

Completeness of Social and Behavioral Determinants of Health in Electronic Health Records: A case study on the Patient-Provided Information from a minority cohort with sexually transmitted diseases

Bhavani Singh Agnikula Kshatriya

Mayo Clinic Rochester

Joyce E Balls-Berry

Washington University in St Louis

William D Freeman

Mayo Clinic Florida

Rui Zhang

University of Minnesota

Yanshan Wang (✉ wang.yanshan@mayo.edu)

Mayo Clinic Rochester <https://orcid.org/0000-0003-4433-7839>

Research

Keywords: social determinants of health, electronic health records, completeness, health equity, minority, sexually transmitted disease

Posted Date: December 10th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-123744/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Introduction

Racially and ethnically diverse minorities often experience the disease burden of sexually transmitted infections or diseases (STD) more often than their White counterparts. Yet, little is known about the connection of STD systematic discrimination, racism, and social and behavioral determinants. Plus, little to no details exist related to how this information is recorded in their Electronic Health Records (EHRs). The objective of this study is to assess the completeness of social and behavioral determinants of health (SDOH) data in the EHRs of a minority cohort with STD.

Materials and Methods

2,993 minority patients diagnosed with a STD at the Mayo Clinic were identified for this study. A natural language processing (NLP) algorithm was applied on the Patient-Provided Information (PPI) in their EHRs to extract SDOH information in six domains that are associated with STD, namely alcohol use, substance use, sexual activity, sexual orientation, housing status, and employment status. The completeness of SDOH was assessed in terms of documentation, breadth, and density.

Results

Our study indicates that nearly half of 2,993 patients did not have SDOH-related PPI records in their EHRs whereas the patients who had SDOH-related PPI records had well-documented records for five out of six SDOH domains, including alcohol use, substance use, sexual activity, housing status, and employment status, except for sexual orientation. A total of 1,504 patients had PPI in their EHRs for at least one of six SDOH domains, which is about 50.3% of the study cohort. Most SDOH domains have a short time span of 1 year, with up to 18 years of record data. Our analysis also indicated that education and age have a significant impact on the recording of SDOH-related PPI records. Patients that are female, older, and higher educated tend to have more SDOH information available in their records.

Discussion and Conclusion

We assessed the completeness of SDOH information recorded in the PPI from patients' EHRs. Due to large amounts of missing SDOH information in the PPI, future research is needed to integrate accurate and robust SDOH related data for downstream research and the impact of systematic discrimination on how this information is collected and interpreted in the EHRs.

Introduction

The perspective that population health outcomes are notably influenced by complex, integrated, and overlapping social structures and economic systems is gaining recognition among scientists and public health professionals. Those social structures economic systems have been widely recognized as social determinants of health (SDOH). According to World Health Organization (WHO), SDOH are the conditions in the places where people live, learn, work and play that affect health risks and outcomes in a wide range.[1] SDOH influences people's access to adequate diet, educational and career opportunities, healthy environmental conditions, and medical care, either in a positive or negative way.[1] They are identified to be the underlying factors for a wide-range of diseases and difficulty in care engagement that are linked directly to health disparities.[2–5] SDOH is an important factor to be considered in order to reduce expenditures due to hospital readmissions.[6 7] Identifying SDOH is key to increasing health equity and quality of life for patients historically impacted by SDOH.[8]

Addressing SDOH improves healthcare drastically.[9] Thus various research studies recommend the importance of understanding SDOH and the need to incorporate interventions to address problems not only at personal level but also at

population level by making policy changes to bring health equity.[10–13] The exploration of a socioecological perspectives with a concentration on SDOH is important to not only improve health but also health equity by recommending policies that affect global population,[14] so that people receive equitable and equal opportunities to improve their health.[1] Though a wide range of studies have been conducted using SDOH data to improve public health and economy,[15] completeness of the SDOH data used to perform these analyses remained unexplored.

