Leveraging Data and Digital Health Technologies to Assess and Impact Social Determinants of Health (SDoH): a State-of-the-Art Literature Review

Kelly J. Thomas Craig, PhD[1]*, Nicole Fusco, ScD[2], Thrudur Gunnarsdottir, PhD[2], Luc Chamberland, MA[2], Jane L. Snowdon, PhD[1], William J. Kassler, MD, MPH[2,3].

[1]IBM® Watson Health®, Center for AI, Research, and Evaluation, 75 Binney Street, Cambridge, MA, 02142 USA, [2]IBM® Watson Health®, 75 Binney Street, Cambridge, MA, 02142 USA, [3] Palantir Technologies, 1555 Blake Street, Denver, CO 80202

Abstract

Objective: Identify how novel datasets and digital health technology, including both analytics-based and artificial intelligence (AI)-based tools, can be used to assess non-clinical, social determinants of health (SDoH) for population health improvement.

Methods: A state-of-the-art literature review with systematic methods was performed on MEDLINE, Embase, and the Cochrane Library databases and the grey literature to identify recently published articles (2013-2018) for evidence-based qualitative synthesis. Following single review of titles and abstracts, two independent reviewers assessed eligibility of full-texts using predefined criteria and extracted data into predefined templates.

Results: The search yielded 2,714 unique database records of which 65 met inclusion criteria. Most studies were conducted retrospectively in a United States community setting. Identity, behavioral, and economic factors were frequently identified social determinants, due to reliance on administrative data. Three main themes were identified: 1) improve access to data and technology with policy – advance the standardization and interoperability of data, and expand consumer access to digital health technologies; 2) leverage data aggregation – enrich SDoH insights using multiple data sources, and use analytics-based and AI-based methods to aggregate data; and 3) use analytics-based and AI-based methods to assess and address SDoH – retrieve SDoH in unstructured and structured data, and provide contextual care management sights and community-level interventions.

Conclusions: If multiple datasets and advanced analytical technologies can be effectively integrated, and consumers have access to and literacy of technology, more SDoH insights can be identified and targeted to improve public health. This study identified examples of AI-based use cases in public health informatics, and this literature is very limited.

Keywords: social determinants of health, artificial intelligence, digital health, data analytics, health policy

Abbreviations: application programming interfaces (API), artificial intelligence (AI), electronic health records (EHRs), machine learning (ML), natural language processing (NLP), return on investment (ROI), social determinants of health (SDoH)

* Correspondence: kelly.jean.craig@ibm.com

DOI: 10.5210/ojphi.v13i3.11081

Copyright ©2021 the author(s)

This is an Open Access article. Authors own copyright of their articles appearing in the Online Journal of Public Health Informatics. Readers may copy articles without permission of the copyright owner(s), as long as the author and OJPHI are acknowledged in the copy and the copy is used for educational, not-for-profit purposes.

Introduction

In the current clinical practice and policy environment, significant attention is focused on non-clinical factors outside of the healthcare system that impact health outcomes. It is generally acknowledged that healthcare alone accounts for only 10-25% of the variance of health outcomes [1]. According to the County Health Rankings model, "modifiable contributions to health include social and economic factors (40%), behavior (30%), clinical care (20%), and environmental factors (10%)" [2]. These health factors are often collectively referred to as social determinants of health (SDoH) even though these non-clinical determinants are not strictly social considerations [3]. Both policymakers [4] and healthcare systems [5] have begun to recognize the importance of factors, such as education, income, housing, food, and the built environment, to help improve health outcomes, reduce disparities, and lower medical costs. Awareness is still required to promote the widespread adoption of integrated healthcare and social services approaches, and long-term investments are needed in the healthcare sector to facilitate referrals for services that address SDoH. Common barriers to social services investments include perceived lack of short-term return on investments, inability to identify and allocate costs to patient care, lack of coverage and reimbursement by payers, and challenges in coordination, communication, and data fluidity between medical and social organizations [6].

There is a paucity of evidence for prioritizing SDoH factors for intervention based on the financial return on investment (ROI). Guidance from the National Academy of Medicine [7] and recommendations provided by the World Health Organization [8] supports the premise that actionable SDoH factors (i.e., those that lend themselves to direct action) will provide strong use cases for integration of social services into ongoing operations. Optimal interventions will target the community level to improve the underlying social and economic conditions contributing to health disparities, rather than mediating individual needs alone [5]. The incorporation of non-clinical support, such as providing food or housing, can improve health when applied on an appropriate scale, but the financial aspects of integrating SDoH interventions remain a challenge.

Despite the financial challenges of integrating SDoH assessment and interventions into healthcare, data and technology are being leveraged to improve both the identification of actionable SDoH factors and population health. The transition to value-based care has promoted the development of collaborative tools utilizing SDoH data and innovative technology to identify and stratify the highest-risk individuals within a community. As the effort to gather SDoH data is prioritized, digital health technologies can be foundational in the collection and assessment of SDoH data, including development of novel sources of data and the platforms by which this information is exchanged between stakeholders. Predictive analytics and other new methods of data analysis including artificial intelligence (AI) can be important tools to identify, assess, and provide insights to address SDoH and their associated health outcomes and related disparities, in order to improve efficiencies in evaluation with the goal of improving care and lowering costs.

The objective of this study was to perform a state-of-the-art literature review to identify how SDoH data and digital health technology are leveraged to improve population health management. Studies of interest were those using large and innovative datasets in addition to studies employing digital health technologies, including those with analytical-based and AI-based methods, to assess and address SDoH factors as a strategy for population health improvement.

Methods

A state-of-the-art literature review with systematic methods was conducted [9]. Searches for all relevant articles were conducted in MEDLINE via PubMed, the Cochrane Library, and Embase according to the summative methods and search strategies outlined in Supplemental Tables 1-7. There was a focus on non-clinical SDoH; health disparity; and health equity terms for our population of interest. A manual search of the bibliographies of full-text articles pertinent to the review was also conducted.

One reviewer screened titles and abstracts of articles identified by the literature searches to select a list of articles to be considered as sources for the report. Two reviewers independently screened all potential full-text citations to determine which sources were included using *a priori* criteria (i.e., included articles must describe digital health technology or novel data used to assess and/or address SDoH). Inclusion and exclusion criteria for (a) large, population-based studies, (b) smaller studies of actionable SDoH factors, and (c) policy pieces including grey literature are outlined in Supplemental Tables 8-10, respectively. Included articles must describe digital health technology or novel data used to assess and/or address SDoH, broadly defined by multiple leading agencies (National Academy, Institute of Medicine; Centers for Disease Control and Prevention; Office of Disease Prevention and Health Promotion; Robert Wood Johnson Foundation; World Health Organization; and Kaiser Family Foundation). Example inclusion criteria are health disparity, tobacco use, educational status, unemployment, and housing instability. Example exclusion criteria are *in vivo* and *in vitro* studies, non-systematic reviews, and global perspective.

Data were extracted into Excel tables from each of the studies meeting the inclusion/exclusion criteria. Studies that met multiple criteria of interest (i.e., data, technology, and/or policy) were categorized and counted more than once. Data were then checked by a second independent reviewer. Supplemental Tables 11-13 provide data extraction organized into policy, digital health technology, and data areas, respectively.

Results

Search Results and Identified Study Characteristics

Literature searches identified 2,714 potentially relevant articles for title abstract screening; 132 were identified for full-text screening, and 65 studies met the inclusion criteria. All search and screening results are presented in Figure 1.

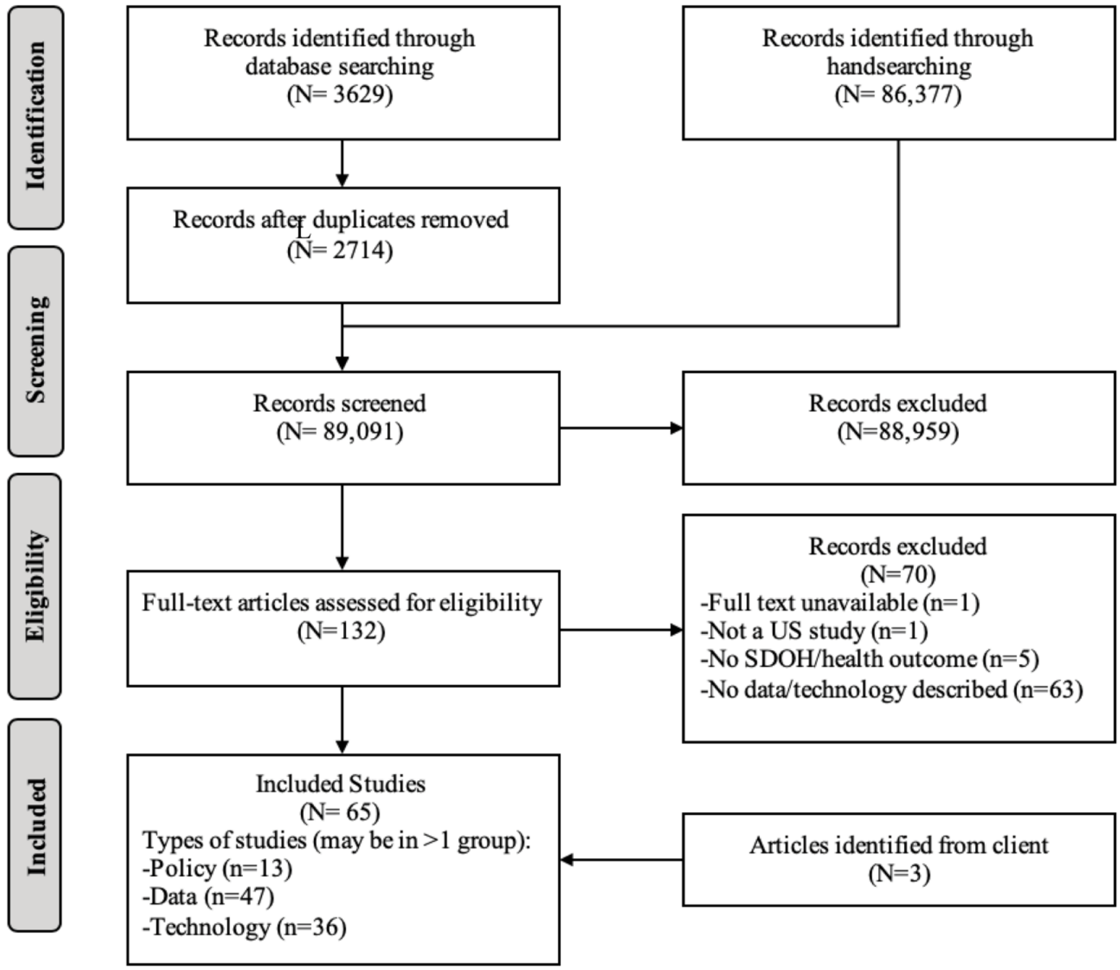

Figure 1. Disposition of articles and screening flow diagram

This flow diagram depicts the process and flow of information (including number of records of identified, included and excluded study numbers, and reasons for exclusion) through the phases of the state-of-the-art literature review.

Supplementary tables 11-13 present the characteristics of the included studies. Most studies were designed retrospectively in US community settings, and all were full-texts. Of the 65 included articles, 13 referenced SDoH in public health policies, 47 presented novel sources and/or novel uses of SDoH data, and 36 described technologies.

Overall, the 13 articles regarding policy focused on access, both improved access to smart mobile devices or broadband internet (n=4), and the measurement, collection, and dissemination of SDOH data (n=9). Of the 36 articles describing technology, most (n=23) pertained to utilization of geocoding to translate patient addresses into spatial SDoH data, but AI (n=9) was also commonly leveraged. Digital health technologies identified in three or fewer articles included social media, other types of software (e.g., modeling or simulation software, transition of care platforms), internet, mobile health (mHealth) sensors and applications (apps), video, and telehealth. Of the 47 articles pertaining to data, most (n=32) combined multiple types of data sources (e.g., electronic health records [EHRs] and census data). Other data sources included EHRs (n=8), census (n=3), state (n=2), federal (n=2), and internet (n=2). Baseline SDoH factor assessments were categorized as: identity, behavioral, economic, neighborhood, physical environment, education, food, housing, social relationships, transportation, health access and quality, employment, community, governmental, and psychosocial. Identity (n=172), behavioral (n=98), and economic (n=53) were the most commonly examined SDoH factors.

Improve Access with Policy

Expand consumer access to digital health technologies

Newer digital health technologies, including telemedicine platforms, mHealth apps, wearable devices, patient portals, EHRs and health information exchanges, and other internet-based technologies and services have the potential to disrupt and improve public health practice. Across intervention types, governmental policy can incentivize behavior in healthcare and public health practice. Therefore, policy-related publications addressing these technologies and SDoH were examined. Findings showed that communities with insufficient resources and individuals with limited health literacy face barriers to adoption [10].

Three articles focused on the need to improve access to digital health technologies to reduce health disparities. In these articles, access to technology was limited to smart mobile devices or broadband internet. Ray et al. (2017) showed disparities in the use of technology for accessing health information by race/ethnicity and income. The authors recommended policies to improve access to technology in these communities [11]. As vulnerable populations gain access to smart phones, grant makers acknowledge that when combined with sound policy strategies, scalable digital technologies supported by communities have the potential to positively impact health disparities that vary by geography [12]. As demonstrated by recent increases in technology usage by racial/ethnic minorities, the digital divide may be narrowing by race/ethnicity, but not by health literacy. Chakkalakal et al. (2014) observed that individuals with limited health literacy were less likely to own and/or utilize these technologies when compared to those with adequate health literacy [13].

Improve standardization and interoperability of data

Recognizing that challenges with effective measurement and collection of SDoH data are barriers to reducing health disparities, numerous public policy guidelines have encouraged the incorporation of SDoH data in health-monitoring systems [14-16]. Four articles described and recommended practices, and in some instances standardization and prioritization, in the measurement of SDoH factors for actionable interventions [17-20]. The Public Health 3.0 initiative prioritized education, safe environments, housing, transportation, economic development, and access to healthy foods in community-level interventions utilizing cross-sector partnerships with stakeholders [17]. Super Church (2015) described how one program, the Healthy Neighborhoods Equity Fund, planned to use data to identify communities suffering disproportionately from health disparities and would benefit most from a community intervention, including improvement in "housing conditions, public safety, employment, transportation, walkability, and access to green space and healthy food" [18]. SDoH factors included in the data assessment were community support and growth potential, transportation access and use, walkability, housing affordability and choice, safety, economic opportunity, recreational areas, food access, indoor air quality, and building and site performance [18]. Cahill et al. (2016) provided recommendations for SDoH factors that should be incorporated into an EHR as a part of meaningful use guidelines. The authors argue that including sexual orientation and gender data (birth sex, preferred name and pronouns, sexual orientation, and sexual practices/behaviors) in EHRs are vital to improving disparities in sexually marginalized groups since this information will better inform providers and improve patient care [19]. Providers and EHR vendors should also be trained on how to collect, interpret, and use this information to best improve care for patients [19]. Lastly, Penman-Aguilar et al. (2016) provided recommended practices for the measurement of SDoH factors at national and sub-national (i.e., state) levels including: *a priori* identification of characteristics of groups that are associated with less power and privilege or lower social position that may demonstrate within-group heterogeneity, and they be measured at multiple levels, i.e., both the individual and community levels. These domains include race/ethnicity, sex, sexual orientation and gender identity, age, level of education, income, wealth, occupation, country of birth, disability status, and geographic location [20]. Additionally, the standardization of methodologies for the measurement of SDoH is an unmet need. However, the methodologies are greatly influenced by how the SDoH information will be disseminated to a particular audience, so analytical methods are varied, which can prevent or limit aggregation of datasets [20].

Despite multiple national programs [21, 22] underscoring the importance of capturing SDoH data in EHRs to guide clinical care, SDOH data collection is currently insufficient. Two policy articles not only assessed policies and programs to improve its collection, but also specified the level of granularity to be collected [17, 23]. Douglas et al. (2015) identified gaps in data collection requirements in the Health Information Technology for Economic and Clinical Health (HITECH) Act of 2009 that drove the national adoption of EHRs. The authors noted deficiencies for collecting data on certain demographics will, at the least, not help resolve health disparities and, at worst, exacerbate them. It was recommended that the Act should require the collection of "more granular race and ethnicity, disability status, and sexual orientation and gender identity" data in EHRs [23]. Demographic data obtained from studies supported by federally funded agencies (i.e., National Academy of Medicine, Centers for Disease Control and Prevention, Department of Health and Human Services [DHHS], etc.) are not comparable due to lack of alignment and program mandates for its expanded collection [23]. Furthermore, inconsistent demographic data collection standards between public health survey data and EHRs limit research interpretation and widen the health disparities gap. In addition to demographic data being more granular, geographic granularity has also been emphasized at the community level. DeSalvo et al. (2017) commented on Public Health 3.0, an initiative by the US DHHS to broaden the scope of public health to directly include communities to address SDoH factors [17]. In efforts to achieve Public Health 3.0, data should be made available to communities, which would enable real-time and geographically granular (i.e., county-level) data to be shared, connected, and interpreted to translate evidence into action [17].

