

Using structured and unstructured data to identify patients' need for services that address the social determinants of health

Joshua R. Vest, PhD, MPH^{1,2*}

Shaun J. Grannis, MD, MS, FACMI^{2,3}

Dawn P. Haut, MD, MPH⁴

Paul Halverson, DrPH, FACHE^{1,2}

Nir Menachemi, PhD, MPH^{1,2}

¹Indiana University Richard M. Fairbanks School of Public Health, Indianapolis, IN

²Regenstrief Institute, Indianapolis, IN

³Indiana University School of Medicine, Indianapolis, IN

⁴Eskenazi Health, Indianapolis, IN

*Corresponding author

Joshua R Vest, PhD, MPH

1050 Wishard Blvd

Indianapolis, IN 46202

317 278 8410

317 274 3443 (f)

joshvest@iu.edu

This work was supported by the Robert Wood Johnson Foundation through the Systems for Action National Coordinating Center, ID 73485.

Number of pages: 11

Word count: 3033

Highlights

- Identifying patients with social determinants of health needs is challenging.
- A majority of patients need services to address social determinants of health.
- Measuring service need from structured data is insufficient.
- Unstructured data are necessary to give a complete picture of patient need.

ABSTRACT

Introduction: Increasingly, health care providers are adopting population health management approaches that address the social determinants of health (SDH). However, effectively

This is the author's manuscript of the article published in final edited form as:

Vest, J. R., Grannis, S. J., Haut, D. P., Halverson, P., & Menachemi, N. (2017). Using structured and unstructured data to identify patients' need for services that address the social determinants of health. *International Journal of Medical Informatics*. <https://doi.org/10.1016/j.ijmedinf.2017.09.008>

identifying patients needing services that address a SDH in primary care settings is challenging. The purpose of the current study is to explore how various data sources can identify adult primary care patients that are in need of services that address SDH.

Methods: A cross-sectional study described patients in need of SDH services offered by a safety-net hospital's federally qualified health center clinics. SDH services of social work, behavioral health, nutrition counseling, respiratory therapy, financial planning, medical-legal partnership assistance, patient navigation, and pharmacist consultation were offered on a co-located basis and were identified using structured billing and scheduling data, and unstructured electronic health record data. We report the prevalence of the eight different SDH service needs and the patient characteristics associated with service need. Moreover, characteristics of patients with SDH services need documented in structured data sources were compared with those documented by unstructured data sources.

Results: More than half (53%) of patients needed SDH services. Those in need of such services tended to be female, older, more medically complex, and higher utilizers of services. Structured and unstructured data sources exhibited poor agreement on patient SDH services need. Patients with SDH services need documented by unstructured data tended to be more complex.

Discussion: The need for SDH services among a safety-net population is high. Identifying patients in need of such services requires multiple data sources with structured and unstructured data.

Key words: medical informatics; social determinants of health; primary health care;

1. INTRODUCTION

Recent changes to reimbursement policies in the United States (US) have actively incentivized preventative care, cost control, accountability for health, and a renewed focus on the quality of care (1). In response, (US) providers have been adopting population health management approaches (2) that address the underlying behavioral, social, contextual, and environmental drivers of health status and health care utilization (3). These factors, typically referred to as social determinants of health (SDH) (4, 5), have typically been outside the scope of medical practice (6). Nonetheless, health care organizations are increasingly offering the services of social workers (7), patient navigators (8), legal experts (9), behavioral health (10) and other professionals that directly address SDH to improve health and reduce costs in the population for which they care.

A key goal of successful population health management programs involves identifying patients at risk for developing poor outcomes due to SDH factors. As such, health care organizations' care delivery processes can benefit by better identifying patients in need of services that may address underlying SDH issues (11). However, identifying such patients is challenging, particularly in the primary care setting. Documentation of both the need for, and actual delivery of, services that address SDH issues tends to occur within electronic health records (EHR) infrequently; and when it does, it is captured in a non-standard format (12). Moreover, services that address a SDH are often not associated with any billing or diagnosis codes further limiting routine, structured collection in EHRs and other administrative systems (13). To overcome this limitation, institutions have implemented patient surveys to collect SDH factors (14-16) but these data, even when computerized, may not be interoperable with other information systems (17) or

their use may be viewed by patients as controversial or stigmatizing (16). In general, most providers do not have access to comprehensive information about patients' SDH services need (18, 19).