With the wide and rapid adoption of electronic health record (EHR) systems in healthcare organizations, there is increased emphasis on use of electronic health records (EHRs) to document SDOH information. Healthcare data collection has been growing rapidly with the introduction of EHRs. Healthcare data has gained the utmost importance in many areas, such as research and payment to perform significant analysis. Patient provided information (PPI) is integrated with the Mayo Clinic EHR system that gathers information from patients systematically and electronically during many aspects of the clinical practice, such as patient care, education, research, and business office functions .[16] PPI involves an intelligent system that automatically determines what questionnaire forms/information are needed and either mails paper forms or electronically sends the forms in the patient portal or on a tablet to be completed online or during physical clinical visit. PPI includes valuable information, for example, family history, past medical history, allergies, depression screening, etc., and most importantly lots of information related to SDOH. PPI is a vital part of healthcare data that is critical not only for patient-centered care but also to identify implicit SDOH factors that determine the health of a community/population.[17] However, the collection of patient provided data may differ from traditional methods of collecting data by health providers, with respect to response rate and data quality. Thus, it becomes crucial to verify the data quality of SDOH information captured in PPI from the EHR, especially the completeness of such information, as incomplete data might make the data useless for diagnosis/research or lead to improper/inaccurate analysis outcomes. Assessing the completeness of SDOH in EHRs is a fundamental factor for its primary or secondary usage yet has been rarely studied.

This study aims to assess the completeness of SDOH information collected in PPI from EHRs by using a case study of a minority cohort with sexually transmitted diseases (STD). Racially and ethnically diverse communities experience systematic discrimination which results in SDOH. Furthermore, this is exacerbated by the lack of prevention and treatment options in the communities. The disease burden is not explained by intrapersonal or community level concern but is often deep-rooted in systematic discrimination, lack of cultural awareness, and lack of access to for prevention and treatment [18]. The purpose of this study is to recommend the SDOH research to incorporate the data completeness factor understanding its importance while performing the analysis. This study has the potential to help the research community focusing on SDOH to improve health equity and public health, as well as to assist the analysis of clinical management for certain diseases such as STD.[19]

Methods

In this case study, the cohort is a minority cohort (≥ 18 -year-old), including African American or Black, Hispanic, Asian, Native American, Pacific Islander, and non-White, who were seen at Mayo Clinic and diagnosed with STD from August 14, 1987 through May 22, 2015. The STD's included genital herpes [International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) codes 054.1, 054.10, 054.13, 054.19], genital warts (ICD-9-CM code 078.11), Chlamydia (ICD-9-CM codes 078.88, 079.88, 079.98, 099.41, 099.53–099.55), gonorrhea (ICD-9-CM codes 098.0, 098.1, 098.10- 098.19, 098.2, 098.30-098.34), syphilis (ICD-9- CM codes 091.0, 095.4, 095.8), and epididymitis/orchitis (ICD-9-CM codes 604, 604.0, 604.9, 098.13, 098.33).[20 21] We used the i2b2 informatics tool to retrieve the minority population cohort. The i2B2 tool is a scalable informatics framework that stores and queries the standardized clinical data from the Mayo Clinical Data Warehouse (CDW). We identified a total of 2,993 patients that had been diagnosed with an STD at Mayo Clinic. Then we retrieved the PPI of this cohort in the same time range. Most PPI is in the form of questionnaires that have been given to patients during their visits. This study was approved by the Mayo Clinic institutional review board (IRB) for human subject research (IRB #20–000908).

Out of a total of 1,647,061 records being retrieved, 3,329 unique questions were found. We defined six SDOH domains that are related to STD, which include alcohol use, substance use, sexual activity, sexual orientation, housing status, and employment status, as shown in Table 1. We developed a rule-based natural language processing (NLP) to identify questions related to the six concerned SDOH domains. The NLP algorithm used a list of keywords to identify the questions in each SDOH topic. The details of this algorithm can be found in the supplemental materials. With the NLP algorithm, we identified and extracted a total of 125 SDOH related questions. Then we manually excluded 7 unique questions as they were not relevant to the SDOH domains, such as “Are you interested in more information about safety (seat belts, smoke detectors, firearms)?” etc. A final dataset containing 118 unique SDOH related questions (42 for alcohol use, 50 for substance use, 16 for sexual activity, 2 for sexual orientation, 2 for housing status, and 6 for employment status) were used in this study.

Table 1
PPI examples of SDOH domains

SDOH Domains	Examples of PPI
alcohol use	How often do you have six or more drinks on one occasion? How often do you have a drink containing alcohol?
substance abuse	Do you have exposure to medications, substances, foods, dusts, and animals to which you are allergic to or have unpleasant side effects to? Have you taken steroid or cortisone-type drugs within the last year?
sexual orientation	Please describe other sexual orientation
sexual activity	Are you sexually active? Do you use condoms?
housing status	With whom do you live?
employment status	What is your current employment status (check all that apply)? What is your current work status?