To impact health outcomes, social determinants must be measured, and the data need to be collected and widely disseminated. Four articles discussed the development, importance, and concerns around sharing SDoH data [24-27]. Krumholz et al. (2016) provided recommendations for data sharing, including 1) foster a culture of data sharing, 2) develop operational functionality for data sharing, and 3) improve data sharing capacity [24]. Perlin et al. (2016) also provided recommendations with a focus on the operationality of systems: interoperability between EHR systems, improved cyber security, and a data strategy that supports a learning health system.

Finally, Smith et al. (2016) described the development and implementation of a data warehouse and EHR system for Federally Qualified Health Centers in Maryland. Important aspects of the system's development included engaging partners, defining quality measures and processes for validating data, and determining effective use of the data for clinical quality improvement. The two biggest challenges in implementation were cost and communication of data and outcomes within an organization and between organizations using population health software platforms [26]. However, privacy and security concerns related to the inclusion of SDoH data in EHRs remain credible threats to their collection and subsequent sharing. At the individual level, many measurements of the SDoH may be highly personal and sensitive. As such, compliance to security and privacy laws, and implementation of good practices to provide stewardship of these data is essential to earn patient trust and their willing disclosure of sensitive information [27].

Data Aggregation

Enrich SDoH insights by using multiple data sources

Innovative methods have been leveraged to mine existing data for new insights related to SDoH. Many (74%) of the articles addressing data on SDoH combined multiple sources, both structured (e.g., machine-readable data) and unstructured (e.g., text). EHRs and US census data were two of the most commonly used sources to glean insights, but other sources included federal, state, and local clinical data (e.g., nationally conducted health surveys, registry data). The use of non-clinical data from public records (e.g., housing, crime, welfare) and novel third-party sources such as internet content from social media (e.g., Twitter, Yelp) were also identified. Notably, when EHR data is combined with other data sources to examine health outcomes, the number of SDoH factors collected are enriched. Compared with EHR data alone, adding disparate data sources captures additional SDoH factors (i.e., neighborhood, food, and education) and actionable insights which provide opportunities for intervention at the community level.

Analytics and AI-based methods to aggregate data

The expanded scope of a population health approach, which incorporates non-clinical perspectives, leads to an ecosystem composed of multiple, diverse data sources that need to be integrated to obtain actionable insights at the individual and community levels. Geocoding, a geospatial analytic technique, is one way to link these datasets by matching locations.

Geocoding (n=23) was the most frequent technology identified to assess and address SDoH. Geocoding was broadly applied to link individual (most commonly, EHR) and community datasets (e.g., US Census, state, or local data) to provide a meaningful, contextual, and geographic analysis of SDoH, and was used primarily to capture identity, economic, neighborhood, and behavioral factors across a variety of health outcomes. For example, Masho et al. (2017) used ArcGIS software to perform geocoding to link three datasets: 10-year birth registry data, 2010 US Census data, and crime statistics from a local police department [28]. The linked data enabled the inclusion of more actionable SDoH, including neighborhood factors like safety which are not typically captured within health data. The analytical assessments identified the association of social factors with health outcomes; youth violence was associated with preterm birth despite controlling for other variables at the individual and community levels [28]. Geocoding can be used to develop targeted contextual and geographic interventions; in this instance, to reduce community violence and preterm births in vulnerable populations.

The use of geocoding, which typically involves specification of latitude and longitude by application programming interfaces (APIs) from mapping databases, has more benefits than traditional linking via county or zip code. This level of geocoding provided more granularity and facilitated the collection of more neighborhood and behavioral-related SDoH factors. Zip code linking primarily provided information regarding community level measurements, such as income and percent race/ethnicity [29, 30]. Other mechanisms of linking data included using probabilistic software, "Link Plus," [31] or a custom probabilistic algorithm to match different datasets [32]. However, these types of linking required the removal of protected health information before the data could be disclosed to investigators, which is a barrier to their use.

Compared to standalone assessments, technology used to link datasets collects previously unattainable SDoH factors. Combined data sets with geocoded EHRs can identify new SDoH insights; for example, Perzynski et al. (2017) assessed whether SDoH affect healthcare utilization via patient portals from outpatient clinics in an urban public health care system. EHR data was geocoded to provide granularity at the census-tract (i.e., subdivision of county) levels and combined with FCC 2013 Form 477 data (e.g., broadband internet usage) and American Community Survey (ACS) data (e.g., neighborhood income and education level) [33]. These data sources cumulatively revealed access to technology was significantly associated with patient portal usage; the elderly, uninsured, and minority patients were less likely to use the portal. This approach was insightful because it used both novel data sources and used data in new ways to identify the association of an SDoH factor with healthcare utilization.

While advanced analytical tools in healthcare have been primarily used to link existing data streams, some researchers have applied these tools to create novel datasets from non-traditional, non-clinical sources. For example, Nguyen et al. (2017) leveraged geographically tagged (geotagged, i.e., with geocoded data embedded in the media) social media data from Twitter and Yelp to create a national food environment database [34]. The authors applied analytical tools including AI techniques to discover patterns and emerging health-related issues in aggregated datasets with geographically identified metadata. To assess sentiment regarding food consumption, Machine Learning for LanguagE (MALLET) was used to analyze geotagged Twitter content (text and images) from its API. Using data from the US Department of Agriculture national nutrient database, algorithms were applied to calculate caloric density of popular food and alcohol tweet content and to create a detailed view of the food environment using Yelp's API. The social media derived data were then linked to data from the ACS, Behavioral Risk Factor Surveillance System, and National Vital Statistics Report to assess state-level health outcome data that included chronic health conditions, all-cause mortality, and homicide rate, respectively. Data were then assessed to determine whether state-level food environment variables obtained from social media were associated with health outcomes. The data were further analyzed with additional granularity to assess the relationship between county-level social media indicators and county health outcomes. High caloric density food tweets and more burger Yelp listings were related to higher all-cause mortality, diabetes, obesity, high cholesterol, and fair/poor self-rated health. More alcohol tweets and Yelp bars and pub listings were related to higher state-level binge drinking, but, curiously, lower mortality and lower percent reporting fair/poor self-rated health. The primary goal of Nguyen et. al (2017) was to use social media data to assess attitudes, norms, and behavioral control activities of a community; however, the study produced a novel, web-based national database that captured social-environmental features at the county level to examine potential impacts on health [34]. Social media and/or AI technologies were paired with geographic information in five of the included studies assessing SDoH.

Analytics and AI-Based Methods to Identify and/or Assess SDoH

The application of AI, such as machine learning (ML) or natural language processing (NLP), can augment the retrieval of SDoH hidden in unstructured data (e.g., text found in clinical notes or social services documents). Nine articles used AI methods (n=7, ML; n=2, NLP) to facilitate the extraction of SDoH terms from EHR data, provide care management insights, and/or provide connectivity of disparate datasets using API as earlier described [34]. Oreskovic et al. (2017) used keyword searching by NLP in an analytics platform to identify SDoH factors in EHRs related to increased psychosocial risk, whereby patients may be eligible for enrollment in a care coordination program [35]. In this instance, the authors determined that psychosocial risk factors are associated with higher healthcare utilization and costs, and worsening clinical outcomes among Medicaid patients. This modeling, used in conjunction with the care coordination program, allowed providers and healthcare systems to assess and manage their risk pool after quantifying and triaging psychosocial risk. Similarly, Jamei et al. (2017) built a neural network model to analyze SDoH data from EHRs of a large health system to identify high-risk patients and predict all-cause risk of 30-day hospital readmission [36]. The limited SDoH derived data (alcohol, drug, and tobacco use) from the EHRs was supplemented with geocoded 2010 census data to the block-level and matched patients' addresses for linking purposes. The authors noted that the predictive power of the model needed to be tested in data sources that contain more granular and structured SDoH, such as that collected using SDoH screening tools, and suggested that NLP [37] could extract additional SDoH measurements from case notes of individual patients. However, Navathe et al. (2018) demonstrated that seven common SDoH factors (tobacco use, alcohol use, drug abuse, depression, housing instability, fall risk, and poor social support) were more readily identified in the unstructured physician notes of an EHR using NLP extraction when compared to screening administrative sources, such as claims and structured EHR data [38]. Automated methods for analyzing physician notes enabled better identification of social needs of patients at risk for readmission. Programs, particularly those with finite resources, greatly benefit from this novel approach to assessing SDoH data.

Discussion

This review of the literature acknowledges that access to data along with innovative datasets and digital technologies improve population health strategies for SDoH assessment (Figure 2). Most studies focused on identity, economic, and sociobehavioral factors in relation to health outcomes. Demographic factors (e.g., identity and economic) are more easily and frequently collected than more actionable factors. However, interventions for behavioral-related SDoH factors are more feasible and have demonstrated examples of high ROI (ROI range, $7.60 to $16.70 returned for every US dollar spent) to promote healthy eating and weight loss [39, 40]. Policy recommendations [8, 17] to prioritize other actionable SDoH factors such as education, food, and housing in targeted interventions show reduced costs [41-43].

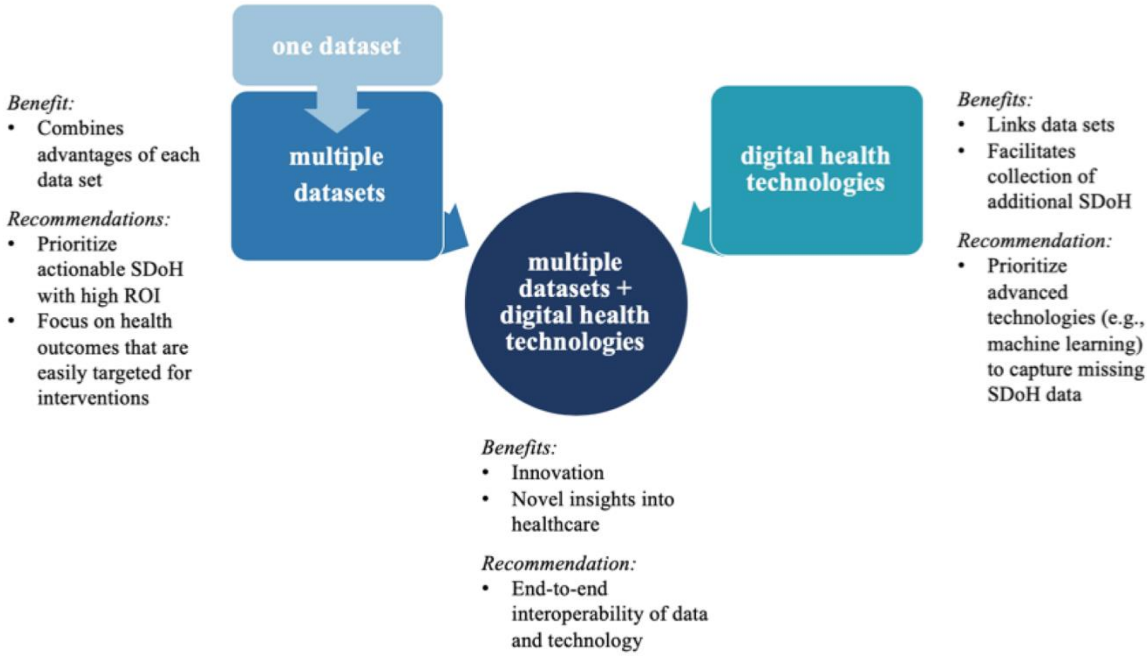

Figure 2. Using multiple datasets with digital health technologies adds the most value

Abbreviations: ROI, return on investment; SDoH, social determinants of health.

Combining data sets may be challenging, but the result can beneficially address SDoH. Combining multiple data sources allows for the study of novel SDoH by linking data less commonly used in healthcare research, such as access to healthy foods or neighborhood safety, with health outcome data. However, linking datasets owned by separate entities can be difficult because of patient privacy concerns; to be compliant with privacy laws, personal health information has to be removed from datasets before they are shared with external researchers, so individual patient data cannot be matched across datasets. Several studies utilizing technology to facilitate data linking through geocoding were identified but may be limited by constructive identification.

Technology can facilitate the integration of multiple data sources to assess SDoH for new insights. For example, technologies with advanced analytics methods, including the use of AI-based algorithms, has the ability to provide personalized care to patients and support comprehensive, effective, and thoughtful care management. The development of predictive models can assist decision makers in cost-saving analyses to more effectively schedule and optimize hospital resources by identifying high-risk patients and correctly determining where resources will provide the most benefit (e.g., providing timely intervention to reduce hospital readmission in high-risk patients). The use of advanced analytics and AI-based services, including cloud-based AI analytics microservices, to address SDoH is paramount to generate data insights to inform decision-making.

Despite its value, technology can also be a barrier. Digital health technology access is not a commonly measured SDoH factor, and lack of access can both result in biased data collection (e.g., only collecting data from patients with technology access) and contribute to further disparities because of the "digital divide," whereby lack of technology literacy and broadband access or mobile phones worsen economic and social inequalities. Notably, policy considerations focused on the need to improve access to digital health technology to reduce health disparities.

Furthermore, incentive-based policies can help create positive habits and behavior and break negative habits in the short-term; however, sustained behavior change continues to be a challenge for many.

There are some limitations to this review. Despite exhaustive search methods, *a priori* inclusion criteria, and dual screening of full-texts under review, the included studies were ranked by greater generalizability, longer follow-up, larger sample sizes, and articles published between 2013-2018.

Future Directions

As healthcare transitions to value-based care delivery, harnessing data sources and leveraging technology to collect SDoH will be essential. The ability to leverage big data for population health management has the potential to improve health outcomes, bridge gaps in care, and reduce costs. The integration of data from disparate sources to understand the composition of the population and stratify individuals according to risk scores will be transformative; the identification of underlying factors that influence patient and community health allows for more practical and meaningful care. The collection of SDoH data, particularly those with high potential ROI, is essential for a more holistic view of the patient to be reflected in patient records and the consolidation of this information across care teams. Clinical enablement tools to capture SDoH factors and to obtain missing SDoH data need to be improved and broadly integrated. Digital health technology and data have the potential to augment and scale labor-intensive and manual processes to identify social needs for the patient, whereby healthcare providers can connect them to the appropriate resources to overcome those barriers to health. With the evolution of technology and value-based care for patient management, it will require the collaboration between clinical and social care teams to improve health disparities related to SDoH with the goal of improving health equity.

Conclusions

If multiple datasets and advanced analytical technologies can be effectively integrated, and consumers have access to and literacy of technology, more SDoH insights can be identified and targeted to improve public health. This study identified examples of AI-based use cases in public health informatics, and this literature is very limited.

Acknowledgements

We would like to thank Bill Marder and Ron Ozminkowski for project guidance; Rachel Faller for screening and abstraction; Brett South, Megan Sands-Lincoln, Mason Russell, and Laura Morgan for critical reading; and Dave Liederbach and Kyu Rhee for project support.

Financial Disclosure

This research study was supported by IBM® Watson Health®.

Competing Interests

The authors are or were employed by IBM® Corporation and have no conflicts germane to this study.

References

1. Blatt, E., et al., *Addressing social determinants and their impact on healthcare.* IBM Cúram Research Institute, February, 2013.

2. *What is health? Health is more than healthcare. County Health Rankings & Roadmaps.* May 24, 2018]; Available from: http://www.countyhealthrankings.org/what-is-health.

3. Magnan S. *Social Determinants of Health 101 for Health Care: Five Plus Five*. NAM Perspectives: Expert Voices in Health & Health Care 2017 [cited 2019 December 2]; Available from: https://nam.edu/social-determinants-of-health-101-for-health-care-five-plus-five/.

4. Kassler WJ, et al. 2017. Population health measurement at Centers for Medicare & Medicaid Services: Bridging the gap between public health and clinical quality. *Popul Health Manag*. 20(3), 173-80. PubMed https://doi.org/10.1089/pop.2016.0038

5. Solomon LKM. 2018. Health Care Steps Up to Social Determinants of Health: Current Context. *Perm J*. (22), 18-139. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6207436/. https://doi.org/10.7812/TPP/18-139

6. KPMG Government Institute. *Investing in social services as a core strategy for health care organizations: Developing the business case*, T.C. Fund, Editor. 2018.