The purpose of the current study is to explore how effectively various data sources support identifying adult primary care patients that need services that address SDH. Specifically, we examine the separate and joint contributions of structured and unstructured data in identifying patients' needs for SDH services. In addition, we characterized the patients in need for SDH services and how patient characteristics differed according to SDH services need determined by structured and unstructured data sources.

2. METHODS

This cross-sectional study described patients in need of SDH services offered by a safety-net hospital's federally qualified health center (FQHC) clinics. SDH services offered on a co-located basis were identified using structured billing and scheduling data, and unstructured EHR data. We report the prevalence of different SDH service needs and the patient characteristics associated with need. Next, we compared the characteristics of patients with SDH services need documented by structured data sources with those documented by unstructured data sources.

2.1 Sample & Setting

We identified needs for services addressing SDH in a sample of 73,085 adult (>18 years) patients seeking care from Eskenazi Health between April 2012 and July 2016. Eskenazi Health is the Indianapolis, IN metropolitan area public safety-net provider with a 315-bed hospital and 10 FQHC sites. Eskenazi Health offered SDH services on a co-located basis at the primary care sites

during the study period, i.e. Eskenazi Health was not relying on referrals to external providers and agencies for these services. All patients had at least one outpatient encounter during the study period.

2.2 Data

Information reflected multiple information systems linked through common patient identifiers. The primary data sources were Eskenazi Health's home grown EHR and the local community health information exchange (HIE). We also accessed appointment data from four different outpatient registration and scheduling systems.

2.3 Measuring the need for SDH services

SDH services of interest were: social work, behavioral health, nutrition counseling, respiratory therapy, financial planning, medical-legal partnership assistance, patient navigation, and pharmacist consultation. Realizing that patients' need for services can be met or unmet, we operationalized need broadly as any documentation that a health care professional judged that any of the aforementioned services were appropriate, recommended, or potentially beneficial to the patient. Therefore, we considered any received services as indicative of need as well as scheduled appointments regardless of whether they were kept or not. Need for each SDH service was operationalized as binary yes or no variables.

The eight SDH services had different workflows and documentation processes, which required searching multiple information sources to identify patients in need of services. First, we examined registration and scheduling systems for any patient appointments for nutritional counseling, behavioral health, respiratory therapy services, patient navigation, and financial

counseling. The appointment data reflected all kept, canceled, and “no show” visits. Second, we queried diagnosis and billing codes from the EHR and the HIE for ICD-9, ICD-10, and CPT procedure codes associated with behavioral health, nutritionist, respiratory therapy, and pharmacist consultation. Third, we reviewed the EHR’s unstructured data (i.e. orders and notes) for additional documentation of SDH services need. We searched provider orders for all eight of SDH services using keywords informed by a review of the literature and professional society’s documentation recommendations (see Appendix for sources and a tabular list of SDH by data source). Additionally, because the EHR progress notes contained a specific section for social worker documentation, we used natural language processing to identify instances of social worker contact with patients (see Appendix). We excluded any notes associated with inpatient admission periods and limited the notes to only those signed by social workers practicing in the outpatient setting. For each service, the source of documented need could be structured (i.e. billing and appointments), unstructured (i.e. orders and notes), or documented by both. Additionally, we created a summary variable of any documented need based on the presence of need for any of the eight SDH services.

2.4 Additional measures

The EHR and HIE data provided patient demographics such as age, race/ethnicity, and gender. We also created binary indicators for the 20 most common chronic conditions (20) and tobacco use (21), as well calculated the Charlson comorbidity index using diagnosis codes (22). We also counted each patient’s total number of emergency department encounters, primary care visits, and hospitalizations during the entire study period. For the secondary analysis of patients with SDH services need (see below), we limited the above measures to the data available prior to the

earliest date of documented need. This alternative method of measurement enabled comparisons of prior patient history between the different sources of information (i.e. structured and unstructured) available at the time of SDH services need documentation.