Since all the six SDOH domains are associated with STD’s, we tried to assess the completeness of STD patients’ records in terms of 1) documentation, 2) breadth, 3) density, followed by Nicole et al’s study on the completeness of EHRs for secondary use, as defined below.[22]

Documentation

Assess whether records of the six SDOH domains are well-documented in patients’ PPI records. This assessment tells us if the SDOH related PPI are well-documented in the entire EHRs.

Breadth

Assess whether a patient’s records contain all the six SDOH domains. Analyzing breadth of SDOH data can represent the availability of SDOH domains in each patient’s records.

Density

Assess the frequency of SDOH data points over time. This assessment evaluates the time span of the SDOH domains, for example, how many patients had SDOH questions recorded for a time span of 2 years.

Since many SDOH information are being recorded with no answers or invalid answers, we also evaluated valid SDOH records (e.g., SDOH questions with valid answers). In addition, we assessed the association between SDOH domains and patients' gender, age, education level, and living area using t-test. The living area is defined as urban, suburban, and rural based on patients' zip codes. This assessment shows if any of the demographics have an influence on the completeness of the SDOH information in the EHRs.

Results

We first demonstrate the demographics of the minority cohort in Table 2. Male and female patients are reasonably balanced in the study cohort, with slightly more female patients. The majority minority in our cohort is Black or African American. The cohort is relatively young with 34% from 18–34 years old due to the study disease. Education status is unknown for more than half of patients (56.6%). For those with education status, most patients (31.8%) are college-educated (e.g., Some College or 2 year degree, Four year college graduate, or Post graduate studies). 36.8% of patients lives in suburban area. The education and living area are consistent with regional statistics in the Rochester area[23].

Table 2
Demographics of the study cohort

Demographic	Number (n = 2993)	Percentage
Gender		
Male	1442	48%
Female	1551	52%
Race		
American Indian/Alaskan Native	116	4%
Asian	309	10%
Black or African American	1140	38%
Native Hawaii/Pacific Islander	27	1%
White	545	18%
Choose Not to Disclose	15	1%
Other	775	26%
Unable to Provide	1	0
Unknown	65	2%
Ethnicity		
Hispanic or Latino	1032	34%
Not Hispanic or Latino	1686	56%
Choose Not to Disclose	23	1%
Unable to Provide	2	0
Unknown	250	8%
Age		
18–34 years old	1029	34%
35–44 years old	745	25%
45–54 years old	504	17%
55–64 years old	350	12%
65–74 years old	222	7%
75–84 years old	104	3%
>= 85 years old	39	1%
Education		
Some College or 2 year degree	464	15.5%
Four year college graduate	242	8.1%
Post graduate studies	245	8.2%
Some high school, but did not graduate	222	7.4%

Demographic	Number (n = 2993)	Percentage
8th grade or less	120	4%
High School graduate or GED	435	14.5%
Unknown	1694	56.6%
Living area		
Urban	283	9.5%
Suburban	1100	36.8%
Rural	11	0.3%
Unknown	1599	53.5%

Documentation

Tables 3 shows the documentation completeness of six SDOH domains. First, the majority of the SDOH-related PPI are on substance use (42.6%) and a very small portion on sexual orientation (0.1%). This is due to the difference of number of questions available in each SDOH domain. Second, the number of valid records is actually acceptable for most SDOH domains. We can see that the SDOH domain with most valid records is housing status with 90.8% valid rate whereas the least SDOH domain is substance use with only 65% valid rate.