7. Long PA. M; Milstein, A; Anderson, G; Lewis Apton, K; Lund Dahlberg, M; Whicher, D., *Effective Care for High-Need Patients: Opportunities for Improving Outcomes, Value, and Health.* 2017, National Academy of Medicine: Washington, DC.

8. Dyakova, M., et al., *Investment for health and well-being: a review of the social return on investment from public health policies to support implementing the Sustainable Development Goals by building on Health 2020.* 2017.

9. Grant MJ, Booth A. 2009. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Info Libr J*. 26(2), 91-108. doi:https://doi.org/10.1111/j.1471-1842.2009.00848.x. PubMed

10. Terry NP. 2016. Will the internet of things transform healthcare. *Vand. J. Ent. & Tech. L.* 19, 327.

11. Ray R, et al. 2017. Missed Opportunity? Leveraging Mobile Technology to Reduce Racial Health Disparities. *J Health Polit Policy Law*. 42(5), 901-24. doi:https://doi.org/10.1215/03616878-3940477. PubMed

12. Graham GN, Ostrowski M, Sabina AB. 2016. Population health-based approaches to utilizing digital technology: a strategy for equity. *J Public Health Policy*. 37(Suppl 2), 154-66. doi:https://doi.org/10.1057/s41271-016-0012-5. PubMed

13. Chakkalakal RJ, et al. 2014. Disparities in using technology to access health information: race versus health literacy. *Diabetes Care*. 37(3), e53-54. PubMed https://doi.org/10.2337/dc13-1984

14. Daniel H, et al. 2018. Addressing social determinants to improve patient care and promote health equity: An american college of physicians position paper. *Ann Intern Med*. 168(8), 577-78. doi:https://doi.org/10.7326/M17-2441. PubMed

15. Department of Health and Human Services (US). *Healthy People 2020*. 2018 Nov 19, 2018 Nov 7, 2018]; Available from: https://www.healthypeople.gov/2020/default.

16. Sheiham A. 2009. Closing the gap in a generation: health equity through action on the social determinants of health. A report of the WHO Commission on Social Determinants of Health (CSDH) 2008. *Community Dent Health*. 26(1), 2-3. PubMed

17. DeSalvo KB, et al. 2017. Peer Reviewed: Public Health 3.0: A Call to Action for Public Health to Meet the Challenges of the 21st Century. *Prev Chronic Dis*. •••, 14.

18. Super Church M, & Using Data to Address Health Disparities and Drive Investment in Healthy Neighborhoods. 2015. National. *Acad Med*. •••, 1-5.

19. Cahill SR, et al. 2016. Inclusion of Sexual Orientation and Gender Identity in Stage 3 Meaningful Use Guidelines: A Huge Step Forward for LGBT Health. *LGBT Health*. 3(2), 100-02. doi:https://doi.org/10.1089/lgbt.2015.0136. PubMed

20. Penman-Aguilar A, et al. 2016. Measurement of Health Disparities, Health Inequities, and Social Determinants of Health to Support the Advancement of Health Equity. *J Public Health Manag Pract*. 22(Suppl 1), S33-42. doi:https://doi.org/10.1097/PHH.0000000000000373. PubMed

21. National Academy of Medicine. *Capturing Social and Behavioral Domains in Electronic Health Records, Phase 1*. Washington, DC: National Academies Press, 2014.

22. National Academy of Medicine. *Capturing Social and Behavioral Domains in Electronic Health Records, Phase 2*. Washington, DC: National Academies Press, 2014.

23. Douglas MD, et al. 2015. Missed policy opportunities to advance health equity by recording demographic data in electronic health records. *Am J Public Health*. 105(Suppl 3), S380-88. doi:https://doi.org/10.2105/AJPH.2014.302384. PubMed

24. Krumholz HM, Terry SF, Waldstreicher J. 2016. Data acquisition, curation, and use for a continuously learning health system. *JAMA*. 316(16), 1669-70. doi:https://doi.org/10.1001/jama.2016.12537. PubMed

25. Perlin JB. 2016. Health information technology interoperability and use for better care and evidence. *JAMA*. 316(16), 1667-68. doi:https://doi.org/10.1001/jama.2016.12337. PubMed

26. Smith EA, et al. 2016. Peer Reviewed: Using Health Information Technology and Data to Improve Chronic Disease Outcomes in Federally Qualified Health Centers in Maryland. *Prev Chronic Dis*. •••, E178. PubMed

27. National Academy of Medicine. *Capturing Social and Behavioral Domains and Measures in Electronic Health Records: Phase 2*. 2014, National Academies Press Washington, DC.

28. Masho SW, et al. 2017. Understanding the role of violence as a social determinant of preterm birth. *Am J Obstet Gynecol*. 216(2), 183.e1-7. doi:https://doi.org/10.1016/j.ajog.2016.10.001. PubMed

29. Beyer KM, et al. 2016. New spatially continuous indices of redlining and racial bias in mortgage lending: links to survival after breast cancer diagnosis and implications for health disparities research. *Health Place*. 40, 34-43. doi:https://doi.org/10.1016/j.healthplace.2016.04.014. PubMed

30. Sills MR, et al. 2016. Association of Social Determinants With Children's Hospitals' Preventable Readmissions Performance. *JAMA Pediatr*. 170(4), 350-58. doi:https://doi.org/10.1001/jamapediatrics.2015.4440. PubMed

31. Marino M, et al. 2016. Receipt of Preventive Services After Oregon's Randomized Medicaid Experiment. *Am J Prev Med*. 50(2), 161-70. doi:https://doi.org/10.1016/j.amepre.2015.07.032. PubMed

32. Rhoads KF, et al. 2015. How do integrated health care systems address racial and ethnic disparities in colon cancer? *J Clin Oncol*. 33(8), 854-60. doi:https://doi.org/10.1200/JCO.2014.56.8642. PubMed

33. Perzynski AT, et al. 2017. Patient portals and broadband internet inequality. *J Am Med Inform Assoc*. 24(5), 927-32. doi:https://doi.org/10.1093/jamia/ocx020. PubMed

34. Nguyen QC, et al. 2017. Social media indicators of the food environment and state health outcomes. *Public Health*. 148, 120-28. doi:https://doi.org/10.1016/j.puhe.2017.03.013. PubMed

35. Oreskovic NM, et al. 2017. Optimizing the Use of Electronic Health Records to Identify High-Risk Psychosocial Determinants of Health. *JMIR Med Inform*. 5(3). PubMed https://doi.org/10.2196/medinform.8240

36. Jamei M, et al. 2017. Predicting all-cause risk of 30-day hospital readmission using artificial neural networks. *PLoS One*. 12(7), e0181173. doi:https://doi.org/10.1371/journal.pone.0181173. PubMed

37. Navathe AS. *Identifying Patients at High-Risk for Readmission Using Socio-Behavioral Patient Characteristics*. in *Academy Health Annual Research Meeting*. 2016. Boston, MA.

38. Navathe AS, et al. 2018. Hospital Readmission and Social Risk Factors Identified from Physician Notes. *Health Serv Res*. 53(2), 1110-36.

https://onlinelibrary.wiley.com/doi/abs/10.1111/1475-6773.12670. PubMed https://doi.org/10.1111/1475-6773.12670

39. Michaud TL, et al. 2016. Effect of Workplace Weight Management on Health Care Expenditures and Quality of Life. *J Occup Environ Med*. 58(11), 1073-78. doi:https://doi.org/10.1097/JOM.0000000000000864. PubMed

40. Michaud TL, et al. 2017. Cost effectiveness and return on investment of a scalable community weight loss intervention. *Prev Med*. 105, 295-303. doi:https://doi.org/10.1016/j.ypmed.2017.10.011. PubMed

41. Ly A, Latimer E. 2015. Housing First Impact on Costs and Associated Cost Offsets: A Review of the Literature. *Can J Psychiatry*. 60(11), 475-87. doi:https://doi.org/10.1177/070674371506001103. PubMed

42. Martin SL, et al. 2018. Simply delivered meals: a tale of collaboration. *Am J Manag Care*. 24(6), 301-04. https://www.ncbi.nlm.nih.gov/pubmed/29939505. PubMed

43. Ramon I, et al. 2018. Early Childhood Education to Promote Health Equity: A Community Guide Economic Review. *J Public Health Manag Pract*. 24(1), e8-15. doi:https://doi.org/10.1097/PHH.0000000000000557. PubMed

44. Ahmed L, et al. *Technology-Enabled Transitions of Care in a Collaborative Effort between Clinical Medicine and Public Health: A Population Health Case Report.* National Academy of Medicine, 2016: p. 1-7.

45. Angier H, et al. 2014. Using geographic information systems (GIS) to identify communities in need of health insurance outreach: An OCHIN practice-based research network (PBRN) report. *J Am Board Fam Med*. 27(6), 804-10. http://www.jabfm.org/content/27/6/804.full.pdf. PubMed https://doi.org/10.3122/jabfm.2014.06.140029

46. Aoyagi H, Ogunseitan OA. 2015. Toxic releases and risk disparity: A spatiotemporal model of industrial ecology and social empowerment. *Int J Environ Res Public Health*. 12(6), 6300-18. https://res.mdpi.com/ijerph/ijerph-12-06300/article_deploy/ijerph-12-06300.pdf?filename=&attachment=1. PubMed https://doi.org/10.3390/ijerph120606300

47. Basu S, et al. 2017. Development and validation of Risk Equations for Complications Of type 2 Diabetes (RECODe) using individual participant data from randomised trials. *Lancet Diabetes Endocrinol*. 5(10), 788-98. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5769867/pdf/nihms921271.pdf. PubMed https://doi.org/10.1016/S2213-8587(17)30221-8

48. Bauer SR, Monuteaux MC, Fleegler EW. 2015. Geographic Disparities in Access to Agencies Providing Income-Related Social Services. *J Urban Health*. 92(5), 853-63. doi:https://doi.org/10.1007/s11524-015-9971-2. PubMed

49. Bejan CA, et al. 2018. Mining 100 million notes to find homelessness and adverse childhood experiences: 2 case studies of rare and severe social determinants of health in electronic health

records. *J Am Med Inform Assoc*. 25(1), 61-71. https://watermark.silverchair.com/ocx059.pdf?token=AQECAHi208BE49Ooan9kkhW_Ercy7Dm3ZL_9Cf3qfKAc485ysgAAAaIwggGeBgkqhkiG9w0BBwagggGPMIIBiwIBADCCAYQGCSqGSIb3DQEHATAeBglghkgBZQMEAS4wEQQMPurqRt-zQhuq4AoPAgEQgIIBVR85PML06Bpe-9S74tl5qFBdRfe3tC1ebYu9PKVKd-y1AYq1-rglf-UxOwEVQZIQ8GR1ZK-ZCqJl9GVINMdTkoNzgpcMJKDYA0_J-h3VMFGTnE8UT0vqdoLhiw8Zo3NpO2phDCxiHGPGzTehuiKVB5waHxyE_qwxht3_5l8iLNYCCkxlNbNeYhUa23LqQByVNiAoYo5TaGGgvvoM3AhKdHRiq4BPexQIzvEbXO31YjbcZBCJXydNEGSo3LNDGYGUXsUZIX8CPsAfifbW86OkWWZudxnPOAKqSUc4FVooUhB1y9JobgJX9bt92gCWQGj6u_nL21iO58UAxZC24_sxceMsa_mLDmdu4vb9P-4tEiwgdHTfGZb3chKokluZA1Lor-XWEBQ8-vjpuOMa5SPG78ydd7I3vVnNh5T9u_yaQxgsuaLdyUMsxsqSE14NHHBm7Y27yNmq. PubMed https://doi.org/10.1093/jamia/ocx059

50. Botticello AL, et al. 2016. To What Extent Do Neighborhood Differences Mediate Racial Disparities in Participation After Spinal Cord Injury? *Arch Phys Med Rehabil*. 97(10), 1735-44. doi:https://doi.org/10.1016/j.apmr.2016.04.007. PubMed

51. Dai H, Hao J, Catley D. 2017. Vape shop density and socio-demographic disparities: A US census tract analysis. *Nicotine Tob Res*. 19(11), 1338-44. https://academic.oup.com/ntr/article-abstract/19/11/1338/3074901?redirectedFrom=fulltext. PubMed https://doi.org/10.1093/ntr/ntx063

52. Drewnowski A, et al. 2016. Geographic disparities in Healthy Eating Index scores (HEI-2005 and 2010) by residential property values: Findings from Seattle Obesity Study (SOS). *Prev Med*. 83, 46-55. doi:https://doi.org/10.1016/j.ypmed.2015.11.021. PubMed

53. Gebreab SY, et al. 2017. Neighborhood social and physical environments and type 2 diabetes mellitus in African Americans: The Jackson Heart Study. *Health Place*. 43, 128-37. doi:https://doi.org/10.1016/j.healthplace.2016.12.001. PubMed

54. Henly S, et al. 2017. Disparities in digital reporting of illness: A demographic and socioeconomic assessment. *Prev Med*. 101, 18-22. https://ac.els-cdn.com/S009174351730172X/1-s2.0-S009174351730172X-main.pdf?_tid=d67b9495-3e6e-4e67-888f-0b9fbae3ea58&acdnat=1528587204_2aecc5dd8407da2f6faa7954b40aaee4. PubMed https://doi.org/10.1016/j.ypmed.2017.05.009

55. Hosgood HD, et al. 2013. Spatial and temporal distributions of lung cancer histopathology in the state of Maine. *Lung Cancer*. 82(1), 55-62. https://www.lungcancerjournal.info/article/S0169-5002(13)00301-2/fulltext. PubMed https://doi.org/10.1016/j.lungcan.2013.06.018

56. Insaf TZ, Talbot T. 2016. Identifying areas at risk of low birth weight using spatial epidemiology: A small area surveillance study. *Prev Med*. 88, 108-14. https://ac.els-cdn.com/S0091743516300457/1-s2.0-S0091743516300457-main.pdf?_tid=87f94b9f-b195-47e4-8766-b32263b73025&acdnat=1528587218_d4aa221d0ef92ab6849b8d604b51484b. PubMed https://doi.org/10.1016/j.ypmed.2016.03.019

57. Kramer MR, Dunlop AL, Hogue CJ. 2014. Measuring women's cumulative neighborhood deprivation exposure using longitudinally linked vital records: a method for life course MCH research. *Matern Child Health J*. 18(2), 478-87. doi:https://doi.org/10.1007/s10995-013-1244-7. PubMed

58. Leach HJ, et al. 2016. An exploratory decision tree analysis to predict cardiovascular disease risk in African American women. *Health Psychol*. 35(4), 397-402. doi:https://doi.org/10.1037/hea0000267. PubMed

59. Lee, J.E., J.H. Sung, and M. Malouhi, *Statistical Validation of a Web-Based GIS Application and Its Applicability to Cardiovascular-Related Studies.* Int J Environ Res Public Health, 2015. **13**(1): p. ijerph13010002. https://doi.org/10.3390/ijerph13010002

60. Nguyen QC, et al. 2017. Geotagged US Tweets as Predictors of County-Level Health Outcomes, 2015–2016. *Am J Public Health*. 107(11), 1776-82. PubMed https://doi.org/10.2105/AJPH.2017.303993

61. Noyes P, et al. 2014. Cycling in the city: an in-depth examination of bicycle lane use in a low-income urban neighborhood. *J Phys Act Health*. 11(1), 1-9. doi:https://doi.org/10.1123/jpah.2011-0429. PubMed

62. Pan I, et al. 2017. Machine learning for social services: A study of prenatal case management in Illinois. *Am J Public Health*. 107(6), 938-44. PubMed https://doi.org/10.2105/AJPH.2017.303711

63. Piccolo RS, et al. 2015. The role of neighborhood characteristics in racial/ethnic disparities in type 2 diabetes: results from the Boston Area Community Health (BACH) Survey. *Soc Sci Med*. 130, 79-90. doi:https://doi.org/10.1016/j.socscimed.2015.01.041. PubMed

64. Prussing C, et al. 2013. Geo-epidemiologic and molecular characterization to identify social, cultural, and economic factors where targeted tuberculosis control activities can reduce incidence in Maryland, 2004-2010. *Public Health Rep*. 128, 104-114. PubMed https://doi.org/10.1177/00333549131286S314

65. Salow AD, et al. 2018. Associations of neighborhood-level racial residential segregation with adverse pregnancy outcomes. *Am J Obstet Gynecol*. 218(3), 351.e1-7. doi:https://doi.org/10.1016/j.ajog.2018.01.022. PubMed