2.5 Analyses

We described the sources of information on SDH services, prevalence of need, and patient characteristics using frequencies and means. We compared patient characteristics in need of SDH services to those without an identified need using X^2 and t-tests. To assess the level of agreement between structured and unstructured information sources, we calculated kappa coefficients for documented need of any service by individual SDH services.

As a secondary analysis, we compared the characteristics of patients with SDH services need documented by structured and unstructured data. In a multinomial probit model, we examined the patient factors associated with having SDH service need documented by unstructured data, and by both structured and unstructured data, in comparison to need documented by structured data only. A best fitting model was identified using a backward elimination approach using the Bayesian Information Criterion with regression coefficients expressed as marginal effects.

3.1 RESULTS

Overall, 53% of patients (n=38,563) were in need of at least one SDH service during the study period (Table 1). Patients identified needing SDH services tended to be female, older, more medically complex, and higher utilizers of services. Specifically, the prevalence of nearly all chronic conditions examined was higher among those documented to be in need of services that addressed a SDH. For example, diagnoses of hypertension (53% vs. 35%), diabetes (32% vs.

14%), and depression (31% vs. 12%) were significantly higher in the in patients in need of SDH services (all $p < 0.05$). In addition, the patients with documented SDH services need also more often had a history of substance abuse (20% vs. 14%; $p < 0.0001$) and tobacco usage (29% vs. 20%; $p < 0.0001$).

Patients in need of social determinants of health services were identified from both structured and unstructured data (Table 2). Agreement on patients' need for SDH services between structured and unstructured data tended to be low. The overall prevalence of any SDH services need based on structured data only was 33% compared to 42% for unstructured data only ($\kappa = 0.35$). The most commonly identified SDH service was nutritional counseling service (35%) followed by behavioral health (21%), and social work (13%). All other social services were less prevalent. For the individual services that address SDH, the estimated prevalence of patient need was also highly variable between structured and unstructured data. For example, social work, medical-legal partnership, and pharmacist consultation were completely undocumented in structured data. For services documented by both structured and unstructured data, the agreement on individual patient need was generally very low.

The association between patient characteristics and the type of data documenting SDH service need (i.e. structured or unstructured) was determined in a multinomial probit model (Table 3). Patients with SDH services need identified only through unstructured data sources (Column A) differed significantly from patients with SDH services need identified only in structured data (the reference group). For example, patients with SDH services need documented by unstructured data tended to be more complex according to Charlson scores (marginal effect = 1.50; 95% CI=0.52, 2.49) than patients whose SDH service needs were documented by structured data.

In addition, patients with SDH services need documented by unstructured data only were associated with substance use (marginal effect = 2.92; 95%CI= 0.87, 4.98), tobacco use (3.36; 95%CI=1.65, 4.88), and the behavioral health conditions of depression (marginal effect = 8.70; 95%CI=7.17, 10.22) and schizophrenia (marginal effect = 7.42; 95%CI=2.82, 11.04). High prevalence chronic conditions of asthma and diabetes were negatively associated with SDH services need documented by unstructured data only. Moreover, the number of outpatient encounters was negatively associated with SDH service need being documented by unstructured data only (marginal effect = -0.78; 95% confidence interval (CI) = -0.85, -0.44). Stated alternatively, the more frequently a patient had outpatient encounters, the more likely structured data documenting that patient's SDH services need was present.

The multinomial probit model also described the association between patient characteristics and SDH services need documented by both structured and unstructured data (Table 3 column B) in comparison to patients with SDH services need identified only in structured data (the reference group). Several of these associations were in opposite directions than the association observed for patients with SDH services need identified only through unstructured data sources. For example, asthma (marginal effect = 6.89; 95%CI=4.33, 9.44) and diabetes (marginal effect = 9.82; 95%CI=8.05, 11.80) were both positively associated with patients' SDH services need being documented by both structured and unstructured data. Also, histories of substance abuse (marginal effect = -7.93; 95%CI=-10.18, -5.68) and tobacco use (marginal effect = -3.37; 95%CI = -5.11, -1.62) were negatively associated with having both structured and unstructured documentation of SDH services need compared to only structured documentation. Again, the more health care encounters observed, including outpatient visits and hospitalizations, the more likely the patient's SDH service need would be documented by structured data.