Table 3
Documentation Completeness of SDOH domains in the PPI from EHRs

SDOH Domains	Total number of records for each SDOH domain/Total number of records (%)	Total number of valid records for each SDOH domain/Total number of records for each SDOH domain (%)
Alcohol use	(26055/101113) - 25.8	(21808/26055) - 83.7
Substance use	(43065/101113) - 42.6	(28007/43065) - 65
Sexual activity	(18760/101113) - 18.6	(14001/18760) - 74.6
Sexual orientation	(142/101113) - 0.1	(100/142) - 70.4
Housing status	(5824/101113) - 5.8	(5286/5824) - 90.8
Employment status	(7267/101113)- 7.2	(6230/7267) - 85.7

Breadth

A total of 1,504 patients had PPI in their EHRs for at least one of six SDOH domains, which is about 50.3% of the study cohort. We examined how many SDOH questions were being recorded and answered for each patient. We found that no patients had all 118 unique SDOH questions with valid answers in their EHRs. The maximum number of unique SDOH questions with valid answers was 64 by only 2 patients, whereas the minimum number of unique SDOH questions with valid answers was 17. Tables S1 and S2 in the supplemental materials list the number of patients for a given range of being recorded and being answered SDOH questions, respectively. We can see that only a small number of patients had valid SDOH records (i.e., with valid answers).

Specific to each SDOH domain, all the 1,504 patients had recorded SDOH questions in alcohol use followed by those in substance use, whereas only 85 patients had recorded those in sexual orientation. Interestingly, out of the patients who had recorded SDOH questions, most of them had answered most questions in their records (i.e., valid records).

Table 4
Breadth Completeness of SDOH domains in the PPI from EHRs

SDOH domain	Number of patients with records for each SDOH domain /Total number of patients (%)	Number of patients with valid records for each SDOH domain/ Number of patients with records for each SDOH domain (%)
Alcohol use	(1504/1504) - 100	(1490/1504) - 99.1
Substance use	(1503/1504)- 99.9	(1502/1503) - 99.9
Sexual activity	(1490/1504) - 99.1	(1484/1490) - 99.6
Sexual orientation	(85/1504) - 5.7%	(65/85) - 76.5%
Housing status	(1450/1504) - 96.4%	(1401/1450) - 96.6%
Employment status	(1411/1504) - 93.8%	(1388/1411) - 98.4%

Density

Tables S3-S4 in the supplemental materials show the density completeness in terms of time span of the SDOH-related PPI in the EHRs. We find that majority of patients only have a short time span of SDOH records, despite the time span may increase up to 18 years. Among six SDOH domains, sexual orientation has only 6 years of records.

Association

The association between six SDOH domains and patients' demographics is shown in Table 5. Since gender and biological sex were identical for all except one patient, we just used the gender information. We find that all SDOH domains have significant association with gender, age, education level, and living area, except sexual orientation and living area. Table S5, S6, S7, and S8 in the supplemental materials list the number of patients for each SDOH domain in terms of gender, age, education level, and living area, respectively. Female patients have more SDOH information recorded in their EHRs than male patients. This is consistent with literature [22]. Table S6 shows that majority of the documentation was found among the age group of 18–34, with less percentage among the age group 75–84 and > 85. From Table S7, we see that majority of the documentation was found among some college or 2-year degree group, whereas the least documentation was found among the 8th grade or less education level patients. Since most patients in our cohort lived in suburban area, Table S8 shows a limited comparison for different living areas. Substance abuse was well-documented in most living areas, whereas sexual orientation was the least documented SDOH topic. This is consistent with the documentation completeness results. However, the impact of geographical location on SDOH documentation must be further explored since the population size among urban, sub-urban, and rural areas varied greatly in our study. In conclusion, patients that are female, older, and higher educated tend to have more SDOH information available in their records.

Table 5
Association between six SDOH domains and demographics

	Gender	Age	Education level	Living area
Alcohol use	statistic=-107.80, p-value = 0.0	Statistic = 498.04, p-value = 0.0	statistic = 155.20, p-value = 0.0	statistic = 34.19, p-value = 1.1390022855597689e-253
Substance use	statistic=-66.21, p-value = 0.0	statistic = 630.51, p-value = 0.0	statistic = 234.17, p-value = 0.0	statistic = 68.25, p-value = 0.0
Sexual orientation	Statistic=-18.32 p-value = 6.189142045686077e-50	Statistic = 49.19, p-value = 1.2938927300359609e-155	statistic = 8.29, p-value = 1.429410186851639e-15	statistic = 0.89, p-value = 0.3735677772834919
Sexual activity	Statistic=-69.71 p-value = 0.0	Statistic = 417.29, p-value = 0.0	statistic = 145.86, p-value = 0.0	statistic = 33.98, p-value = 1.1215742367200186e-249
Housing status	Statistic=-63.25, p-value = 0.0	Statistic = 244.60, p-value = 0.0	statistic = 75.04, p-value = 0.0	statistic = 12.67, p-value = 1.4213069000654502e-36
Employment status	statistic=-59.39, p-value = 0.0	Statistic = 292.29, p-value = 0.0	statistic = 81.78, p-value = 0.0	statistic = 15.94, p-value = 8.791608332045375e-57