66. Sharifi M, et al. 2016. The role of neighborhood characteristics and the built environment in understanding racial/ethnic disparities in childhood obesity. *Prev Med*. 91, 103-09. doi:https://doi.org/10.1016/j.ypmed.2016.07.009. PubMed

67. Shim JM, Shin E, Johnson TP. 2013. Self-rated health assessed by web versus mail modes in a mixed mode survey: the digital divide effect and the genuine survey mode effect. *Med Care*. 51(9), 774-81. doi:https://doi.org/10.1097/MLR.0b013e31829a4f92. PubMed

68. Sills MR, et al. 2017. Adding Social Determinant Data Changes Children's Hospitals' Readmissions Performance. *J Pediatr*. 186, 150-157.e1. doi:https://doi.org/10.1016/j.jpeds.2017.03.056. PubMed

69. Silverman BG, et al. 2015. A systems approach to healthcare: Agent-based modeling, community mental health, and population well-being. *Artif Intell Med*. 63(2), 61-71. doi:https://doi.org/10.1016/j.artmed.2014.08.006. PubMed

70. Threatt, T.B. and E.D. Ward, *Telehealth for diabetes self-management education and support in an underserved, free clinic population: A pilot study*. J Am Pharm Assoc (2003), 2017. **57**(3): p. 402-406. https://doi.org/10.1016/j.japh.2017.01.019

71. Tomayko EJ, et al. 2015. Linking electronic health records with community-level data to understand childhood obesity risk. *Pediatr Obes*. 10(6), 436-41. doi:https://doi.org/10.1111/ijpo.12003. PubMed

72. Zenk SN, et al. 2014. Food shopping behaviours and exposure to discrimination. *Public Health Nutr*. 17(5), 1167-76. doi:https://doi.org/10.1017/S136898001300075X. PubMed

73. Asada Y, Yoshida Y, Whipp AM. 2013. Summarizing social disparities in health. *Milbank Q*. 91(1), 5-36. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3607125/pdf/milq0091-0005.pdf. PubMed https://doi.org/10.1111/milq.12001

74. Castrucci BC, et al. 2015. What gets measured gets done: an assessment of local data uses and needs in large urban health departments. *J Public Health Manag Pract*. 21(Suppl 1), S38. PubMed https://doi.org/10.1097/PHH.0000000000000169

75. Cheng ER, et al. 2016. Association of missing paternal demographics on infant birth certificates with perinatal risk factors for childhood obesity. *BMC Public Health*. 16, 453. doi:https://doi.org/10.1186/s12889-016-3110-1. PubMed

76. Flood TL, et al. 2015. Electronic health records and community health surveillance of childhood obesity. *Am J Prev Med*. 48(2), 234-40. doi:https://doi.org/10.1016/j.amepre.2014.10.020. PubMed

77. French DD, et al. 2017. Payment Reform Needed to Address Health Disparities of Undiagnosed Diabetic Retinopathy in the City of Chicago. *Ophthalmol Ther*. 6(1), 123-31. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5449291/pdf/40123_2016_Article_72.pdf. PubMed https://doi.org/10.1007/s40123-016-0072-4

78. Goldner M, et al. 2013. The intersection of gender and place in online health activities. *J Health Commun*. 18(10), 1235-55. doi:https://doi.org/10.1080/10810730.2013.778364. PubMed

79. Goyal NK, et al. 2015. Risk prediction for adverse pregnancy outcomes in a medicaid population. *J Womens Health (Larchmt)*. 24(8), 681-88. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4543486/pdf/jwh.2014.5069.pdf. PubMed https://doi.org/10.1089/jwh.2014.5069

80. Houle JN. 2014. Mental health in the foreclosure crisis. *Soc Sci Med*. 118, 1-8. doi:https://doi.org/10.1016/j.socscimed.2014.07.054. PubMed

81. Jones A. 2013. Segregation and cardiovascular illness: the role of individual and metropolitan socioeconomic status. *Health Place*. 22, 56-67. doi:https://doi.org/10.1016/j.healthplace.2013.02.009. PubMed

82. Kasthurirathne SN, et al. 2018. Assessing the capacity of social determinants of health data to augment predictive models identifying patients in need of wraparound social services. *J Am Med Inform Assoc*. 25(1), 47-53. doi:https://doi.org/10.1093/jamia/ocx130. PubMed

83. Pearson-Stuttard, J., et al., *Reducing US cardiovascular disease burden and disparities through national and targeted dietary policies: A modelling study*. PLoS Medicine, 2017. **14**(6 Article Number).

84. Roth C, et al. 2014. Community-level determinants of obesity: harnessing the power of electronic health records for retrospective data analysis. *BMC Med Inform Decis Mak*. 14, 36. doi:https://doi.org/10.1186/1472-6947-14-36. PubMed

85. Steiner JF, et al. 2018. Food Insecurity in Older Adults in an Integrated Health Care System. *J Am Geriatr Soc*., 1017. PubMed https://doi.org/10.1111/jgs.15285

Supplementary Material

Supplemental Table 1. Medline Search (via PubMed)

Search no.	**Facet**	**Search terms**	**Search results (May 16, 2018)**
1	Identify studies on social determinants of health.	"Social Determinants of Health" [MeSH] OR "social determinants of health"[tiab] OR "health status disparities"[MeSH] OR "health status disparities"[tiab] OR "health equity"[MeSH] OR "health equity"[tiab] OR "social determinant" OR "social determinants" OR "health disparity" OR "health disparities"	28,882
2		"tobacco use"[tiab] OR "alcohol and illicit drug use"[tiab] OR "physical activity"[tiab] OR "diet"[tiab] OR "obesity"[tiab] OR "health literacy"[MeSH] OR "health literacy"[tiab] OR "early intervention education"[MeSH] OR "educational status"[MeSH] OR "educational status"[tiab] OR "high school graduation"[tiab] OR "early childhood education"[tiab] OR "language proficiency"[tiab] OR "employment"[tiab] OR "unemployment"[tiab] OR "socioeconomic factors"[MeSH] OR "socioeconomic factors"[tiab] OR "socioeconomic disparities"[tiab]	1,497,418

		OR “income”[tiab] OR “social support”[tiab] OR “built environment”[tiab] OR “food insecurity”[tiab] OR “food security”[tiab] OR “access to healthy food”[tiab] OR incarcerat*[tiab] OR “crime”[tiab] OR “violence”[tiab] OR “civil rights”[MeSH] OR “civil rights”[tiab] OR “civic participation”[tiab] OR “gender”[tiab] OR “discrimination”[tiab] OR “walkability”[tiab] OR “housing instability”[tiab] OR “quality of housing”[tiab] OR “environmental health”[MeSH] OR “environmental health”[tiab] OR “environmental conditions”[tiab] OR “transportation”[tiab] OR “urbanization”[tiab] OR “air quality”[tiab]	
3		#1 AND #2	15,755
4		“social determinants of health”[tiab] OR "social determinants of health"[MeSH]	3,913
5		#3 OR #4	17,386
6	Identify studies that use data analytics.	“analytic*”[tiab] OR “database*”[tiab] OR “data interpretation, statistical”[MeSH] OR “data analysis”[tiab] OR “data mining”[tiab] OR “electronic health records” [MeSH] OR “electronic health record”[tiab] OR “personal health record”[tiab] OR “data set*”[tiab] OR “survey*”[tiab] OR “standards-based assessment*” OR “claims data”[tiab] OR “electronic medical record”[tiab]	861,933
7	Identify studies that use data analytics and report on social determinants of health.	#5 AND #6	4058
8	Exclude non-US studies.	#7 NOT (“Australia”[Mesh] OR “Canada”[Mesh] OR “Mexico”[Mesh] OR “Europe”[Mesh] OR “China”[Mesh] OR “Russia”[Mesh] OR “Africa”[Mesh] OR “Asia”[Mesh] OR “South America”[Mesh] OR “iran”[tiab] OR “Africa”[tiab])	2,409
9	Exclude *in vivo* and *in vitro* studies.	#8 NOT (“*in vivo*” OR “*in vitro*”)	2,408
10		Review[pt] NOT (Cochrane OR systematic or meta-analy*)	2,129,768

11	Exclude non-systematic reviews.	#9 NOT #10	2,374
12	Exclude other inappropriate study designs.	#11 NOT ("case reports"[pt] OR "case report" OR "case series" OR "editorial"[pt] OR "letter"[pt])	2,367
13	Exclude studies with a global perspective.	#12 NOT "global health"[tiab]	2340
14	Filter for human studies.	#13 AND [humans]	1859
15	Filter for publications with abstracts.	#14 AND [abstract]	1848
16	Filter by publication date (last five years).	#15 AND 2013-2018 [dp]	1032
17	Filter for publications written in English.	#16 AND [English]	1029

Supplemental Table 2. Cochrane Library Search

Search no.	Facet	Search terms	Search results (May 18, 2018)
1	Identify studies on social determinants of health.	Social determinants of health[Mesh] OR "social determinants of health" OR Health status disparities[Mesh] OR "health status disparities" OR Health equity[Mesh] OR "health equity" OR "Social determinant*" OR "health disparity" OR "health disparities"	1057
2		"tobacco use" OR "alcohol and illicit drug use" OR "physical activity" OR "diet" OR "obesity" OR "health literacy"[MeSH] OR "health literacy" OR "early intervention education"[MeSH] OR "educational status"[MeSH] OR "educational status" OR "high school graduation" OR "early childhood education" OR "language proficiency" OR "employment" OR	123,755

		"unemployment" OR "socioeconomic factors"[MeSH] OR "socioeconomic factors" OR "socioeconomic disparities" OR "income" OR "social support" OR "built environment" OR "food insecurity" OR "food security" OR "access to healthy food" OR incarcerat* OR "crime" OR "violence" OR "civil rights" OR "civil rights" OR "civic participation" OR "gender" OR "discrimination" OR "walkability" OR "housing instability" OR "quality of housing" OR "environmental health"[MeSH] OR "environmental health" OR "environmental conditions" OR "transportation" OR "urbanization" OR "air quality"	
3		#1 AND #2	646
4		Social determinants of health[Mesh] OR "social determinants of health"	150
5		#3 OR #4	698
6	Identify studies that use data analytics.	"analytic*" OR "database*" OR "data interpretation, statistical"[MeSH] OR "data analysis" OR "data mining" OR "electronic health records" [MeSH] OR "electronic health record" OR "personal health record" OR "data set*" OR "survey*" OR "standards-based assessment*" OR "claims data" OR "electronic medical record"	124,309
7	Identify studies that use data analytics and report on social determinants of health.	#5 AND #6	226
8	Exclude *in vivo* and *in vitro* studies.	#7 NOT ("*in vivo*" OR "*in vitro*")	219
9	Filter by publication date (last five years).	#8 AND (2013-2018)	126
10*	Filter to include only trials.	#9 AND trials	71
11*	Filter to include only technology assessments.	#9 AND (tech assessments)	4

12*	Filter to include systematic reviews.	#9 AND (reviews)	51

*Note: Rows 10, 11, and 12 all appear in the final results of the search; a total of 126 records were identified from the Cochrane Library search.

Supplemental Table 3. Embase Search

Search no.	**Facet**	**Search terms**	**Search results (May 18, 2018)**
1	Identify studies on social determinants of health.	'social determinants of health'/exp OR 'social determinants of health':ab,ti OR 'health disparity'/exp OR 'health disparit*':ab,ti OR 'health equity'/exp OR 'health equity':ab,ti OR 'social determinant*'	29,137
2		'tobacco use'/exp OR 'tobacco use':ab,ti OR 'alcohol and illicit drug use':ab,ti OR 'physical activity'/exp OR 'physical activity':ab,ti OR 'diet'/exp OR 'diet':ab,ti OR 'obesity'/exp OR 'obesity':ab,ti OR 'health literacy'/exp OR 'health literacy':ab,ti OR 'early childhood intervention'/exp OR 'educational status'/exp OR 'educational status':ab,ti OR 'high school graduate':ab,ti OR 'early childhood education':ab,ti OR 'language ability'/exp OR 'language proficiency':ab,ti OR 'employment':ab,ti OR 'unemployment':ab,ti OR 'socioeconomics'/exp OR 'socioeconomic factors':ab,ti OR 'socioeconomic disparities':ab,ti OR 'income':ab,ti OR 'social support':ab,ti OR 'built environment':ab,ti OR 'food insecurity'/exp OR 'food security':ab,ti OR 'access to healthy food':ab,ti OR 'incarcerat*':ab,ti OR 'crime':ab,ti OR 'violence':ab,ti OR 'civil rights'/exp OR 'civil rights':ab,ti OR 'civic participation':ab,ti OR 'gender':ab,ti OR 'discrimination':ab,ti OR 'walkability':ab,ti OR 'housing instability':ab,ti OR 'quality of housing'/ab,ti OR 'environmental health'/exp OR 'environmental health':ab,ti OR 'environmental conditions':ab,ti OR 'transportation':ab,ti OR 'urbanization':ab,ti OR 'air quality':ab,ti	2,654,021
3		#1 AND #2	15,802
4		'social determinants of health'/exp OR 'social determinants of health':ab,ti	5,326
5		#3 OR #4	18,046

6	Identify studies that use data analytics.	‘analytic*’:ab,ti OR ‘database’:ab,ti OR ‘statistical analysis’/exp OR ‘data analysis’:ab,ti OR ‘data mining’:ab,ti OR ‘electronic health record’/exp OR ‘electronic health record’:ab,ti OR ‘personal health record’:ab,ti OR ‘data set*’:ab,ti OR ‘survey’:ab,ti OR ‘standards-based assessment*’:ab,ti OR ‘claims data’:ab,ti OR ‘electronic medical record’/exp OR ‘electronic medical record’:ab,ti	3,029,553
7	Identify studies that use data analytics and report on social determinants of health.	#5 AND #6	5,946
8	Exclude non-US studies.	#7 NOT (‘Australia and New Zealand’/exp OR ‘Canada’/exp OR ‘Mexico’/exp OR ‘Europe’/exp OR ‘China’/exp OR ‘Russia’/exp OR ‘Africa’/exp OR ‘Asia’/exp OR ‘South America’/exp OR ‘Iran’/exp)	3,574
9	Exclude *in vivo* and *in vitro* studies.	#8 NOT (‘*in vivo*’ OR ‘*in vitro’*)	3,567
10	Exclude non-systematic reviews.	[review]/lim NOT (Cochrane OR systematic OR ‘meta-analy*’)	2,275,010
11		#9 NOT #10	3,471
12	Exclude other inappropriate study designs.	‘case report’/exp OR ‘case report*’ OR ‘case study’/exp OR ‘case series’	2,460,562
13		#11 NOT #12	3,457
14	Exclude studies with a global perspective.	#13 NOT (‘global health’:ab,ti)	3,422
15	Filter for human studies.	#14 AND [humans]/lim	3,295
16	Filter for publications with abstracts.	#15 AND [abstracts]/lim	3,235
17	Filter by publication date (last 5 years).	#16 AND [2013-20180/py	2,216

18	Filter for publications written in English.	#17 AND [English]/lim	2,195

Supplemental Table 4. Medline Search (via PubMed)

Search no.	**Facet**	**Search terms**	**Search results (May 23, 2018)**
1	Identify articles on social determinants of health.	"Social Determinants of Health" [MeSH] OR "social determinants of health"[tiab] OR "health status disparities"[MeSH] OR "health status disparities"[tiab] OR "health equity"[MeSH] OR "health equity"[tiab] OR "social determinant" OR "social determinants" OR "health disparity" OR "health disparities" OR healthcare disparities[MeSH] OR minority health[MeSH]	38,924
2	Identify articles on health policy.	Health Policy[MeSH] OR health policy[tw] OR healthcare policy[tw]	109,876
3	Identify studies that use digital health technology.	Electronic Health Records[Mesh] OR "electronic health record*"[tiab] OR "electronic registry"[tiab] OR "electronic registries"[tiab] OR "Health Information Exchange"[Mesh] OR "health information exchange*"[tiab] OR "Decision Support Systems, Clinical"[Mesh] OR "clinical decision support system*"[tiab] OR "Social Networking"[Mesh] OR "social networking"[tiab] OR "Health Records, Personal"[Mesh] OR "personal health record*"[tiab] OR "electronic health information"[tiab] OR "electronic health communication*"[tiab] OR "patient monitor*"[tiab] OR "Patient Portals"[Mesh] OR "patient portal*"[tiab] OR "Telemedicine"[Mesh] OR "telemedicine"[tiab] OR telehealth[tiab] OR electronic medical record[tiab] OR data analytics[tw] OR data mining[tiab] OR data interpretation, statistical[MeSH] OR data analysis[tiab]	171,779
4	Identify policy studies that use digital health technology and report on social	#1 AND #2 AND #3	39

	determinants of health.		
5	Exclude non-US studies.	#4 NOT (Australia[Mesh] OR Canada[Mesh] OR Mexico[Mesh] OR Europe[Mesh] OR China[Mesh] OR Russia[Mesh] OR Africa[Mesh] OR Asia[Mesh] OR South America[Mesh] OR Iran[tiab] OR Canada[tiab] OR New Zealand[tiab] OR United Kingdom[tiab] OR Great Britain[tiab])	24
6	Exclude studies with a global perspective.	#5 NOT (global health[MeSH] OR global health[tiab]	23
7	Filter for publications with abstracts.	#6 AND [abstract]	22
8	Filter by publication date (last 5 years).	#7 AND 2013-2018 [dp]	16
9	Filter for publications written in English.	#8 AND [English]	16

