Journal of Clinical and Translational Science

www.cambridge.org/cts

Development of a social and environmental determinants of health informatics maturity model

Juan C. Espinoza[1,2], Shruti Sehgal[2], Jimmy Phuong[3,4], Neil Bahroos[5], Justin Starren[2], Adam Wilcox[6] and Daniella Meeker[7]

[1]Stanley Manne Children's Research Institute, Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, IL, USA; [2]Northwestern University Feinberg School of Medicine, Chicago, IL, USA; [3]Division of Biomedical and Health Informatics, University of Washington, Seattle, WA, USA; [4]Harborview Injury Prevention Research Center, University of Washington, Seattle, WA, USA; [5]University of Southern California Keck School of Medicine, Los Angeles, CA, USA; [6]Institute for Informatics, Department of Medicine, Washington University in St. Louis, St. Louis, MO, USA and [7]Department of Biomedical Informatics & Data Science, Yale University School of Medicine, New Haven, CT, USA

Implementation, Policy and Community Engagement Research Article

Cite this article: Espinoza JC, Sehgal S, Phuong J, Bahroos N, Starren J, Wilcox A, and Meeker D. Development of a social and environmental determinants of health informatics maturity model. *Journal of Clinical and Translational Science* **7**: e266, 1–10. doi: 10.1017/cts.2023.691

Received: 31 July 2023
Revised: 4 November 2023
Accepted: 29 November 2023

Keywords:
Social determinants of health; informatics; maturity models; health equity; clinical and translational research

Corresponding author:
J. C. Espinoza, MD;
Email: jespinozasalomon@luriechildrens.org

Abstract

Introduction: Integrating social and environmental determinants of health (SEDoH) into enterprise-wide clinical workflows and decision-making is one of the most important and challenging aspects of improving health equity. We engaged domain experts to develop a SEDoH informatics maturity model (SIMM) to help guide organizations to address technical, operational, and policy gaps. *Methods:* We established a core expert group consisting of developers, informaticists, and subject matter experts to identify different SIMM domains and define maturity levels. The candidate model (v0.9) was evaluated by 15 informaticists at a Center for Data to Health community meeting. After incorporating feedback, a second evaluation round for v1.0 collected feedback and self-assessments from 35 respondents from the National COVID Cohort Collaborative, the Center for Leading Innovation and Collaboration's Informatics Enterprise Committee, and a publicly available online self-assessment tool. *Results:* We developed a SIMM comprising seven maturity levels across five domains: data collection policies, data collection methods and technologies, technology platforms for analysis and visualization, analytics capacity, and operational and strategic impact. The evaluation demonstrated relatively high maturity in analytics and technological capacity, but more moderate maturity in operational and strategic impact among academic medical centers. Changes made to the tool in between rounds improved its ability to discriminate between intermediate maturity levels. *Conclusion:* The SIMM can help organizations identify current gaps and next steps in improving SEDoH informatics. Improving the collection and use of SEDoH data is one important component of addressing health inequities.

© The Author(s), 2023. Published by Cambridge University Press on behalf of The Association for Clinical and Translational Science. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial licence (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original article is properly cited. The written permission of Cambridge University Press must be obtained prior to any commercial use.

Introduction

Social determinants of health (SDoH) are defined by the World Health Organization as "the conditions in which people are born, grow, work, live, and age, and the wider set of forces and systems shaping the conditions of daily life" and include factors like socioeconomic status, education, neighborhood and physical environment, employment, social support networks, and access to health care [1]. Despite significant advances in high-quality healthcare access, health inequities in the United States persist by race, ethnicity, sexual orientation, gender identity, and disability, as well as by economic and community-level factors such as geographic location, poverty status, and employment. Socioeconomic status is perhaps the central concept that brings together the set of social determinants that shape health and plays a critical role in driving disparate health outcomes [2]. One study estimated that socioeconomic factors alone may account for 47% of health outcomes, while health behaviors, clinical care, and the physical environment account for 34, 16, and 3% of health outcomes, respectively [3].