Discussion

This is the first study, to the best of our knowledge, that assesses the completeness of SDOH records in the PPI from EHRs. We used a minority cohort of 2,993 patients with STDs. Our results show that nearly half of the patients did not have SDOH-related PPI recorded in their EHRs, although the SDOH information is vital to the STD. There are many underlying reasons. One of them might be due to the varied clinical guidelines and questionnaires during a clinical visit overtime. The lack of SDOH-related PPI in EHRs indicates an additional layer of data completeness evaluation to be considered prior to using PPI in downstream analysis and research for public health. But amongst the patients who had recorded SDOH-related PPI, five out of six SDOH domains were complete, except for sexual orientation. This result indicates that alternative data resources in EHR should be considered for extraction of sexual orientation (e.g., clinical notes) and more sophisticated methods should be used (e.g., NLP).

SDOH-related PPI records were higher recorded among females than males. It might be due to that females usually have more health issues compared to males, leading to more hospital visits with a higher number of EHRs, which is consistent with the literature.[24] Most of the SDOH-related PPI were recorded and answered by well-educated people. It might be because of a greater awareness of the importance of PPI medically and socially, whereas this awareness might be less in the not well-educated population. Thus, it is important to find the means to educate the importance of PPI among not-well-educated patients, as well as the rural population, as the number of patients who recorded PPI in rural areas was also significantly lower than urban and suburban areas. Another major factor for no records of PPI might be the quantity of PPI questionnaires and/or the way of displaying the questionnaires for the patients, which might make them feel tiresome and uninterested in filling it out during the waiting time before their interaction with the physician. This can be specifically among elderly patients.

This study provides a useful conclusion that the SDOH information stored in structured PPI from EHRs has significant information gaps, which makes it non-straightforward for secondary use in epidemiological research studies in a population. The information gap might also be due to various formats of the existing SDOH information in different resources in EHRs. There is a need to have universal data standards to appropriately capture, retrieve and analyze SDOH data.[25] It is also important to design methods like NLP, ML to collect SDOH data from various other sources, such as clinical notes of EHR etc.

This study has several limitations corresponding to data and the level assessment of SDOH domains. We have used the PPI data from a single institution. Hence, there is a possibility for inconsistency of the results if extended to generalize on PPI data from other institutions. The assessment is performed at the topic level, i.e., alcohol use related, substance use related etc., but we did not dive into the intricacies of each question, i.e., what type of questions are considered less by the patients, why and so on. We were unable to examine the link to systematic discrimination and SDOH for this study. It is documented that some efforts are in place to address the STD health disparities for Blacks but this limited effort has not been evaluated in all minority communities [18].

For the future work, this study presents the scope of performing research in various areas, such as developing natural language processing (NLP) based approach to derive sexual orientation related PPI information from clinical notes, as the availability of structured data is very low. This approach of evaluating data completeness can be extended to other diseases PPI, such as depression, instead of only the current SDOH domains evaluated in this study. The SDOH outcomes could help in designing policies/strategies to remove barriers of health equity and provide better patient care for individuals. [26]

Conclusion

SDOH plays a vital role in the research studies to determine ways to bring health equity among the population. Our study indicates that nearly half of 2,993 patients did not have SDOH-related PPI records in their EHRs whereas the patients who had SDOH-related PPI records had well-documented records for five out of six SDOH domains, including alcohol use, substance use, sexual activity, housing status, and employment status, except for sexual orientation. A total of 1,504 patients had PPI in their EHRs for at least one of six SDOH domains, which is about 50.3% of the study cohort. Most SDOH domains have a short time span of 1 year, with up to 18 years of record data. Our analysis also indicated that education and age have a significant impact on the recording of SDOH-related PPI records. Patients that are female, older, and higher educated tend to have more SDOH information available in their records. The impact of geographical location on SDOH documentation has to be further explored as our study cohort is imbalanced among urban, sub-urban, and rural areas. It is also important to consider the ethnoracial factors linked directly to systematic discrimination and lack of healthcare. Due to large amounts of missing SDOH information in the PPI, future research is needed to collect and integrate accurate and robust SDOH related data for downstream research from other EHR resources, such as the unstructured clinical notes, to have a complete picture of SDOH variables for patients.

