doi:10.1016/j.amepre.2018.12.012
3. Lushniak BD, Alley DE, Ulin B, Graffunder C. The National Prevention Strategy: leveraging multiple sectors to improve population health. *Am J Public Health*. 2015;105(2):229-231. doi:10.2105/AJPH.2014.302257
4. Nelson K, Schwartz G, Hernandez S, Simonetti J, Curtis I, Fihn SD. The association between neighborhood environment and mortality: results from a national study of veterans. *J Gen Intern Med*. 2017;32(4):416-422. doi:10.1007/s11606-016-3905-x
5. Gundlapalli AV, Redd A, Bolton D, et al. Patient-aligned care team engagement to connect veterans experiencing homelessness with appropriate health

- care. *Med Care*. 2017;55 Suppl 9 Suppl 2:S104-S110. doi:10.1097/MLR.0000000000000770
6. Rash CJ, DePhilippis D. Considerations for implementing contingency management in substance abuse treatment clinics: the Veterans Affairs initiative as a model. *Perspect Behav Sci*. 2019;42(3):479-499. doi:10.1007/s40614-019-00204-3.
 7. Ompad DC, Galea S, Caiaffa WT, Vlahov D. Social determinants of the health of urban populations: methodologic considerations. *J Urban Health*. 2007;84(3 Suppl):i42-i53. doi:10.1007/s11524-007-9168-4
 8. Hatef E, Rouhizadeh M, Tia I, et al. Assessing the availability of data on social and behavioral determinants in structured and unstructured electronic health records: a retrospective analysis of a multilevel health care system. *JMIR Med Inform*. 2019;7(3):e13802. doi:10.2196/13802
 9. Conway M, Keyhani S, Christensen L, et al. Moonstone: a novel natural language processing system for inferring social risk from clinical narratives. *J Biomed Semantics*. 2019;10(1):6. doi:10.1186/s13326-019-0198-0
 10. Adler NE, Cutler DM, Fielding JE, et al. Addressing social determinants of health and health disparities: a vital direction for health and health care. Discussion Paper. *NAM Perspectives*. National Academy of Medicine, Washington, DC. doi:10.31478/201609t
 11. Cottrell EK, Dambrun K, Cowburn S, et al. Variation in electronic health record documentation of social determinants of health across a national network of community health centers. *Am J Prev Med*. 2019;57(6):S65-S73. doi:10.1016/j.amepre.2019.07.014
 12. Committee on the Recommended Social and Behavioral Domains and Measures for Electronic Health Records, Board on Population Health and Public Health Practice, Institute of Medicine. *Capturing Social and Behavioral Domains and Measures in Electronic Health Records: Phase 2*. National Academies Press (US); 2015.
 13. Gottlieb L, Tobey R, Cantor J, Hessler D, Adler NE. Integrating Social And Medical Data To Improve Population Health: Opportunities And Barriers. *Health Aff (Millwood)*. 2016;35(11):2116-2123. doi:10.1377/hlthaff.2016.0723
 14. Centers for Medicare and Medicaid Service, Office of Minority Health. Z codes utilization among medicare fee-for-service (FFS) beneficiaries in 2017. Published January 2020. Accessed December 8, 2020. <https://www.cms.gov/files/document/cms-omh-january2020-zcode-data-highlightpdf.pdf>
 15. Kharrazi H, Wang C, Scharfstein D. Prospective EHR-based clinical trials: the challenge of missing data. *J Gen Intern Med*. 2014;29(7):976-978. doi:10.1007/s11606-014-2883-0
 16. Weiskopf NG, Weng C. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. *J Am Med Inform Assoc*. 2013;20(1):144-151. doi:10.1136/amiajnl-2011-000681
 17. Anzaldi LJ, Davison A, Boyd CM, Leff B, Kharrazi H. Comparing clinician descriptions of frailty and geriatric syndromes using electronic health records: a retrospective cohort study. *BMC Geriatr*. 2017;17(1):248. doi:10.1186/s12877-017-0645-7
 18. Chen T, Dredze M, Weiner JP, Kharrazi H. Identifying vulnerable older adult populations by contextualizing geriatric syndrome information in clinical notes of electronic health records. *J Am Med Inform Assoc*. 2019;26(8-9):787-795. doi:10.1093/jamia/ocz093
 19. Raphael E, Gaynes R, Hamad R. Cross-sectional analysis of place-based and racial disparities in hospitalisation rates by disease category in California in 2001 and 2011. *BMJ Open*. 2019;9(10):e031556. doi:10.1136/bmjopen-2019-031556