4. DISCUSSION

A US, urban safety-net population demonstrated a frequent need for the types of services that address SDH services need. The high prevalence of need indicates health care providers should develop strategies to support the demand for SDH services. At the same time, health care providers must also be attentive to the documentation, storage, and retrieval of information about SDH services delivered within their organization. Even within an integrated delivery system with a long history of EHR usage and HIE participation, identifying SDH services need required multiple information sources with both structured and unstructured data.

We estimated that more than half of the population needed SDH services. To the best of our knowledge, ours was the first attempt to quantify the need for SDH services in US primary care, although others have documented a similarly high prevalence of SDH associated risk factors (15, 23). This consistency suggests that the need for SDH services in the US is likely high - particularly for those providers caring for vulnerable patient populations. Problematically, physicians and medical care providers are not trained nor equipped to deliver such services (24), and widespread dissemination of the SDH paradigm arrived later in the US than in other parts of the world (25). As a further complication, structural barriers exist in the US as health care, public health, and social services providers operate as independent systems comprised of varying mixes of public, non-profit, and for-profit organizations each with different financing mechanisms, different goals, and no overarching coordination.

In light of the high prevalence of need for SDH services coupled with the trend towards greater accountability for population health, it is incumbent upon health care organizations in the US to create the infrastructure to make such services accessible. In the current study, SDH services were offered directly by the health care organization under the FQHC model; other health care organizations seeking to address patients' SDH could follow a similar approach by either expanding care team members with skills in addressing SDH (26) or co-locating services (11). Direct service offering has the benefits of facilitating patient access (27) and improved coordination (28). Also, the shift to value-based payment models in the US may make it financially feasible for health care organizations to offer such (predominately) preventative and health management services in expectation of avoiding downstream patient costs (29, 30). An alternative approach to addressing patients' SDH services need is through referral partnerships with community and social service organizations (31), such as required by the Patient Center Medical Home model (32). Even though the evidence on long-term patient health improvement is limited (33), external community and social service organizations remain an avenue for providers to improve health behaviors and address social needs. Additionally, addressing the differential funding streams between health care and social service is a challenge, but it is not insurmountable (11).

Second, these findings highlight the importance of unstructured data in health care organizational activities and population health services research. Several previous studies have demonstrated the value of unstructured data in such diverse settings as genetics research (34), identifying dialysis patients (35), attribution of diabetes patients (36) and risk prediction (37). Likewise, as these findings indicate, unstructured data are necessary to sufficiently document the services delivered by providers. The estimated prevalence of SDH need in this population was more than

1 in 2, however, if we used only structured data the estimate would have been 1 in 3. Organizations relying solely on structured data, e.g. procedures codes and appointments, will likely underestimate needed services that directly address SDH. Such underestimation could be problematic for health care administrators' resource allocation and operations planning, but undercounting patients in need is not the sole issue. Importantly, our findings suggest that unstructured data was associated with identifying the more complex and potentially more difficult patients, e.g. those with more comorbidities, mental health comorbidities, substance abuse, and potentially less ongoing primary care. Survey research suggests that patient complexity is associated with increased SDH service use (38) and unstructured data may be the means for providers to systematically identify the patients that have greater need for, or that would benefit more from SDH services.

While valuable, unstructured data are complex, difficult to extract, and variable across health IT systems (39). This work, then, joins a growing chorus of research calling for more flexible, effective and efficient tools and methods for extracting and standardizing unstructured data in the social, behavioral, and contextual domains (11, 17, 40). In primary care settings, such increased data accessibility could facilitate utilization risk prediction, point-of-care decision support, or automated referrals to services (41, 42). For health care organizations, better information on SDH needs could be leveraged to identify the appropriate social services organizations necessary to deliver patient centered services. For policymakers, aggregating such information would support policy evaluation, help assess health equity, and support identifying population's underlying health needs. Moreover, as researchers, policy makers, and technology experts consider data collection methods and data standards around SDH, the role of documenting SDH service delivery cannot be ignored. Understanding the importance of SDH in patient risk and

compliance are critical, but the better capture and standardization of SDH services will support effective referrals to a wide range of services in support of population health and prepare health care organizations for the day when more non-clinical services become reimbursable (42). The low capture rates for many types of SDH services across both structured and unstructured data suggests that we are only beginning to understand the demand for specific SDH services. Better characterization of this demand will come with more complete and accurate structured and unstructured data capture.