Declarations

Ethics approval and consent to participate

This study was a retrospective study of existing records. The study and a waiver of informed consent were approved by Mayo Clinic Institutional Review Board in accordance with 45 CFR 46.116 (Approval #20-000908).

Consent to publish

Not applicable; the manuscript does not contain individual level of data.

Availability of data and materials

The EHR dataset are not publicly available due to the privacy of patients.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the Mayo Clinic Center for Health Equity and Community Engagement Research Award. The funders had no role in the design of the study, or collection, analysis, and interpretation of data and in preparation of the manuscript. The views presented in this report are not necessarily representative of the funder's views and belong solely to the authors.

Authors' contributions

YW conceptualized the research, led the analyses and interpretation of data. BSAK performed analyses and interpretation of data. JEB participated in the conception and design of the study. WDF and RZ made substantial contributions to the data acquisition. BSAK and YW drafted the manuscript. BSAK, JEB, WDF, RZ, and YW critically revised the manuscript for important intellectual content. All authors read and approved the final manuscript. All authors agreed to be accountable for all aspects of the work.

Acknowledgements

None.

Disclaimer

The view(s) expressed herein are those of the author(s) and do not reflect the official policy or position of Mayo Clinic.

References

1. Social Determinants of Health. Office of Disease Prevention and Health Promotion. Accessed at: <https://www.healthypeople.gov/2020/topics-objectives/topic/social-determinants-of-health> 10/26/2020.
2. Tapp J, Hudson T. Sexually Transmitted Infections Prevalence in the United States and the Relationship to Social Determinants of Health. *Nurs. Clin. North Am.* 2020;**55**(3):283-93 doi: 10.1016/j.cnur.2020.05.001[published Online First: Epub Date].
3. Stubbs VC, Rajasekaran K, Cannady SB, Newman JG, Ibrahim SA, Brant JA. Social determinants of health and survivorship in parotid cancer: An analysis of the National Cancer Database. *Am. J. Otolaryngol.* 2020;**41**(1):102307 doi: 10.1016/j.amjoto.2019.102307[published Online First: Epub Date].
4. Federico MJ, McFarlane AE, 2nd, Szeffler SJ, Abrams EM. The Impact of Social Determinants of Health on Children with Asthma. *J Allergy Clin Immunol Pract* 2020;**8**(6):1808-14 doi: 10.1016/j.jaip.2020.03.028[published Online First: Epub Date].
5. Williams AD, So A, Synnestvedt M, et al. Mammographic breast density decreases after bariatric surgery. *Breast Cancer Res. Treat.* 2017;**165**(3):565-72
6. Friedman C. Social determinants of health, emergency department utilization, and people with intellectual and developmental disabilities. *Disabil Health J* 2020:100964 doi: 10.1016/j.dhjo.2020.100964[published Online First: Epub Date].