Supplemental Table 5. Medline Search (via PubMed)

Search no.	**Facet**	**Search terms**	**Search results (May 31, 2018)**
1	Identify articles on social determinants of health.	Search ("Social Determinants of Health" [MeSH] OR "social determinants of health"[tiab] OR "health status disparities"[MeSH] OR "health status disparities"[tiab] OR "health equity"[MeSH] OR "health equity"[tiab] OR "social determinant" OR "social determinants" OR "health disparity" OR "health disparities")	29,065

2		Search ("tobacco use"[tiab] OR "alcohol and illicit drug use"[tiab] OR "physical activity"[tiab] OR "diet"[tiab] OR "obesity"[tiab] OR "health literacy"[MeSH] OR "health literacy"[tiab] OR "early intervention education"[MeSH] OR "educational status"[MeSH] OR "educational status"[tiab] OR "high school graduation"[tiab] OR "early childhood education"[tiab] OR "language proficiency"[tiab] OR "employment"[tiab] OR "unemployment"[tiab] OR "socioeconomic factors"[MeSH] OR "socioeconomic factors"[tiab] OR "socioeconomic disparities"[tiab] OR "income"[tiab] OR "social support"[tiab] OR "built environment"[tiab] OR "food insecurity"[tiab] OR "food security"[tiab] OR "access to healthy food"[tiab] OR incarcerat*[tiab] OR "crime"[tiab] OR "violence"[tiab] OR "civil rights"[MeSH] OR "civil rights"[tiab] OR "civic participation"[tiab] OR "gender"[tiab] OR "discrimination"[tiab] OR "walkability"[tiab] OR "housing instability"[tiab] OR "quality of housing"[tiab] OR "environmental health"[MeSH] OR "environmental health"[tiab] OR "environmental conditions"[tiab] OR "transportation"[tiab] OR "urbanization"[tiab] OR "air quality"[tiab])	1,501,371
3		Search (#1 AND #2)	15,845
4		Search ("social determinants of health"[tiab] OR "social determinants of health"[MeSH])	3,947
5		Search (#3 OR #4)	17,493

6	Identify studies that use digital health technology.	Search ("Artificial Intelligence"[Mesh] OR "artificial intelligence"[tiab] OR "machine intelligence"[tiab] OR "computational intelligence"[tiab] OR "Machine Learning"[Mesh] OR "machine learning"[tiab] OR "machine-learning"[tiab] OR "Computer Security"[Mesh] OR "data security"[tiab] OR "cybersecurity"[tiab] OR "cyber security"[tiab] OR "data protect*"[tiab] OR "data encrypt*"[tiab] OR "Cloud Computing"[Mesh] OR "cloud computing"[tiab] OR "cloud process*"[tiab] OR "cognitive comput*"[tiab] OR "Patient Portals"[Mesh] OR "patient web portal*"[tiab] OR "patient web-portal*"[tiab] OR "patient portal*"[tiab] OR "web portal*"[tiab] OR "mobile technolog*"[tiab] OR "Telemedicine"[Mesh] OR "telemedicine"[tiab] OR "telehealth*"[tiab] OR "mobile health"[tiab] OR "mHealth"[tiab] OR "eHealth"[tiab] OR "m-Health"[tiab] OR "mobile-health"[tiab] OR "telecommunication*"[tiab] OR "Decision Support Systems, Clinical"[Mesh] OR "clinical decision support"[tiab] OR "decision support system"[tiab])	131,429
7	Identify studies that use digital health technology and report on social determinants of health.	Search (#5 AND #6)	112
8	Exclude non-US studies.	Search (#7 NOT ("Australia"[Mesh] OR "Canada"[Mesh] OR "Mexico"[Mesh] OR "Europe"[Mesh] OR "China"[Mesh] OR "Russia"[Mesh] OR "Africa"[Mesh] OR "Asia"[Mesh] OR "South America"[Mesh] OR "iran"[tiab] OR "Africa"[tiab]))	102
9	Exclude *in vivo* and *in vitro* studies.	Search (#8 NOT ("in vivo" OR "in vitro"))	102
10	Exclude non-systematic reviews.	Search (Review[pt] NOT (Cochrane OR systematic or meta-analy*))	2,133,577
11		Search (#9 NOT #10)	93
12	Exclude other inappropriate study designs.	Search (#11 NOT ("case reports"[pt] OR "case report" OR "case series" OR "editorial"[pt] OR "letter"[pt]))	90

13	Exclude studies with a global perspective.	Search (#12 NOT "global health"[tiab])	89
14	Filter for human studies.	Search (#12 NOT "global health"[tiab]) Filters: Humans	51
15	Filter for publications with abstracts.	Search (#12 NOT "global health"[tiab]) Filters: Abstract; Humans	49
16	Filter by publication date (last 5 years).	Search (#12 NOT "global health"[tiab]) Filters: Abstract; Publication date from 2013/01/01; Humans	37
17	Filter for publications written in English.	Search (#12 NOT "global health"[tiab]) Filters: Abstract; Publication date from 2013/01/01; Humans; English	37

Supplemental Table 6. Cochrane Library Search

Search no.	Facet	Search Terms	Search Results (June 1, 2018)
1	Identify articles on social determinants of health.	Social determinants of health[Mesh] OR “social determinants of health”:ti,ab,kw OR Health status disparities[Mesh] OR “health status disparities”:ti,ab,kw OR Health equity[Mesh] OR “health equity”:ti,ab,kw OR “Social determinant*”:ti,ab,kw OR “health disparity”:ti,ab,kw OR “health disparities”:ti,ab,kw	920
2		“tobacco use”:ti,ab,kw OR “alcohol and illicit drug use”:ti,ab,kw OR “physical activity”:ti,ab,kw OR “diet”:ti,ab,kw OR “obesity”:ti,ab,kw OR “health literacy”[MeSH] OR “health literacy”:ti,ab,kw OR “early intervention (education)”[MeSH] OR “educational status”[MeSH] OR “educational status”:ti,ab,kw OR “high school graduation”:ti,ab,kw OR “early childhood education”:ti,ab,kw OR “language proficiency”:ti,ab,kw OR “employment”:ti,ab,kw OR	121,365

		"unemployment":ti,ab,kw OR "socioeconomic factors"[MeSH] OR "socioeconomic factors":ti,ab,kw OR "socioeconomic disparities":ti,ab,kw OR "income":ti,ab,kw OR "social support":ti,ab,kw OR "built environment":ti,ab,kw OR "food insecurity":ti,ab,kw OR "food security":ti,ab,kw OR "access to healthy food":ti,ab,kw OR incarcerat*:ti,ab,kw OR "crime":ti,ab,kw OR "violence":ti,ab,kw OR "civil rights":ti,ab,kw OR "civic participation":ti,ab,kw OR "gender":ti,ab,kw OR "discrimination":ti,ab,kw OR "walkability":ti,ab,kw OR "housing instability":ti,ab,kw OR "quality of housing":ti,ab,kw OR "environmental health"[MeSH] OR "environmental health":ti,ab,kw OR "environmental conditions":ti,ab,kw OR "transportation":ti,ab,kw OR "urbanization":ti,ab,kw OR "air quality":ti,ab,kw	
3		#1 AND #2	533
4		Social determinants of health[Mesh] OR "social determinants of health":ti,ab,kw	83
5		#3 OR #4	551
6	Identify studies that use digital health technology.	"artificial intelligence"[MeSH] OR "artificial intelligence":ti,ab,kw OR "machine intelligence":ti,ab,kw OR "computational intelligence":ti,ab,kw OR "machine learning"[MeSH] OR "machine learning":ti,ab,kw OR "machine-learning":ti,ab,kw OR "computer security"[MeSH] OR "data security":ti,ab,kw OR "cybersecurity":ti,ab,kw OR "cyber security":ti,ab,kw OR "data protect*":ti,ab,kw OR "data encrypt*":ti,ab,kw OR "cloud computing"[MeSH] OR "cloud computing":ti,ab,kw OR "cloud process*":ti,ab,kw OR "cognitive comput*":ti,ab,kw OR "patient portals"[MeSH] OR "patient web portal*":ti,ab,kw OR "patient web-	7,930

		portal*":ti,ab,kw OR "patient portal*":ti,ab,kw OR "web portal*":ti,ab,kw OR "mobile technolog*":ti,ab,kw OR "telemedicine"[MeSH] OR "telemedicine":ti,ab,kw OR "telehealth*":ti,ab,kw OR "mobile health":ti,ab,kw OR "mHealth":ti,ab,kw OR "eHealth":ti,ab,kw OR "m-Health":ti,ab,kw OR "mobile-health":ti,ab,kw OR "telecommunication*":ti,ab,kw OR "Decision support systems, clinical"[MeSH] OR "clinical decision support":ti,ab,kw OR "decision support system":ti,ab,kw	
7	Identify studies that use digital health technology and report on social determinants of health.	#5 AND #6	17
8	Exclude *in vivo* and *in vitro* studies.	#7 NOT ("in vivo" OR "in vitro")	17
9	Filter by publication date (last 5 years).	#8 AND Publication Year from 2013 to 2018	15

Supplemental Table 7. Embase Search

Search no.	Facet	Search Terms	Search Results (June 4, 2018)
1	Identify articles on social determinants of health.	'social determinants of health'/exp OR 'social determinants of health':ab,ti OR 'health disparity'/exp OR 'health disparit*':ab,ti OR 'health equity'/exp OR 'health equity':ab,ti OR 'social determinant*'	29,426
2		'tobacco use'/exp OR 'tobacco use':ab,ti OR 'alcohol and illicit drug use':ab,ti OR 'physical activity'/exp OR 'physical activity':ab,ti OR 'diet'/exp OR 'diet':ab,ti OR 'obesity'/exp OR 'obesity':ab,ti OR 'health literacy'/exp OR 'health literacy':ab,ti OR 'early childhood intervention'/exp OR 'educational status'/exp OR 'educational status':ab,ti OR 'high school graduate':ab,ti OR 'early childhood education':ab,ti OR 'language ability'/exp OR 'language proficiency':ab,ti OR 'employment':ab,ti OR	2,741,489

		'unemployment':ab,ti OR 'socioeconomics'/exp OR 'socioeconomic factors':ab,ti OR 'socioeconomic disparities':ab,ti OR 'income':ab,ti OR 'social support':ab,ti OR 'built environment':ab,ti OR 'food insecurity'/exp OR 'food security':ab,ti OR 'access to healthy food':ab,ti OR 'incarcerat*':ab,ti OR 'crime':ab,ti OR 'violence':ab,ti OR 'civil rights'/exp OR 'civil rights':ab,ti OR 'civic participation':ab,ti OR 'gender':ab,ti OR 'discrimination':ab,ti OR 'walkability':ab,ti OR 'housing instability':ab,ti OR 'quality of housing':ab,ti OR 'environmental health'/exp OR 'environmental health':ab,ti OR 'environmental conditions':ab,ti OR 'transportation':ab,ti OR 'urbanization':ab,ti OR 'air quality':ab,ti	
3		#1 AND #2	15,949
4		'social determinants of health'/exp OR 'social determinants of health':ab,ti	5,439
5		#3 OR #4	18,251
6	Identify studies that use digital health technology.	'artificial intelligence'/exp OR 'artificial intelligence':ab,ti OR 'machine intelligence':ab,ti OR 'computational intelligence':ab,ti OR 'machine learning'/exp OR 'machine learning':ab,ti OR 'machine-learning':ab,ti OR 'computer security'/exp OR 'data security':ab,ti OR 'cybersecurity':ab,ti OR 'cyber security':ab,ti OR 'data protect*':ab,ti OR 'data encrypt*':ab,ti OR 'cloud computing'/exp OR 'cloud computing':ab,ti OR 'cloud process*':ab,ti OR 'cognitive comput*':ab,ti OR 'medical record'/exp OR 'patient web portal*':ab,ti OR 'patient web-portal*':ab,ti OR 'patient portal*':ab,ti OR 'web portal*':ab,ti OR 'mobile technolog*':ab,ti OR 'telemedicine'/exp OR 'telemedicine':ab,ti OR 'telehealth*':ab,ti OR 'mobile health':ab,ti OR 'mHealth':ab,ti OR 'eHealth':ab,ti OR 'm-Health':ab,ti OR 'mobile-health':ab,ti OR 'telecommunication*':ab,ti OR 'clinical	393,620

		decision support system'/exp OR 'clinical decision support':ab,ti OR 'decision support system':ab,ti	
7	Identify studies that use digital health technology and report on social determinants of health.	#5 AND #6	398
8	Exclude non-US studies.	#7 NOT ('Australia and New Zealand'/exp OR 'Canada'/exp OR 'Mexico'/exp OR 'Europe'/exp OR 'China'/exp OR 'Russia'/exp OR 'Africa'/exp OR 'Asia'/exp OR 'South America'/exp OR 'Iran'/exp)	330
9	Exclude *in vivo* and *in vitro* studies.	#8 NOT ('in vivo' OR 'in vitro')	330
10	Exclude non-systematic reviews.	[review]/lim NOT (Cochrane OR systematic OR 'meta-analy*')	2,270,891
11		#9 NOT #10	310
12	Exclude other inappropriate study designs.	'case report'/exp OR 'case report*' OR 'case study'/exp OR 'case series'	2,495,228
13	Exclude studies with a global perspective.	#11 NOT #12	306
14		#13 NOT ('global health':ab,ti)	305
15	Filter for human studies.	#14 AND [humans]/lim	295
16	Filter for publications with abstracts.	#15 AND [abstracts]/lim	263
17	Filter by publication date (last 5 years).	#16 AND [2013-2018]/py	212
18	Filter for publications written in English.	#17 AND [English]/lim	211

Supplemental Table 8. Inclusion/Exclusion Criteria for Large, Population-Based Studies

Inclusion Criteria	Exclusion Criteria
The publication is in English.	The publication is in a language other than English.
The publication contains an abstract.	The publication does not contain an abstract.

The publication date is within the last five years (2013-2018).	The publication date is not within the last five years.
The publication describes a **primary study** (prospective or retrospective; observational or experimental; comparative or non-comparative), or **systematic review** with or without meta-analyses (if systematic review is of only United States studies).	The publication describes a study design other than primary study or systematic review (eg, case report, case series, editorial, narrative review).
The publication describes a study that takes place in the United States or pertains to policy in the United States.	The publication describes a study that takes place or pertains to policy in a country other than the United States.
The publication focuses on a social, behavioral, or environmental determinant of health factor (as predictors).	The publication is not focused on a social, behavioral, or environmental determinant of health factor; the publication is focused on access to care or clinical care without addressing a social, behavioral, or environmental determinant.
The publication reports an association of factors related to health outcomes (allows for both positive, null, and negative associations).	The publication does not report a health-related outcome.
The publication reports a NEW data source (electronic health records (EHRs), claims data, administrative data, geocoded data (GIS), geographic data, personal health records) that may use tools (surveys, scales, assessments, questionnaires, etc.) in a large population.	The publication reports a previously identified data source (Behavioral Risk Factor Surveillance System, National Health Interview Survey, American Census, American Community Survey, National Health and Nutrition Examination Survey) or uses tools (surveys, scales, assessments, questionnaires, etc.) in a small population.
The study reports rigorous data: the data set has a sample size of greater than 10,000 and provides granularity (meaning the authors provide details about the levels of data utilized).	The study does not report rigorous data: The data set has a sample size less than 10,000 or does not provide granularity.

Supplemental Table 9. Inclusion/Exclusion Criteria for the Smaller Studies of Actionable SDoH Factors

Inclusion Criteria	**Exclusion Criteria**
The publication is in English.	The publication is in a language other than English.
The publication contains an abstract.	The publication does not contain an abstract.
The publication date is within the last five years (2013-2018).	The publication date is not within the last five years.