4.2 Limitations

The generalizability of these findings may be limited in terms of patient population and differences in documentation practices across providers. Additionally, several of SDH services included in this study were not offered over the course of the entire study period. Importantly, our primary outcome of SDH services need only reflects a partial list of all the potential SDH issues that influence patient health (e.g. government policies and environmental conditions); moreover, our measure was not assessed against a gold standard. We did not assess the quality of either the structured or unstructured data, which can be highly variable. Further, we only examined data sources currently used in the FQHC's clinical care setting. Information systems such as environmental monitoring, census data, educational information systems, and other public health information systems are critically important to understanding SDH, but were not considered in the current analysis. Lastly, we only assessed SDH services need; patients referred to services may have refused the referral or opted not to keep the appointment.

5. CONCLUSION

The need for SDH services among safety-net populations is high. Identifying patients in need of such services requires multiple data sources with structured and unstructured data. Health care organizations preparing for population health activities and payment reform need to identify approaches for both ensuring patient access to SDH services as well as documenting SDH services delivery.

Authors' Contributions

Study conception: JV, SG, PH, NM. Data collection and analysis: JV, SG, DH, NM.

Interpretation and writing: JV, SG, DH, PH, NM.

Statement on Conflicts of Interest

The authors have no conflicts of interest to declare.

Summary Table

- A majority of patients may be in need of the services that address the social determinants of health. As a result, health care organizations will need to create the infrastructure to make such services accessible to patients.
- Measuring service need from structured data is insufficient; unstructured data are necessary to give a complete picture of patient need.

Acknowledgements

This work was supported by the Robert Wood Johnson Foundation through the Systems for Action National Coordinating Center, ID 73485. The authors thank Jennifer Williams and the Regenstrief Institute Data Core.

References

1. Kocher R, Emanuel EJ, DeParle NM. The affordable care act and the future of clinical medicine: The opportunities and challenges. *Annals of Internal Medicine*. 2010;153(8):536-9.
2. Livingston S. Paying for population health 2017 [cited 2017 2017-04-12]. Available from: <http://www.modernhealthcare.com/article/20170211/MAGAZINE/302119984>.
3. Fitzpatrick T, Rosella LC, Calzavara A, Petch J, Pinto AD, Manson H, et al. Looking Beyond Income and Education: Socioeconomic Status Gradients Among Future High-Cost Users of Health Care. *American Journal of Preventive Medicine*. 2015;49:161-71.
4. Williams DR, Costa MV, Odunlami AO, Mohammed SA. Moving upstream: how interventions that address the social determinants of health can improve health and reduce disparities. *Journal of public health management and practice : JPHMP*. 2008;14 Suppl:S8-17.
5. Commission on Social Determinants of Health. Closing the gap in a generation: Health equity through action on the social determinants of health. Geneva, Switzerland: World Health Organization, 2008.
6. Solberg LI. Theory vs Practice: Should Primary Care Practice Take on Social Determinants of Health Now? No. *Ann Fam Med*. 2016;14(2):102-3.
7. Stanhope V, Videka L, Thorning H, McKay M. Moving toward integrated health: an opportunity for social work. *Soc Work Health Care*. 2015;54(5):383-407.
8. Hedlund N, Risendal BC, Pauls H, Valverde PA, Whitley E, Esparza A, et al. Dissemination of Patient Navigation Programs Across the United States. *Journal of Public Health Management and Practice*. 2014;20:E15-E24.