7. Lax Y, Martinez M, Brown NM. Social Determinants of Health and Hospital Readmission. *Pediatrics* 2017;**140**(5) doi: 10.1542/peds.2017-1427[published Online First: Epub Date]].
8. Braveman P, Gottlieb L. The social determinants of health: it's time to consider the causes of the causes. *Public Health Rep.* 2014;**129 Suppl 2**:19-31 doi: 10.1177/00333549141291S206[published Online First: Epub Date]].
9. Taylor LA, Tan AX, Coyle CE, et al. Leveraging the Social Determinants of Health: What Works? *PLoS One* 2016;**11**(8):e0160217 doi: 10.1371/journal.pone.0160217[published Online First: Epub Date]].
10. Lathrop B. Moving Toward Health Equity by Addressing Social Determinants of Health. *Nurs. Womens Health* 2020;**24**(1):36-44 doi: 10.1016/j.nwh.2019.11.003[published Online First: Epub Date]].
11. Tallon MM, Kendall GE, Priddis L, Newall F, Young J. Barriers to Addressing Social Determinants of Health in Pediatric Nursing Practice: An Integrative Review. *J. Pediatr. Nurs.* 2017;**37**:51-56 doi: 10.1016/j.pedn.2017.06.009[published Online First: Epub Date]].
12. Davis SL, Chapa DW. Social Determinants of Health: Knowledge to Effective Action for Change. *The Journal for Nurse Practitioners* 2015;**11**(4):424-29 doi: <https://doi.org/10.1016/j.nurpra.2015.01.029>[published Online First: Epub Date]].
13. Singh GK, Daus GP, Allender M, et al. Social Determinants of Health in the United States: Addressing Major Health Inequality Trends for the Nation, 1935-2016. *Int J MCH AIDS* 2017;**6**(2):139-64 doi: 10.21106/ijma.236[published Online First: Epub Date]].
14. Marmot M. Social determinants of health inequalities. *Lancet* 2005;**365**(9464):1099-104 doi: 10.1016/S0140-6736(05)71146-6[published Online First: Epub Date]].
15. Viner RM, Ozer EM, Denny S, et al. Adolescence and the social determinants of health. *Lancet* 2012;**379**(9826):1641-52 doi: 10.1016/S0140-6736(12)60149-4[published Online First: Epub Date]].
16. Eton DT, Beebe TJ, Hagen PT, et al. Harmonizing and consolidating the measurement of patient-reported information at health care institutions: a position statement of the Mayo Clinic. *Patient Relat Outcome Meas* 2014;**5**:7-15 doi: 10.2147/prom.S55069[published Online First: Epub Date]].
17. Pakhomov SV, Jacobsen SJ, Chute CG, Roger VL. Agreement between patient-reported symptoms and their documentation in the medical record. *Am. J. Manag. Care* 2008;**14**(8):530-9
18. Secondary. <https://www.cdc.gov/std/stats17/minorities.htm>.
19. Frier A, Devine S, Barnett F, Dunning T. Utilising clinical settings to identify and respond to the social determinants of health of individuals with type 2 diabetes-A review of the literature. *Health Soc Care Community* 2020;**28**(4):1119-33 doi: 10.1111/hsc.12932[published Online First: Epub Date]].
20. Chung SD, Lin YK, Huang CC, Lin HC. Increased risk of prostate cancer following sexually transmitted infection in an Asian population. *Epidemiol. Infect.* 2013;**141**(12):2663-70 doi: 10.1017/S0950268813000459[published Online First: Epub Date]].
21. Dixon BE, Rahurkar S, Ho Y, Arno JN. Reliability of administrative data to identify sexually transmitted infections for population health: a systematic review. *BMJ Health Care Inform* 2019;**26**(1) doi: 10.1136/bmjhci-2019-100074[published Online First: Epub Date]].
22. Weiskopf NG, Hripcsak G, Swaminathan S, Weng C. Defining and measuring completeness of electronic health records for secondary use. *J. Biomed. Inform.* 2013;**46**(5):830-6 doi: 10.1016/j.jbi.2013.06.010[published Online First: Epub Date]].
23. Rocca WA, Grossardt BR, Brue SM, et al. Data Resource Profile: Expansion of the Rochester Epidemiology Project medical records-linkage system (E-REP). *Int J Epidemiol* 2018;**47**(2):368-68j doi: 10.1093/ije/dyx268[published Online First: Epub Date]].
24. Boerma T, Hosseinpoor AR, Verdes E, Chatterji S. A global assessment of the gender gap in self-reported health with survey data from 59 countries. *BMC Public Health* 2016;**16**:675 doi: 10.1186/s12889-016-3352-y[published Online First: Epub Date]].

25. Cantor MN, Thorpe L. Integrating Data On Social Determinants Of Health Into Electronic Health Records. Health Aff. (Millwood) 2018;**37**(4):585-90 doi: 10.1377/hlthaff.2017.1252[published Online First: Epub Date].
26. Deferio JJ, Breitinger S, Khullar D, Sheth A, Pathak J. Social determinants of health in mental health care and research: a case for greater inclusion. J. Am. Med. Inform. Assoc. 2019;**26**(8-9):895-99 doi: 10.1093/jamia/ocz049[published Online First: Epub Date].

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SuplTables.docx](#)
- [SuplTables.docx](#)
- [SuplTables.docx](#)
- [regextoassigntopics.txt](#)
- [regextoassigntopics.txt](#)
- [regextoassigntopics.txt](#)