In recent years, both public health and not-for-profit organizations have increasingly been focused on addressing SDoH and their impact on each individual's overall wellness and the ability of people to access healthcare services. For example, the Department of Health and Human Services (HHS) launched Healthy People 2030 [4], which has made a broader audience aware of SDoH issues. The COVID-19 pandemic has also shed light on the systemic drivers of health inequities in our society as well as innovative technology solutions to help mitigate the effects of the pandemics [5–9], specifically, the need for clinical data sharing and the collection and integration of high-quality SDoH data. The pandemic has also revealed the coupled role of the clinical health system and the non-clinical aspects that influence patient well-being and ability to maintain physical and mental wellness. The role of the social environment and

structural barriers inherent to access to care and supportive resources has led to a focus on community-level effects and information captured by different spatial data products, sometimes termed as social and environmental determinants of health (SEDoH) to emphasize the role of the environment [10].

Fundamentally, to address health disparities and improve patient health outcomes, health systems need the ability to gather, track, quantify, and use SEDoH data. Unfortunately, many of our health information systems and medical terminologies were not designed to capture this information [11]. This, combined with the fact that our understanding of SEDoH is still evolving, has resulted in many organizations making little or no progress in meaningfully gathering SEDoH data. One way to address the current state is through maturity models. Maturity models are self-evaluation tools that reflect best practices and can help organizations identify gaps, prioritize investments, and develop institutional roadmaps and strategies for growth. They are "based on the premise that people, organizations, functional areas, processes, etc., evolve through a process of development and growth towards a more advanced maturity accomplishing several stages [12]." Examples of clinical informatics maturity models include The HIMSS Electronic Medical Record Adoption Model [13], the Continuity of Care Maturity Model [14], and the Quintegra maturity model for electronic healthcare (eHMM) [15]. There has also been active development of research informatics maturity models, focusing on domains like research data warehouses and research data sharing [16–18]. Models exist covering a variety of domains, including the PACS (picture archiving and communication systems) [19] and Healthcare Analytics Adoption Model for data analysis [20], but no models focus on SEDoH information that will facilitate organizational advancements. The 2019 Parkland Center for Clinical Innovation SDOH Maturity Model [21] is intended to guide healthcare system executives when designing population health programs and developing patient-specific treatment plans, but it is focused on the behaviors and actions of leadership teams, rather than informatics.

Informed by our prior work focused on SEDoH [9,22–24] and supported by the Center for Data to Health (CD2H), we set out to develop and evaluate an SEDoH informatics maturity model. The goal of this project is to create a tool that can help organizations identify gaps and prioritize investments in their policies, processes, and technologies related to the capture and use of SEDoH data. Ultimately, better data and data-driven decision-making will help improve health outcomes for underserved and marginalized populations.

Materials and methods

Development phase

To develop the SEDoH informatics maturity model (SIMM), we leveraged the guidance provided by CD2H in developing self-assessment maturity tools, aligned with CD2H Informatics Maturity and Best Practices Core's mission of supporting development and dissemination of best practices in data use and informatics to the Clinical and Translational Science Award (CTSA) community [25].

Establishing the maturity model was a multi-step process (Fig. 1). First, we established a core expert group consisting of developers, informatics champions, and subject matter experts at the University of Southern California, Children's Hospital Los Angeles, and University of Washington. The next step involved extensive training on maturity model processes and procedures, including a review of over 30 existing informatics maturity models curated by author AW and colleagues. Conceptually, the model was informed by HHS' Healthy People 2030 and the socioecological model of health (Fig. 2). The core group identified the different dimensions/domain axes of the maturity model and iteratively reviewed them with external reviewers from CTSA hubs/CD2H leadership. This was followed by defining the maturity levels for the model. Seven levels were chosen instead of the typical five to provide additional granularity and specificity on the various intermediate steps toward SEDoH data maturity across the five domains. In all cases, we used a consensus approach in which all topics and levels were discussed until we arrived at a unanimous decision.