9. Sandel M, Hansen M, Kahn R, Lawton E, Paul E, Parker V, et al. Medical-Legal Partnerships: Transforming Primary Care By Addressing The Legal Needs Of Vulnerable Populations. *Health Affairs*. 2010;29:1697-705.
10. Davis M, Balasubramanian BA, Waller E, Miller BF, Green LA, Cohen DJ. Integrating Behavioral and Physical Health Care in the Real World: Early Lessons from Advancing Care Together. *The Journal of the American Board of Family Medicine*. 2013;26(5):588-602.
11. Gottlieb L, Sandel M, Adler NE. Collecting and applying data on social determinants of health in health care settings. *JAMA Internal Medicine*. 2013;173(11):1017-20.
12. DeVoe JE, Bazemore AW, Cottrell EK, Likumahuwa-Ackman S, Grandmont J, Spach N, et al. Perspectives in Primary Care: A Conceptual Framework and Path for Integrating Social Determinants of Health Into Primary Care Practice. *The Annals of Family Medicine*. 2016;14:104-8.
13. Lewis JH, Whelihan K, Navarro I, Boyle KR, SDH Card Study Implementation Team. Community health center provider ability to identify, treat and account for the social determinants of health: a card study. *BMC Family Practice*. 2016;17:121.
14. Garg A, Butz AM, Dworkin PH, Lewis RA, Serwint JR. Screening for Basic Social Needs at a Medical Home for Low-Income Children. *Clinical Pediatrics*. 2009;48:32-6.
15. Page-Reeves J, Kaufman W, Bleecker M, Norris J, McCalmont K, Ianakieva V, et al. Addressing Social Determinants of Health in a Clinic Setting: The WellRx Pilot in Albuquerque, New Mexico. *J Am Board Fam Med*. 2016;29(3):414-8.
16. Pinto AD, Glattstein-Young G, Mohamed A, Bloch G, Leung F-H, Glazier RH. Building a Foundation to Reduce Health Inequities: Routine Collection of Sociodemographic Data in Primary Care. *Journal of the American Board of Family Medicine : JABFM*. 2016;29:348-55.

17. Institute of Medicine. Capturing social and behavioral domains in electronic health records: Phase 2. Washington, DC2014.
18. Gottlieb LM, Tirozzi KJ, Manchanda R, Burns AR, Sandel MT. Moving Electronic Medical Records Upstream: Incorporating Social Determinants of Health. *American Journal of Preventive Medicine*. 2015;48:215-8.
19. Glasgow RE, Kaplan RM, Ockene JK, Fisher EB, Emmons KM. Patient-reported measures of psychosocial issues and health behavior should be added to electronic health records. *Health Aff (Millwood)*. 2012;31(3):497-504.
20. Goodman RA, Posner SF, Huang ES, Parekh AK, Koh HK. Defining and measuring chronic conditions: imperatives for research, policy, program, and practice. *Preventing chronic disease*. 2013;10:E66.
21. Wiley LK, Shah A, Xu H, Bush WS. ICD-9 tobacco use codes are effective identifiers of smoking status. *Journal of the American Medical Informatics Association : JAMIA*. 2013;20:652-8.
22. Charlson ME, Charlson RE, Peterson JC, Marinopoulos SS, Briggs WM, Hollenberg JP. The Charlson comorbidity index is adapted to predict costs of chronic disease in primary care patients. *Journal of Clinical Epidemiology*. 2008;61:1234-40.
23. Bolen SD, Sage P, Perzynski AT, Stange KC. No moment wasted: the primary-care visit for adults with diabetes and low socio-economic status. *Prim Health Care Res Dev*. 2016;17(1):18-32.
24. Loeb DF, Binswanger IA, Candrian C, Bayliss EA. Primary care physician insights into a typology of the complex patient in primary care. *Ann Fam Med*. 2015;13(5):451-5.

25. Raphael D. Social determinants of health: present status, unanswered questions, and future directions. *International journal of health services : planning, administration, evaluation.* 2006;36(4):651-77.
26. Barr VJ, Robinson S, Marin-Link B, Underhill L, Dotts A, Ravensdale D, et al. The Expanded Chronic Care Model: An Integration of Concepts and Strategies from Population Health Promotion and the Chronic Care Model. *Healthcare Quarterly.* 2003;7(1):73-82.
27. Garg A, Jack B, Zuckerman B, MG M, B Z, MG M, et al. Addressing the Social Determinants of Health Within the Patient-Centered Medical Home. *JAMA.* 2013;309:2001.
28. Berry LL, Rock BL, Smith Houskamp B, Brueggeman J, Tucker L. Care Coordination for Patients With Complex Health Profiles in Inpatient and Outpatient Settings. *Mayo Clinic Proceedings.* 2013;88(2):184-94.
29. Bachrach D, Pfister H, Wallis K, Lipson M. ADDRESSING PATIENTS' SOCIAL NEEDS An Emerging Business Case for Provider Investment. New York, NY: Commonwealth Fund, 2014.
30. DeMilto L, M N. Using Social Determinants of Health Data to Improve Health Care and Health: A Learning Report Princeton, NJ: Robert Wood Johnson Foundation,, 2016.
31. Blewett LA, Owen RA. Accountable Care for the Poor and Underserved: Minnesota's Hennepin Health Model. *American Journal of Public Health.* 2015;105(4):622-4.
32. Jackson GL, Powers BJ, Chatterjee R, Prvu Bettger J, Kemper AR, Hasselblad V, et al. The Patient-Centered Medical Home. *Annals of Internal Medicine.* 2013;158:169.
33. Porterfield DS, Hinnant LW, Kane H, Horne J, McAleer K, Roussel A. Linkages between clinical practices and community organizations for prevention: a literature review and environmental scan. *Am J Prev Med.* 2012;42(6 Suppl 2):S163-71.