The publication describes a **primary study** (prospective or retrospective; observational or experimental; comparative or non-comparative), or **systematic review** with or without meta-analyses (if systematic review is of only United States studies).	The publication describes a study design other than primary study or systematic review (eg, case report, case series, editorial, narrative review).
The publication describes a study that takes place in the United States or pertains to policy in the United States.	The publication describes a study that takes place or pertains to policy in a country other than the United States.
The publication focuses on an **actionable** social, behavioral, or environmental determinant of health factor (as predictors). *E.g., housing and stability, transportation, early education, utility assistance, interpersonal safety, social support, and food insecurity.*	The publication is not focused on a social, behavioral, or environmental determinant of health factor; the publication is focused on access to care or clinical care without addressing a social, behavioral, or environmental determinant.
The publication reports an association of **actionable** factors related to health outcomes (allows for both positive, null, and negative associations).	The publication does not report a health-related outcome.
The publication uses NOVEL data sources and/or digital health technology (electronic health records, artificial intelligence, machine-learning, advanced analytics, patient portals, national surveys, insurance claims data, advanced analytics, etc.).	The publication does not use data sources or technology or describes data already identified by determinantsofhealth.org.

Supplemental Table 10. Inclusion/Exclusion Criteria for Policy Pieces (including grey literature)

Inclusion Criteria	Exclusion Criteria
The publication is in English.	The publication is in a language other than English.
The publication date is within the last five years (2013-2018).	The publication date is not within the last five years.
The publication describes a **policy piece**.	The publication does not describe a policy piece.
The publication describes a study that takes place in the United States or pertains to policy in the United States.	The publication describes a study that takes place or pertains to policy in a country other than the United States.

The publication focuses on a social, behavioral, or environmental determinant of health factor (as predictors).	The publication is not focused on a social, behavioral, or environmental determinant of health factor; the publication is focused on access to care or clinical care without addressing a social, behavioral, or environmental determinant.
The publication addresses the impact of digital health technology (electronic health records, artificial intelligence, machine-learning, advanced analytics, patient portals) on social determinants of health.	The publication does not address the impact of technology on social determinants of health.

Note that priority will be given to more recently published articles in higher-tier journals, with longer follow-up, larger sample sizes, and greater generalizability.

Supplemental Table 11. Policy articles

Study	Design	Setting	Geographic Location	Policy Under Study	Author Conclusions
Cahill et al. (2016)[19]	N/A	N/A	US	Federal guideline requiring sexual orientation and gender identity data in EHRs	"Although the recent ONC final rule presents an important opportunity, we must still research ways to implement SO/GI questions looking at both consumer and provider perspectives to ensure that the data are collected correctly and used to provide quality care that meets the unique health needs of each of the constituent populations within the LGBT community. It is essential that clinical staff be trained in how to collect and use the data to improve quality of care and better document, address, and reduce LGBT health disparities—including risk behaviors, low rates of accessing preventive screenings, disease burden, and treatment outcomes. The organizations we represent—The Fenway Institute, CAP, and the Center of Excellence for Transgender Health at the University of California at San Francisco—offer resources for providers and

					clinical staff seeking to implement SO/GI questions in their EHRs and to use the data to improve patient care."
Chakkalakal et al. (2014)[13]	Retrospective	Clinical	Southern US	Existence of disparities in technology access and increasing access to and use of technology	"Increased reliance on technology to promote patient health may have limited value if certain groups lack access and/or the skills to leverage these tools. Efforts are needed to engage individuals with limited HL in the development of technology-based interventions that they would use."
DeSalvo et al. (2017)[17]	N/A	Community	US	Public Health 3.0 recommendations	"With the Public Health 3.0 framework, we envision a strong local public health infrastructure in all communities and its leaders serving as Chief Health Strategists that partner with stakeholders across a multitude of sectors on the ground to address the social determinants of health. With equity and social determinants of health as guiding principles, every person and every organization can take shared accountability to ensure

					the conditions in which everyone can be healthy regardless of race, ethnicity, gender identity, sexual orientation, geography, or income level. If successful, such transformation can form the foundation from which we build an equitable health-promoting system — in which stable, safe, and thriving community is a norm rather than an aberration. The Public Health 3.0 initiative seeks to inspire transformative success stories such as those already witnessed in many pioneering communities across the country. The challenge now is to institutionalize this expanded approach to community based public health practice and replicate these triumphs across all communities, for the health of all people."
Douglas et al. (2015)[23]	N/A	Community	US	Health Information Technology for Economic and Clinical Health (HITECH) Act	"The use of EHRs to identify and reduce health disparities is promising, but limited by the type of demographic data that is currently collected. To recognize HITECH's policy priority of reducing health disparities, more granular race and ethnicity data, disability status, and sexual

					orientation and gender identity must be collected in EHRs."
Graham et al. (2016)[12]	N/A	Community	US	Incorporation of technology into addressing health inequities	"When combined with sound policy strategies, emerging, scalable, digital technologies will likely become powerful allies for improving health and reducing health disparities."
Krumholz et al. (2016)[24]	N/A	N/A	US	Policies related to enabling data sharing to fuel a learning health system	“At this vital juncture in health care and research, the secure sharing of data has great potential. However, achievement of such a grand strategy for change will require unprecedented levels of collaboration among and communication between all stakeholders in the health system, and systems to evaluate effects and iterate for improvement.”
National Academy of Medicine (2014)[27]	N/A	N/A	US	sharing and privacy concerns of SDoH in EHR	"These “psychosocial vital signs” include four measures that are already widely collected (race/ethnicity, tobacco use, alcohol use, and residential address) and eight additional measures (education, financial resource strain, stress, depression, physical activity,

					social isolation, intimate partner violence, and neighborhood median household income). While recognizing the additional time needed to collect such data and act upon it, the committee concluded that the health benefits of addressing these determinants outweigh the added burden to providers, patients, and health care systems."
Penman-Aguilar et al. (2016)[20]	N/A	N/A	US	Recommendations for measuring SDoH and inequalities	"Over the decades since the clarion call for health equity was raised domestically by the Secretary's Task Force Report on Black and Minority Health and internationally by the writings of Margaret Whitehead and others, much has been learned about how to measure health disparities, health inequities, and social determinants of health at the national level to support the advancement of health equity. Nevertheless, there is still much to learn and implement, and the challenges of health equity persist. As the field of health equity continues to evolve, we anticipate that the present discussion will contribute to the laying of a foundation for

					standard practice in the monitoring of national progress toward achievement of health equity.”
Perlin et al. (2016)[25]	N/A	N/A	US	Policies related to enabling data sharing to gain economic and clinical benefits of EHR	Key points from summary recommendations for vital directions: “1. Commit to end-to-end interoperability extending from devices to EHR systems. 2. Aggressively address cyber security vulnerability. 3. Develop a data strategy that supports a learning health system.”
Ray et al. (2017)[11]	Retrospective	Community	US	use of mobile technology for health policy (develop health literacy, improve health outcomes, and reduce health disparities)	“Blacks and Latinos, compared to whites, were more likely to trust online newspapers to get health information. Blacks also were more likely than whites to use the Internet to access health information when in the midst of a strong need event. However, minorities who are privately insured were more likely than their uninsured counterparts to rely on the Internet. These findings are important considering that federally insured persons who are connected to mobile devices had the highest probability of

					reliance on the Internet as a go-to source of health information. In sum, these findings lend credence that mobile technologies are important for achieving greater racial equity in health behavior and health outcomes."
Smith et al. (2016)[26]	N/A	Clinical	Southern US	Policies related to developing data sharing to support healthcare centers serving vulnerable populations	"The innovative data warehouse project in Maryland can inform and transform the quality of health care delivered to the state's most vulnerable populations. However, the project is still in its early stages and has yet to translate this tremendous potential into real-world improvements. Future research should revisit the data warehouse efforts to further evaluate its reach and impact. Future research should focus not only on progress on clinical outcomes and the delivery of preventive services but also on changes in clinical practice resulting from data sharing, benchmarking, and collaboration around quality improvement."
Super Church (2015)[18]	N/A	N/A	Northeastern US	Programs related to leveraging data for	Collaborative partnerships are fundamental to the success of Healthy Neighborhood Equity

				neighborhood improvement	Funds. Without these collaborations, access to the data sources needed to effectively target investments and measure the impact of our investments over time would be impossible
Terry (2016)[10]	N/A	N/A	US	Areas of the Internet of (Health) Things that should be regulated	"While the resultant IoHT has great promise (some dystopian predictions aside), policymakers and regulators have failed to articulate strong and consistent regulation regarding data protection, efficacy, or safety. Currently, apps, wearables, and IoT hardware and software are only lightly regulated. Regarding data protection, the explanation is as simple as it is unfortunate. Outside of the HIPAA "zone," the protection of healthcare information is negligible. The quality and safety situation is more nuanced. The FDA has the power to regulate this area yet has taken something of a hands-off approach, although it seems increasingly concerned about the security of medical devices. Filling in the gaps on, hopefully, only a temporary basis, the FTC is increasingly intervening with regard to apps and devices that are ineffective or threaten

					privacy. If these technologies are to transform, or even disrupt, our existing healthcare systems, they deserve to be overseen by a consistent and well-thought-out regulation."

Abbreviations: CAP, Center for American Progress; EHR, electronic health record; HITECH, Health Information Technology for Economic and Clinical Health; HL, health literacy; IoHT, internet of health things; LGBT, lesbian, gay, bisexual, and transgender; N/A, not applicable; ONC, Office of the National Coordinator for Health Information Technology; SDoH, social determinants of health; SO/GI, sexual orientation and gender identity; US, United States.

Supplemental Table 12. Digital health technology

Study	Design	Setting	Geographic Location	Type of Technology	Sample Size of Data Set	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
Ahmed et al. (2016)[44]	Prospective	Clinical	Midwestern US	Software	N/R	Sex, Age, Ethnicity, Race	Hospital readmission; Emergency department visits	"This collaborative effort between clinical and public health entities aligns with present day health reform efforts that call for projects aimed at closing the gap between clinical care for the patient and the health of the population. Emerging capabilities in health information technology may serve as fertile ground for future collaborative efforts

								between clinical medicine and public health, provide for sustainable and scalable infrastructures to support ongoing integration, and support collaborative efforts to improve individual and population health, reduce costs, and improve the care experience.”
Angier et al. (2014)[45]	Retrospective	Clinical	US	Geocoding	228,224	Geography, Neighborhood environment (zip)	Healthcare utilization	“EHR data can be imported into a web-based GIS mapping tool to visualize patient information. Using EHR data, we were able to

								observe smaller areas than could be seen using only publicly available data.”
Aoyagi (2015)[46]	Retrospective	Community	Western US	Geocoding	203	Race, Ethnicity, Poverty, Owner occupied	Toxic emissions	“Our results support the notions that (i) environmental exposures such as TRI emissions are clustered throughout Los Angeles, and possibly, similar urban areas; and (ii) improvements in environmental quality due to TRI emissions and toxicity decreases tend to benefit populations that are more economically

								and socially empowered.”
Basu et al. (2017)[47]	Retrospective	Clinical	US	Machine learning	Look AHEAD, 4,760 DPPOS, 1,018 ACCORD, 9,635	Age, Gender/sex, Race, Ethnicity, Tobacco use, Drug use (non-illicit)	CVD	RECODe might improve estimation of risk of complications for patients with type 2 diabetes
Bauer et al. (2015)[48]	Retrospective	Community	Northeastern US	Geocoding	561,754	Race, Poverty, Unemployment rate, Neighborhood environment	Healthcare utilization	“Eliminating poverty is an important goal of society and increasing access to income-related social services is one strategy to reduce poverty. This cross-sectional analysis identifies block groups and neighborhoods in the Boston area with limited geographic

								access to agencies providing income-related social services despite a large population in need. It is important to note that although income-related social service agencies are unevenly distributed in Boston, the distribution does skew towards those areas with more concentrated poverty. City planning should take into consideration the geographic location of populations in need when deciding

								where to establish new social service agency locations.”
Bejan et al. (2018)[49]	Retrospective	Clinical	US	Machine learning	2,634,057	Homelessness, Mental health, Tobacco use	Mental health	"We provide an efficient solution for mining homelessness and ACE information from EHRs, which can facilitate large clinical and genetic studies of these social determinants of health."
Botticello et al. (2016)[50]	Retrospective	Clinical	US	Geocoding	8,351	Race, Ethnicity, Geography, Neighborhood environment, Access to (healthy) food, Urbanization, SES	Spinal cord injury (SCI)	“Neighborhood characteristics may be critical in understanding race disparities in community outcomes after SCI. It is important to identify

						Advantage Index		barriers to community reintegration after SCI that may result in inequalities in health, disability, and quality of life. This is especially important among historically disadvantaged and marginalized groups of people residing in areas with adverse conditions who have less personal and economic resources to overcome environmental barriers and are in greater need of interventions.

								Neighborhood differences are modifiable. Research focused on understanding the role of residential context in the experience of disability will allow us to address persistent inequalities in health by improving the environment with informed public policy.”
Dai et al. (2017)[51]	Retrospective	Community	US	Geocoding; Social media	72,758	Race, Ethnicity, Age, Poverty, Owner occupied, Household size	Tobacco use	“At the national level, there are inequalities of vape shop density by some socio-demographic characteristics and heterogeneity between urban

								and nonurban areas.”
Drewnowski et al. (2016)[52]	Prospective	Community	Northwestern US	Geocoding	2,001	Educational attainment, Income, Own/rent, Home (property) value	Diet	“Residential property values may capture socioeconomic disparities better than the conventional measures of education and income. Ability to geo-localize residential property values opens the door to valuable studies on the geography of diets and health.”

Gebreab et al. (2017)[53]	Prospective	Community	Southern US	Geocoding	Total, 5301 Cross-sectional analysis, 4,693 Longitudinal analysis, 3,670	Age, Diet/obesity, Physical activity, Social cohesion, Safety and violence (crime), Access to (healthy) food	Diabetes	In conclusion, our findings provide longitudinal evidence that neighborhoods with greater density of unfavorable food stores may increase the risk of developing T2DM among African Americans independent of individual-level risk factors and neighborhood social cohesion. Our findings also showed neighborhoods with better social cohesion may be protective of future development of T2DM

									independent of individual-level risk factors. In addition, we found strong association between neighborhood problems and prevalence of T2DM independent of individual-level risk factors and GIS-based measures. Additional research is needed to corroborate our findings using rigorous longitudinal studies or natural experiments or randomized trials. If corroborated by future studies, these

								findings suggest that modification of neighborhood environments might be an important strategy to consider for the prevention of T2DM in African Americans.
Henly et al. (2016)[54]	Retrospective	Internet	US	Machine learning; Social media	N/R	Income, Educational attainment, Grocery or retail food store	Foodborne illness	"These results suggest that well-known health disparities might also be reflected in the online environment."
Hosgood et al. (2013)[55]	Retrospective	Community	Northeastern US	Geocoding	9,670	Geography, Gender/sex	Lung cancer	"Our exploratory findings generated hypotheses that environmental exposures and socioeconomic

								factors may con-tribute to lung cancer rates, specifically large cell carcinoma in Maine."
Insaf et al. (2015)[56]	Retrospective	Community	Northeastern US	Geocoding	562,586	Race, Ethnicity, Tobacco use, Geography	Low birth weight (LBW)	"Neighborhood racial composition contributes to disparities in LBW prevalence beyond differences in behavioral and socioeconomic factors. Small-area analyses of LBW can identify areas for targeted interventions and display unique local patterns that should be accounted for in prevention strategies."

Jamei et al. (2017)[36]	Retrospective	Clinical	Western US	Geocoding; Machine learning	335,815	Tobacco use, Alcohol use, Drug use, Neighborhood environment	Hospital readmission; Cost	"In this study, we successfully trained and tested a neural network model to predict the risk of patients' rehospitalization within 30 days of their discharge. This model has several advantages over LACE, the current industry standard, and other proposed models in the literature including (1) significantly better performance in predicting the readmission risk, (2) being based on real-time data from

								EHR, and thus applicable at the time discharge from hospital, and (3) being compact and immune to model drift. Furthermore, to determine the classifier's labeling threshold, we suggested a simple cost-saving optimization analysis."
Kramer et al. (2014)[57]	Retrospective	Community	Southern US	Geocoding	1,815,944	Health status, Race, Neighborhood environment	Preterm birth	"The creation of trajectories from geocoded maternal longitudinally-linked vital records is one method to carry out life course maternal and child health research."