34. Kho AN, Pacheco JA, Peissig PL, Rasmussen L, Newton KM, Weston N, et al. Electronic Medical Records for Genetic Research: Results of the eMERGE Consortium. *Science Translational Medicine*. 2011;3(79):79re1-re1.
35. Abhyankar S, Demner-Fushman D, Callaghan FM, McDonald CJ. Combining structured and unstructured data to identify a cohort of ICU patients who received dialysis. *Journal of the American Medical Informatics Association*. 2014;21(5):801-7.
36. West SL, Blake C, Liu Z, McKoy JN, Oertel MD, Carey TS. Reflections on the use of electronic health record data for clinical research. *Health Informatics Journal*. 2009;15(2):108-21.
37. Wang Y, Ng K, Byrd RJ, Hu J, Ebadollahi S, Daar Z, et al. Early Detection of Heart Failure with Varying Prediction Windows by Structured and Unstructured Data in Electronic Health Records(). Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual Conference. 2015;2015:2530-3.
38. Paino M, Aletraris L, Roman P. The Relationship Between Client Characteristics and Wraparound Services in Substance Use Disorder Treatment Centers. *JOURNAL OF STUDIES ON ALCOHOL AND DRUGS / JANUARY 2016 Stud Alcohol Drugs*. 2016;77:160-9.
39. Capurro D, Yetisgen M, van Eaton E, Black R, Tarczy-Hornoch P. Availability of structured and unstructured clinical data for comparative effectiveness research and quality improvement: a multisite assessment. *EGEMS (Wash DC)*. 2014;2(1):1079.
40. Estabrooks PA, Boyle M, Emmons KM, Glasgow RE, Hesse BW, Kaplan RM, et al. Harmonized patient-reported data elements in the electronic health record: supporting

meaningful use by primary care action on health behaviors and key psychosocial factors. *J Am Med Inform Assoc.* 2012;19(4):575-82.

41. Fiscella K, Tancredi D. Socioeconomic status and coronary heart disease risk prediction. *JAMA.* 2008;300(22):2666-8.

42. Vest JR, Harle CA, Schleyer T, Dixon BE, Grannis SJ, Halverson PK, et al. Getting From Here to There: Health IT Needs for Population Health. *American Journal of Managed Care.* 2016;22(12):827-9.

Table 1. Comparison of an adult safety-net primary care population with and without a documented need for social determinants of health (SDH) services, 2012-2016.