Leach et al. (2016)[58]	Prospective	N/A	US	Geocoding	30	Age, Physical activity, Neighborhood environment	CVD	"This study suggests that for African American women, being younger than 55 years old and having access to many high-quality neighborhood PARs is associated with having multiple CVD risk factors at ideal levels. Implications of these findings include taking into account built environment factors when discussing or addressing lifestyle modification for CVD risk in African American

								women. These findings may be important for reducing racial disparities in CVD risk via increased PA, particularly for residents of low socioeconomic status or minority neighborhoods with few low-quality PARs. These residents could be referred to further CVD risk screening or to PA programs that can address and overcome these negative neighborhood characteristics. ”

Lee et al. (2015)[59]	Retrospective	Community	Southern US	Geocoding; Internet	500	Walkability, Income, Access to (healthy) food, Owner occupied	CVD	"The RTRN DCC web-based GIS application might be useful in CVD-related research in which short-term enrollment or retrospective geocoding is planned, and data capture is the main purpose of using the tool. This tool successfully captured geospatial data for a multi-site hypertension case-control study. This customized tool cut costs for GIS software and personnel, reduced the time needed

								for training and allowed standardization of procedures across sites and on-site geocoding for sites reluctant to release patient data. The RTRN DCC GIS application could also be applied to other fields of epidemiology studies to investigate the association of the community environment with diseases and to the phase IV clinical trials to determine the modifying effect of geospatial factors on trial efficacy.”

Marino et al. (2016)[31]	Retrospective	Clinical	Northwestern US	Software	11,041	Diet/obesity, Insurance status, Tobacco use	Healthcare utilization	Utilizing the Oregon Experiment, a randomized natural experiment, this study demonstrates a causal relationship between Medicaid coverage and receipt of several preventive services in CHC patients, including receipt of breast and cervical cancer screenings as well as screenings for BMI, blood pressure, and smoking, during a 3-year follow-up.

Masho et al. (2017)[28]	Retrospective	Community	Southern US	Geocoding	27,519	Safety and violence (crime)	Preterm birth; Pregnancy complications	This study found a statistically significant association between violence and very preterm births (<32 weeks) after adjusting for individual factors. Individual factors accounted for a large portion of the covariance between youth violence and preterm birth.
Navathe et al. (2018)[38]	Retrospective	Clinical	Northeastern US	Natural language processing/text mining	49,319	Mental health, Drug use, Social support, Housing stability	Hospital readmission	"The seven social risk factors studied are substantially more prevalent than represented in administrative data. Automated

								methods for analyzing physician."
Nguyen et al. (2017)[34]	Retrospective	Internet	US	Geocoding; Machine learning; Social media	Twitter, 79,848,992 Yelp, 505,554	Food security	Mortality; Self-rated health; Diabetes; Chronic conditions	“In this study, we demonstrate that social media can be utilized to create indicators of the food environment that are associated with state-level mortality, health behaviors, and chronic conditions. Social media represents an untapped resource for public health research and intervention.”

Nguyen et al. (2017)[60]	Prospective	Internet	US	Geocoding; Machine learning; Social media	603,363	Diet/obesity, Physical activity, Drug use, Alcohol use	Alcohol-related mortality	"Social media represents a new type of real-time data that may enable public health officials to examine movement of norms, sentiment, and behaviors that may portend emerging issues or outbreaks—thus providing a way to intervene to prevent adverse health events and measure the impact of health interventions."
Noyes (2014)[61]	Prospective	Community	Northeastern US	Video	324	Physical activity	Underweight	"Bicycle lanes were used by local residents of a low-income urban neighborhood. Compared

								with neighborhood residents overall, cyclists reported better health and health behaviors."
Oreskovic et al. (2017)[35]	Retrospective	Clinical	Northeastern US	Natural language processing/text mining	120	Insurance status, Age, Race	Mental health	"This study provides an important step forward for population health management by outlining a new method for identifying the important role that social determinants and mental health play in health outcomes, and offers a promising new approach to stratifying this risk burden on

								a population level."
Pan et al. (2017)[62]	Retrospective	Clinical	Midwestern US	Machine learning	6,457	Health status, Tobacco use, Age, Homelessness, Drug use, Mental health	Pregnancy complications	"Our analysis exhibits the potential for machine learning to move government agencies toward a more data-informed approach to evaluating risk and providing social services. Overall, such efforts will improve the efficiency of allocating resource-intensive interventions."

Piccolo et al. (2015)[63]	Retrospective	Community	Northeastern US	Geocoding	5,502	Income, Poverty, Race, Ethnicity, Neighborhood environment, Grocery or retail food store, Physical activity, Age, Gender/sex, Diet/obesity	Diabetes	"In conclusion, using data from the BACH Survey, we have identified large, significant, neighborhood variability in the prevalence of T2DM. However, the many neighborhood factors we were able to examine did not explain this neighborhood variability, nor did they appear to play a role in the amplification or creation of racial/ethnic disparities in T2DM. While the findings of this study

								overall suggest that neighborhood factors are not a major contributor to racial/ethnic disparities in T2DM, there is a need for further research including data from other geographic locations, capturing both urban and rural areas and locations with both high and low residential segregation.”
Prussing et al. (2013)[64]	Retrospective	Community	Southern US	Geocoding	1,384	Geography, Poverty, Race, Ethnicity, Nativity, Health status, Drug use, Alcohol use	Tuberculosis (TB)	“In Maryland from 2004 to 2010, two distinct geospatial clusters of TB cases were identified, one in Baltimore City and the

								other in Montgomery and Prince George's counties. The TB cases and census tracts that make up these geospatial clusters had distinct demographic, socioeconomic, and risk-factor characteristics that differed from characteristics of the state at large. These TB clusters show a clear distribution of social health inequality."

Ray et al. (2017)[11]	Retrospective	Community	US	Mobile	3,165	Race, Ethnicity	Healthcare utilization; Health literacy	"Blacks and Latinos, compared to whites, were more likely to trust online newspapers to get health information. Blacks also were more likely than whites to use the Internet to access health information when in the midst of a strong need event. However, minorities who are privately insured were more likely than their uninsured counterparts to rely on the Internet. These findings are important considering

								that federally insured persons who are connected to mobile devices had the highest probability of reliance on the Internet as a go-to source of health information. In sum, these findings lend credence that mobile technologies are important for achieving greater racial equity in health behavior and health outcomes.”
Salow et al. (2018)[65]	Retrospective	Clinical	Midwestern US	Geocoding	5,174	Degree of segregation, Poverty, Insurance status, Health status	Preterm birth	“Among non-Hispanic Black women in an urban area, high levels of segregation

								were independently associated with the higher odds of spontaneous preterm birth."
Sharifi et al. (2016)[66]	Retrospective	Clinical	Northeastern US	Geocoding	44,810	Race, Ethnicity, Income, Educational attainment, Food security, Physical activity	Obesity	"In conclusion, this study contributes to the understanding of potential socio-contextual pathways that may underlie alarming disparities in childhood obesity. The results suggest that neighborhood SES is an important driver of disparities in child and adolescent BMI and that built

								environment characteristics also help explain obesity disparities. These results highlight the imperative need to address contextual factors that contribute to disparities in childhood overweight and obesity such as the neighborhoods and built environments in which children live.”
Shim et al. (2013)[67]	Retrospective	Community	US	Internet	39,149	Health status, Technology access, Income, Educational attainment, Poverty	Healthcare utilization	"The alignment between survey mode selection, internet access, and health disparities, as

								well as genuine survey mode characteristics, leads to web–mail differences in SRH. Unless the digital divide and its influences on survey mode selection are resolved and differential genuine mode effects are fully comprehended, we recommend that both modes be simultaneously used on a complementary basis."

Sills et al. (2017)[68]	Retrospective	Clinical	US	Geocoding	445,668	Single parents, Gender/sex, Insurance status, income, Health status	Hospital readmission	"In conclusion, the results of our analysis show that SDoH risk adjustment has substantial impact at the hospital level, where readmission penalties are calculated, despite only a small impact on readmission prediction model performance at the discharge level. The large proportion of hospitals that change rank decile with each SDoH adjustment model reinforces our previous

								finding that SDoH risk adjustment can impact penalties levied for readmissions. For pay for performance measures calculated at the hospital level, and for research on hospital-level performance, our findings support the inclusion of SDoH variables in risk adjustment."
Silverman et al. (2015)[69]	Retrospective	Community	Northeastern US	Software	527,056	Income, Social support	Hospital readmission	"The 3 level [(individuals, organizations, and society] approach appears to be useful to help health administrators sort through

								system complexities to find effective interventions at lower costs."
Threatt et al. (2017)[70]	Prospective	Clinical	Southern US	Telehealth	Total, 33 Free telehealth, 12 Free traditional clinic, 21	Educational attainment, Income, Poverty	Obesity; Blood pressure	"Expanding access to care in populations faced with challenges of socioeconomics, limited education, and lower health literacy is a step toward reducing health disparities and positively affecting care. Mean [hemoglobin] A1C can be improved with telehealth DSME/S services in an underserved,

								free clinic population."
Tomayko et al. (2015)[71]	Retrospective	Clinical	Midwestern US	Geocoding	102,231	Race, Ethnicity, Economic hardship index	Obesity	"The factors contributing to obesity prevalence are extremely complex, and EHI represents only a component of obesity risk. However, our study suggests that the relationship between EHI and its interaction with race/ethnicity that was uncovered in LA in regards to childhood obesity is also evident in Wisconsin, suggesting the

								utility of this composite index score in measures of health. Despite the limitations mentioned, we demonstrated the feasibility of using EHR-based methods, which represent a substantial savings of both time and financial resources compared to traditional data collection methods. In summary, the PHINEX dataset enabled an examination of patient-level demographic information aggregated within defined

								geographic boundaries and for assessment of factors that may contribute to childhood obesity. Understanding how these factors act individually and in combination will allow researchers, practitioners, and public health professionals to tailor intervention programs to local communities and at-risk populations."
Zenk et al. (2013)[72]	Retrospective	Community	Midwestern US	Geocoding	919	Age, Race, Ethnicity, Gender/sex, Grocery or retail food	Diet	"The study suggests that unfair treatment in retail interactions

						store, Food security		warrants investigation as a pathway by which restricted neighborhood food environments and food shopping behaviors may adversely affect health and contribute to health disparities."

Abbreviations: ACCORD, action to control cardiovascular risk in diabetes; ACE, adverse childhood experience; AHEAD, action for health in diabetes; BACH, Boston area community health; BMI, body mass index; CHC, community health center; CVD, cardiovascular disease; DCC, data coordinating center; DPPOS, diabetes prevention program outcome study; DSME/S, diabetes self-management education and support; EHI, electronic health information; EHR, electronic health record; GIS, geographic information system; IV, four; LA, Los Angeles; LACE, length of stay, acuity, comorbidities, ER (emergency room) visits (hospital index); LBW, low birth weight; NR, not reported; PA, physical activity; PAR, physical activity resources; PHINEX, public health information exchange; RECODe, risk equations for complications of type 2 diabetes; RTRN, research centers in minority institutions (RCMI) translational research network; SES, socioeconomic status; SCI, spinal cord injury; SDoH, social determinants of health; SRH, self-rated health; T2DM, type 2 diabetes mellitus; TB, tuberculosis; TRI, toxic release inventory; US, United States.

Supplemental Table 13. Data

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
Angier et al. (2014)[45]	Retrospective	Clinical	US	EHRs	228,224	Geography, Neighborhood environment (zip)	Healthcare utilization	"EHR data can be imported into a web-based GIS mapping tool to visualize patient information. Using EHR data, we were able to observe smaller areas than could be seen using only publicly available data."

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
Aoyagi (2015)[46]	Retrospective	Community	Western US	Study survey; Private data	203	Race, Ethnicity, Poverty, Owner occupied	Toxic emissions	"Our results support the notions that (i) environmental exposures such as TRI emissions are clustered throughout Los Angeles, and possibly, similar urban areas; and (ii) improvements in environmental quality due to TRI emissions and toxicity decreases tend to benefit populations that are more economically and socially empowered."

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
Asada (2013)[73]	Retrospective	Community	US	Census	2,080,894	Race, Ethnicity, Income	Functional limitation	"Our proposed approach offers policy-relevant health disparity information in a comparable and interpretable manner, and currently publicly available data support its application."
Basu et al. (2017)[47]	Retrospective	Clinical	US	Previous study	Look AHEAD, 4760 DPPOS, 1018 ACCORD, 9635	Age, Gender/sex, Race, Ethnicity, Tobacco use, Drug use (non-illicit)	CVD	"RECODe might improve estimation of risk of complications for patients with type 2 diabetes."

Bauer et al. (2015)[48]	Retrospective	Community	Northeastern US	Internet; Private database; Census	561,754	Race, Poverty, Unemployment rate, Neighborhood environment	Healthcare utilization	"Eliminating poverty is an important goal of society and increasing access to income-related social services is one strategy to reduce poverty. This cross-sectional analysis identifies block groups and neighborhoods in the Boston area with limited geographic access to agencies providing income-related social services despite a large population in need. It is important to note that although income-related social service

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
								agencies are unevenly distributed in Boston, the distribution does skew towards those areas with more concentrated poverty. City planning should take into consideration the geographic location of populations in need when deciding where to establish new social service agency locations."

Beyer et al. (2016)[29]	Retrospective	Community	Midwestern US	EHRs; State data; Federal data	1,010	Discrimination and bias, Neighborhood environment	Breast cancer; Mortality	"This work introduces two new environmental measures, drawing from a housing-focused database, that enable the consideration of racial bias in mortgage lending and residential redlining as predictors in health disparities research. This preliminary work indicates that these qualities of neighborhoods may have public health implications, and indicates that more work is needed in this area."

Botticello et al. (2016)[50]	Retrospective	Clinical	US	Federal data; Census	8,351	Race, Ethnicity, Geography, Neighborhood environment, Access to (healthy) food, Urbanization, SES Advantage Index	Spinal cord injury (SCI)	"Neighborhood characteristics may be critical in understanding race disparities in community outcomes after SCI. It is important to identify barriers to community reintegration after SCI that may result in inequalities in health, disability, and quality of life. This is especially important among historically disadvantaged and marginalized groups of people residing in areas with adverse conditions who have less

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
								personal and economic resources to overcome environmental barriers and are in greater need of interventions. Neighborhood differences are modifiable. Research focused on understanding the role of residential context in the experience of disability will allow us to address persistent inequalities in health by improving the environment with informed public policy.”

Castrucci et al. (2015)[74]	Prospective	Community	US	EHRs	45	Geography	Chronic disease	"Monitoring the status of community health is a core function of all public health departments. Public health professionals must have access to current local data on both risk factors and health status to effectively target interventions, wisely allocate resources, and assess the effects of interventions. While the innovations in our largest cities have allowed access to some local data related to chronic disease, LHDs still urgently need

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
								timelier and geographically specific data to efficiently and effectively address the most pressing problems in public health"

Cheng et al. (2016)[75]	Retrospective	Clinical	Northeastern US	EHRs	200,343	Single parents	Tobacco use during pregnancy; Infant birthweight	"In this large cohort study, the availability of paternal data on birth certificates in Massachusetts was independently related to perinatal risk factors for childhood obesity. In adjusted analyses, we observed higher rates of smoking during pregnancy, lower rates of breastfeeding initiation, and reduced infant birthweight among pregnancies without paternal data on the birth certificate as compared to pregnancies involving

								married mothers where paternal data were available. We also observed higher odds of children ever having a WFL ≥ 95th percentile and of crossing ≥ 2 major WFL percentiles in the first 2 years of life, although this association was modestly attenuated after adjusting for maternal characteristics. Our results raise the possibility of using missing paternal data on the infant birth certificate as a practical tool to identify children who may be at greater risk for

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
								certain perinatal precursors of childhood obesity and suggest that efforts to understand and reduce childhood obesity risk factors in early life may need to consider paternal factors.”
Dai et al. (2017)[51]	Retrospective	Community	US	Social media; Internet; Census; Federal data	72,758	Race, Ethnicity, Age, Poverty, Owner occupied, Household size	Tobacco use	“At the national level, there are inequalities of vape shop density by some socio-demographic characteristics and heterogeneity between urban and nonurban areas.”

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
Drewnowski et al. (2016)[52]	Prospective	Community	Northwestern US	Study survey; Local data	2,001	Educational attainment, Income, Own/rent, Home (property) value	Diet	"Residential property values may capture socioeconomic disparities better than the conventional measures of education and income. Ability to geo-localize residential property values opens the door to valuable studies on the geography of diets and health."

Flood et al. (2015)[76]	Retrospective	Clinical	Midwestern US	EHRs; Federal survey; Census	93,130	Age, Gender/Sex, Race, Ethnicity	Obesity	"Future directions include using the PHINEX data set to better understand how racial/ethnic factors interact with community-level covariates. The PHINEX data set is also capable of spatial and longitudinal analysis. Next steps include identifying the communities where childhood weight gain or loss occurs after controlling for other variables. This longitudinal and spatial approach could have implications for urban planning

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
								and community health needs assessments within Wisconsin. In sum, using statistically weighted and adjusted EHR data may provide a cost-effective solution for precise, local data that are actionable at the community level and comparable at a national scale.”

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
French et al. (2016)[77]	Retrospective	Clinical	Midwestern US	EHRs	150,661	Age, Race	Diabetes	"Over 90% of the identified diabetics were African American, 4.5% were white, 2.3% were Hispanic, and 3.2% other race/ethnicity and 37% were uninsured (range: 20% to 68.6%). Among those insured, 32.8% were covered by Medicare and only 10% by Medicaid."

Gebreab et al. (2017)[53]	Prospective	Community	Southern US	Federal data; Previous study	Total, 5301 Cross-sectional analysis, 4,693 Longitudinal analysis, 3,670	Age, Diet/obesity, Physical activity, Social cohesion, Safety and violence (crime), Access to (healthy) food	Diabetes	"In conclusion, our findings provide longitudinal evidence that neighborhoods with greater density of unfavorable food stores may increase the risk of developing T2DM among African Americans independent of individual-level risk factors and neighborhood social cohesion. Our findings also showed neighborhoods with better social cohesion may be protective of future development of T2DM independent of individual-level risk factors. In

								addition, we found strong association between neighborhood problems and prevalence of T2DM independent of individual-level risk factors and GIS-based measures. Additional research is needed to corroborate our findings using rigorous longitudinal studies or natural experiments or randomized trials. If corroborated by future studies, these findings suggest that modification of neighborhood environments might be an

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
								important strategy to consider for the prevention of T2DM in African Americans.”
Goldner et al. (2013)[78]	Retrospective	Community	US	Federal survey; Federal data	7,674	Urbanization, Age, Gender/sex	Healthcare utilization	“Internet use alone is not sufficient for eliminating health disparities among those in rural areas or for women.”

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
Goyal et al. (2015)[79]	Retrospective	Clinical	Midwestern US	EHRs	263	Health status, Tobacco use, Diet/obesity, Alcohol use, Drug use, Mental health	Preterm birth; Pregnancy complications	"Risk scoring based on social determinants can discriminate pregnancy risk within a Medicaid population; however, performance is modest and consistent with prior prediction models. Future research is needed to evaluate whether implementation of risk scoring in Medicaid prenatal care programs improves clinical outcomes"

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
Henly et al. (2016)[54]	Retrospective	Internet	US	Internet; Social Media	N/R	Income, Educational attainment, Grocery or retail food store	Foodborne illness	"These results suggest that well-known health disparities might also be reflected in the online environment."
Hosgood et al. (2013)[55]	Retrospective	Community	Northeastern US	State data; Census	9,670	Geography, Gender/sex	Lung cancer	"Our exploratory findings generated hypotheses that environmental exposures and socioeconomic factors may contribute to lung cancer rates, specifically large cell carcinoma in Maine."

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
Houle (2014)[80]	Retrospective	Community	US	Federal survey; Federal data; Local data; Census	N/R	Income, Poverty, Educational attainment, Race, Ethnicity, Financial issues (foreclosure rate)	Mental health	"The outcomes from this study support the perspective that the foreclosure crisis has the potential to exacerbate existing social disparities in mental health."

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
Insaf et al. (2015)[56]	Retrospective	Community	Northeastern US	Census	562,586	Race, Ethnicity, Tobacco use, Geography	Low birth weight	"Neighborhood racial composition contributes to disparities in LBW prevalence beyond differences in behavioral and socioeconomic factors. Small-area analyses of LBW can identify areas for targeted interventions and display unique local patterns that should be accounted for in prevention strategies."

Jamei et al. (2017)[36]	Retrospective	Clinical	Western US	EHRs; Census	335,815	Tobacco use, Alcohol use, Drug use, Neighborhood environment	Hospital readmission; Cost	"In this study, we successfully trained and tested a neural network model to predict the risk of patients' rehospitalization within 30 days of their discharge. This model has several advantages over LACE, the current industry standard, and other proposed models in the literature including (1) significantly better performance in predicting the readmission risk, (2) being based on real-time data from EHR, and thus applicable at the time discharge from

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
								hospital, and (3) being compact and immune to model drift. Furthermore, to determine the classifier's labeling threshold, we suggested a simple cost-saving optimization analysis.”

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
Jones (2013)[81]	Retrospective	Community	US	Federal survey; Census	200,102	Degree of segregation, Income, Educational attainment	Hypertension	"These findings reveal that SES has differential effects across segregation types and that hypertension in disadvantaged (extremely hyper segregated) areas maybe a function of structural constraints rather than socioeconomic position."

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
Kasthurirathne et al. (2018)[82]	Retrospective	Clinical	Midwestern US	EHRs; Local data; Census	84,317	Income	Mental health; Obesity; Healthcare utilization	"Our results indicate the potential to predict the need for various social services with considerable accuracy and represent a model for reimplementation across other datasets, health outcomes, and patient populations."
Kramer et al. (2014)[57]	Retrospective	Community	Southern US	State data; Census	1,815,944	Health status, Race, Neighborhood environment	Preterm birth	"The creation of trajectories from geocoded maternal longitudinally-linked vital records is one method to carry out life course maternal and child health research."

Leach et al. (2016)[58]	Prospective	N/A	US	Study survey; Crime data	30	Age, Physical activity, Neighborhood environment	CVD	"This study suggests that for African American women, being younger than 55 years old and having access to many high-quality neighborhood PARs is associated with having multiple CVD risk factors at ideal levels. Implications of these findings include taking into account built environment factors when discussing or addressing lifestyle modification for CVD risk in African American women. These findings may be

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
								important for reducing racial disparities in CVD risk via increased PA, particularly for residents of low socioeconomic status or minority neighborhoods with few low-quality PARs. These residents could be referred to further CVD risk screening or to PA programs that can address and overcome these negative neighborhood characteristics."

Lee et al. (2015)[59]	Retrospective	Community	Southern US	Census	500	Walkability, Income, Access to (healthy) food, Owner occupied	CVD	"The RTRN DCC web-based GIS application might be useful in CVD-related research in which short-term enrollment or retrospective geocoding is planned, and data capture is the main purpose of using the tool. This tool successfully captured geospatial data for a multi-site hypertension case-control study. This customized tool cut costs for GIS software and personnel, reduced the time needed for training and allowed standardization

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
								of procedures across sites and on-site geocoding for sites reluctant to release patient data. The RTRN DCC GIS application could also be applied to other fields of epidemiology studies to investigate the association of the community environment with diseases and to the phase IV clinical trials to determine the modifying effect of geospatial factors on trial efficacy.”

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
Marino et al. (2016)[31]	Retrospective	Clinical	Northwestern US	EHRs; State data	11,041	Diet/obesity, Insurance status, Tobacco use	Healthcare utilization	"Utilizing the Oregon Experiment, a randomized natural experiment, this study demonstrates a causal relationship between Medicaid coverage and receipt of several preventive services in CHC patients, including receipt of breast and cervical cancer screenings as well as screenings for BMI, blood pressure, and smoking, during a 3-year follow-up."

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
Masho et al. (2017)[28]	Retrospective	Community	Southern US	State data; Census; Crime data	27,519	Safety and violence (crime)	Preterm birth; Pregnancy complications	"This study found a statistically significant association between violence and very preterm births (<32 weeks) after adjusting for individual factors. Individual factors accounted for a large portion of the covariance between youth violence and preterm birth."

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
Navathe et al. (2018)[38]	Retrospective	Clinical	Northeastern US	EHRs	49,319	Mental health, Drug use, Social support, Housing stability	Hospital readmission	"The seven social risk factors studied are substantially more prevalent than represented in administrative data. Automated methods for analyzing physician."

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
Nguyen et al. (2017)[34]	Retrospective	Internet	US	Social media; Internet; Federal data; Census; State data; Federal survey	Twitter, 79,848,992 Yelp, 505,554	Food security	Mortality; Self-rated health; Diabetes; Chronic conditions	"In this study, we demonstrate that social media can be utilized to create indicators of the food environment that are associated with state-level mortality, health behaviors, and chronic conditions. Social media represents an untapped resource for public health research and intervention."

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
Noyes (2014)[61]	Prospective	Community	Northeastern US	Study survey	324	Physical activity	Underweight	"Bicycle lanes were used by local residents of a low-income urban neighborhood. Compared with neighborhood residents overall, cyclists reported better health and health behaviors."

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
Oreskovic et al. (2017)[35]	Retrospective	Clinical	Northeastern US	EHRs	120	Insurance status, Age, Race	Mental health	"This study provides an important step forward for population health management by outlining a new method for identifying the important role that social determinants and mental health play in health outcomes, and offers a promising new approach to stratifying this risk burden on a population level."

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
Pan et al. (2017)[62]	Retrospective	Clinical	Midwestern US	State data	6,457	Health status, Tobacco use, Age, Homelessness, Drug use, Mental health	Pregnancy complications	"Our analysis exhibits the potential for machine learning to move government agencies toward a more data-informed approach to evaluating risk and providing social services. Overall, such efforts will improve the efficiency of allocating resource-intensive interventions."

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
Pearson-Stuttard et al. (2017)[83]	Retrospective	Community	US	Federal survey; Census	215,861,896	Food security	CVD	"Fiscal strategies targeting diet might substantially reduce CVD burdens. A national 10% F&V subsidy would save by far the most lives, while a 30% F&V subsidy targeting SNAP participants would most reduce socio-economic disparities. A combined policy would have the greatest overall impact on both mortality and socio-economic disparities."

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
Perzynski et al. (2017)[33]	Retrospective	Clinical	Midwestern US	EHRs; Census	243,249	Race, Insurance status, Technology access, Age	Healthcare utilization	"Our study found that for patients at one urban hospital system, differences in home broadband internet access demonstrate a clear negative association with patient portal initiation and use. Overall, patient portal initiation was modest, and activity was systematically lower for blacks, Hispanics, older adults, and persons of low socioeconomic status (Medicaid and uninsured)."

Piccolo et al. (2015)[63]	Retrospective	Community	Northeastern US	Study survey; Census; Crime data; Private data	5,502	Income, Poverty, Race, Ethnicity, Neighborhood environment, Grocery or retail food store, Physical activity, Age, Gender/sex, Diet/obesity	Diabetes	"In conclusion, using data from the BACH Survey, we have identified large, significant, neighborhood variability in the prevalence of T2DM. However, the many neighborhood factors we were able to examine did not explain this neighborhood variability, nor did they appear to play a role in the amplification or creation of racial/ethnic disparities in T2DM. While the findings of this study overall suggest that neighborhood

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
								factors are not a major contributor to racial/ethnic disparities in T2DM, there is a need for further research including data from other geographic locations, capturing both urban and rural areas and locations with both high and low residential segregation.”

Prussing et al. (2013)[64]	Retrospective	Community	Southern US	State data; Census	1,384	Geography, Poverty, Race, Ethnicity, Nativity, Health status, Drug use, Alcohol use	Tuberculosis (TB)	"In Maryland from 2004 to 2010, two distinct geospatial clusters of TB cases were identified, one in Baltimore City and the other in Montgomery and Prince George's counties. The TB cases and census tracts that make up these geospatial clusters had distinct demographic, socioeconomic, and risk-factor characteristics that differed from characteristics of the state at large. These TB clusters show a clear distribution of

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
								social health inequality.”
Rhoads et al. (2015)[32]	Retrospective	Community	Western US	State data	33,593	Race, Ethnicity, Access to healthcare	Cancer; Mortality	“IHS delivered higher rates of evidence based care; was associated with lower 5-year mortality. Racial/ethnic disparities in survival were absent in IHS.”

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
Roth et al. (2014)[84]	Retrospective	Clinical	Midwestern US	EHRs; Private data; Census; Federal data	62,701	Age, Race, Gender/sex, Grocery or retail food store, Educational attainment	Obesity	"Integrating community data into the EHR maximizes the potential of secondary use of EHR data to study and impact obesity prevention and other significant public health issues."
Salow et al. (2018)[65]	Retrospective	Clinical	Midwestern US	EHRs; Census	5,174	Degree of segregation, Poverty, Insurance status, Health status	Preterm birth	"Among Black women in an urban area, high levels of segregation were independently associated with the higher odds of spontaneous preterm birth."

Sharifi et al. (2016)[66]	Retrospective	Clinical	Northeastern US	EHRs; Census	44,810	Race, Ethnicity, Income, Educational attainment, Food security, Physical activity	Obesity	"In conclusion, this study contributes to the understanding of potential socio-contextual pathways that may underlie alarming disparities in childhood obesity. The results suggest that neighborhood SES is an important driver of disparities in child and adolescent BMI and that built environment characteristics also help explain obesity disparities. These results highlight the imperative need to address contextual

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
								factors that contribute to disparities in childhood overweight and obesity such as the neighborhoods and built environments in which children live.”

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
Shim et al. (2013)[67]	Retrospective	Community	US	Federal survey	39,149	Health status, Technology access, Income, Educational attainment, Poverty	Healthcare utilization	"The alignment between survey mode selection, internet access, and health disparities, as well as genuine survey mode characteristics, leads to web–mail differences in SRH. Unless the digital divide and its influences on survey mode selection are resolved and differential genuine mode effects are fully comprehended, we recommend that both modes be simultaneously used on a complementary basis."

Sills et al. (2016)[30]	Retrospective	Clinical	US	EHRs; Private data	179,400	Race, Ethnicity, Insurance status, Geography	Hospital readmission	"The results of our analysis show that adjustment for SDoH changes hospitals' penalty status on a readmissions-based P4P measure. Without adjusting P4P measures for SDoH, hospitals that care for more vulnerable patients may receive penalties in part related to patient factors beyond the control of the hospital and unrelated to the quality of hospital care. Further work to characterize the effects of SDoH on performance

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
								measures may assist efforts to improve care quality and deliver more equitable care.”

Sills et al. (2017)[68]	Retrospective	Clinical	US	EHRs	445,668	Single parents, Gender/sex, Insurance status, income, Health status	Hospital readmission	"In conclusion, the results of our analysis show that SDH risk adjustment has substantial impact at the hospital level, where readmission penalties are calculated, despite only a small impact on readmission prediction model performance at the dis- charge level. The large proportion of hospitals that change rank decile with each SDoH adjustment model reinforces our previous finding that SDoH risk adjustment can impact penalties

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
								levied for readmissions. For pay for performance measures calculated at the hospital level, and for research on hospital-level performance, our findings support the inclusion of SDoH variables in risk adjustment."

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
Steiner et al. (2018)[85]	Retrospective	Clinical	Western US	EHRs; Federal survey	130,208	Food security	Obesity; Quality of life	"Almost 6% of older members in a large, private-sector healthcare delivery system reported that they did not always have the money to buy the food they needed. Food insecurity was associated with minority race or ethnicity and Medicaid insurance coverage, as well as other social determinants of health such as low education and social isolation."

Tomayko et al. (2015)[71]	Retrospective	Clinical	Midwestern US	EHRs; Census	102231	Race, Ethnicity, Economic hardship index	Obesity	"The factors contributing to obesity prevalence are extremely complex, and EHI represents only a component of obesity risk. However, our study suggests that the relationship between EHI and its interaction with race/ethnicity that was uncovered in LA in regards to childhood obesity is also evident in Wisconsin, suggesting the utility of this composite index score in measures of health. Despite the limitations mentioned, we

								demonstrated the feasibility of using EHR-based methods, which represent a substantial savings of both time and financial resources compared to traditional data collection methods. In summary, the PHINEX dataset enabled an examination of patient-level demographic information aggregated within defined geographic boundaries and for assessment of factors that may contribute to childhood obesity. Understanding how these factors act

Study	Design	Setting	Geographic Location	Data Source	Sample Size	SDoH Factor(s)	Health Outcome(s)	Author Conclusions
								individually and in combination will allow researchers, practitioners, and public health professionals to tailor intervention programs to local communities and at-risk populations.”

Zenk et al. (2013)[72]	Retrospective	Community	Midwestern US	Study survey; Census	919	Age, Race, Ethnicity, Gender/sex, Grocery or retail food store, Food security	Diet	"The study suggests that unfair treatment in retail interactions warrants investigation as a pathway by which restricted neighborhood food environments and food shopping behaviors may adversely affect health and contribute to health disparities."