	Total	In need of at least 1 SDH service	No SDH service needs	p
	n=73,085	n= 38,563	n= 34,522	
	%			
Age (mean, sd)	44.6 (15.4)	45.5 (14.7)	43.6 (16.1)	<0.0001
Male gender	34.6	31.4	38.1	<0.0001
Race / ethnicity				
White, non-Hispanic	25.8	28.6	22.7	<0.0001
African American, non-Hispanic	40.1	40.6	39.6	0.007
Hispanic	18.5	18.3	18.7	0.165
Other	5.4	4.9	5.9	<0.0001
Diagnoses				
Hypertension	44.6	53.2	35.0	<0.0001
Congestive heart failure	5.2	6.5	3.8	<0.0001
Coronary artery disease	7.5	9.1	5.8	<0.0001
Cardiac arrhythmias	8.6	10.0	7.1	<0.0001
Hyperlipidemia	21.5	27.0	15.3	<0.0001
Stroke	4.1	5.0	3.2	<0.0001
Arthritis	10.9	13.4	8.1	<0.0001
Asthma	9.0	12.0	5.7	<0.0001
Cancer	8.7	9.5	7.8	<0.0001
COPD	11.0	14.4	7.2	<0.0001
Depression	22.0	30.7	12.2	<0.0001
Diabetes	23.4	32.0	13.7	<0.0001
Hepatitis	4.3	5.1	3.5	<0.0001
HIV	1.0	0.9	1.1	0.0282
Schizophrenia	3.5	4.3	2.6	<0.0001
Substance abuse	17.3	20.3	13.8	<0.0001
Charlson index score (mean, sd)	0.8 (1.3)	1.1 (1.4)	0.6 (1.1)	<0.0001
Tobacco use	24.5	28.8	19.6	<0.0001
Hospitalizations (mean, sd)	1.4 (5.2)	34.1 (30.7)	1.0 (2.9)	<0.0001
ED visits (mean, sd)	7.9 (20.9)	9.3 (24.5)	6.4 (15.7)	<0.0001
Outpatient (mean, sd)	28.2 (28.7)	1.6 (6.5)	21.5 (24.7)	<0.0001

Table 2. Prevalence of the need for social determinants of health services (SDH) in an adult safety-net primary care population, by source of information, 2012-2016.

Service	Documented Need	Structured data only	Unstructured data only	Kappa
Any SDH service	52.8	33.2	41.8	0.35
Social work	12.5	0.0	12.5	-- ¹
Behavioral health	20.5	6.8	16.7	0.17
Nutritional counseling	35.1	24.2	26.6	0.49
Respiratory therapy	4.6	3.0	3.3	0.53
Financial counseling	6.2	5.1	1.2	0.01
Medical-legal partnership	0.5	0.0	0.5	--
Patient navigation	0.5	0.5	0.0	--
Pharmacist consultation	0.2	0.0	0.2	--

¹Kappa could not be calculated

Table 3. Association between adult patient characteristics and social determinants of health (SDH) service need documented by unstructured data only, and both structured and unstructured data, compared to structured data sources only.

	(A)		(B)	
	SDH services need documented by unstructured data only compared to structured data sources		SDH services need documented by structured & unstructured data compared to structured data sources	
	Marginal effect (95% CI)	p	Marginal effect (95% CI)	p
	n= 14,287		n= 16,235	
Age	0.07 (0.02, 0.11)	0.003	0.14 (0.09, 0.18)	<0.001
Male	-0.50 (-1.73, 0.72)	0.420	-6.63 (-7.93, -5.33)	<0.001
Diagnoses				
Hyperlipidemia	0.75 (-0.96, 2.45)	0.393	3.03 (1.26, 4.79)	0.001
Arthritis	4.21 (1.70, 6.72)	<0.001	0.41 (-2.21, 3.03)	0.761
Asthma	-5.90 (-8.40, -3.40)	<0.001	6.89 (4.33, 9.44)	<0.001
Depression	8.70 (7.17, 10.22)	<0.001	3.37 (1.72, 5.01)	<0.001
Diabetes	-11.6 (-13.44, -9.76)	<0.001	9.92 (8.05, 11.80)	<0.001
HIV	-4.56 (-11.44, 2.33)	0.195	-4.94 (-12.11, 2.23)	0.177
Schizophrenia	7.43 (3.82, 11.04)	<0.001	2.73 (-1.23, 6.69)	0.176
Substance abuse	2.92 (0.87, 4.98)	0.005	-7.93 (-10.18, -5.68)	<0.001
Charlson index	1.50 (0.52, 2.49)	0.003	1.75 (0.72, 2.77)	0.001
Tobacco use	3.26 (1.65, 4.88)	<0.001	-3.37 (-5.11, -1.62)	<0.001
Hospitalizations	0.12 (-0.31, 0.55)	0.590	-1.42 (-1.94, -0.90)	<0.001
Outpatient visits	-0.78 (-0.85, -0.71)	<0.001	-0.50 (-0.57, -0.44)	<0.001

¹Compared to structure data sources only