Alignment of Maturity Model with Principles of Implementation Science. Due to the complex nature of adopting SEDoH into healthcare delivery systems, we aimed to better align maturity modeling with principles of Implementation Science. For example, in addition to descriptions of technical assets and processes, we included domains for data collection policies and organizational and strategic maturity. These allow for the self-evaluation of organizational needs that have been highlighted in Implementation Science frameworks such as the Practical, Robust Implementation, and Sustainability Model and the updated Consolidated Framework for Implementation Research that explicitly address organizational and structural determinants of achieving maturity goals, such as policy, management, and resourcing postures [26,27].

Model description

We established a seven-level maturity model to map how SEDoH processes can reach maturity in each of five domains. This model may reflect the evolution, improvement, and transformation of an organization's capacity over time and captures its capabilities at each intermediate level.

The scope of every domain was iteratively defined during the development of the SIMM. During the development phase, the characteristics of maturity levels were established. A generic description of maturity levels outlined in Table 1 served as a foundation for developing maturity statements for every domain. To support the use of the SIMM tool, detailed narrative descriptions were developed for each of the seven maturity levels across all five domains (Table 2). The narratives for every level provide a detailed explanation of the stages from the simplest, ad hoc stage to the advanced, systematized, and optimized level. The goal of this maturity model was to offer an opportunity for institutions to have a structured way to identify their current level of capability or maturity with SEDoH information and find the gap between where they are and where they want to be to remain competitive and innovative. In this model, an institution is defined to be at a specific maturity level for a given domain when the process attributes for the lower maturity level(s) were fully achieved and the attributes for the specific maturity level were fully or largely achieved.

The model also includes a description and quantification of data sources that captures whether institutions are using patient-level SEDoH data (e.g., self-reported assessments) or contextual SEDoH data (e.g., geocoded variables extracted from the American Community Survey [28]). This section of the tool helps organizations catalog their SEDoH data sources and qualify patient-level data with a categorical hierarchy of ad hoc tools,

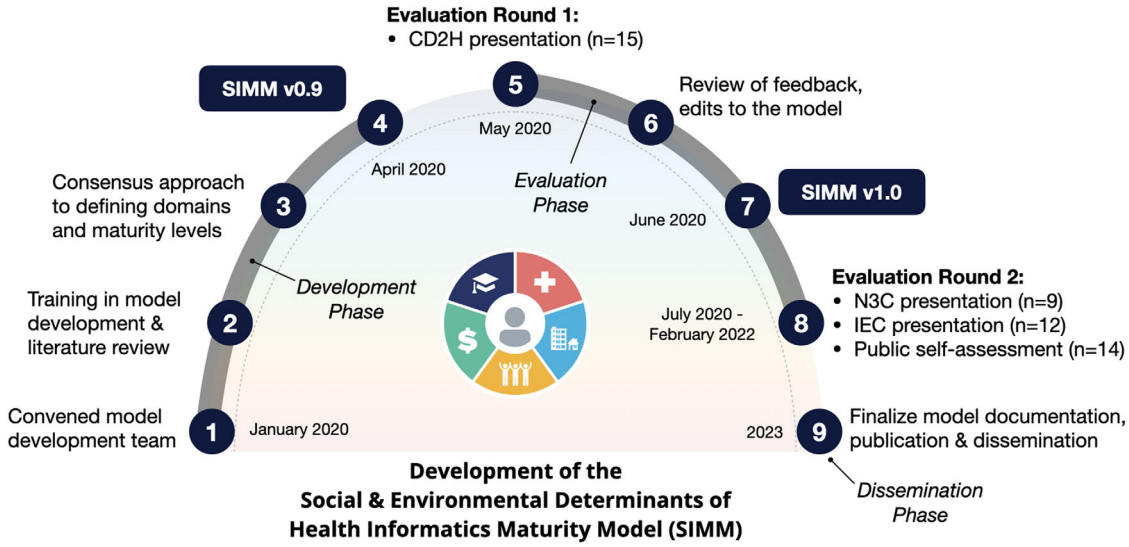

Figure 1. Process diagram illustrating the steps in the development of the social and environmental determinants of health informatics maturity model. SIMM = social & environmental determinants of health informatics maturity model, CD2H = Center for Data to Health, N3C = National COVID Cohort Collaborative, IEC = the Center for Leading Innovation and Collaboration's Informatics Enterprise Committee.

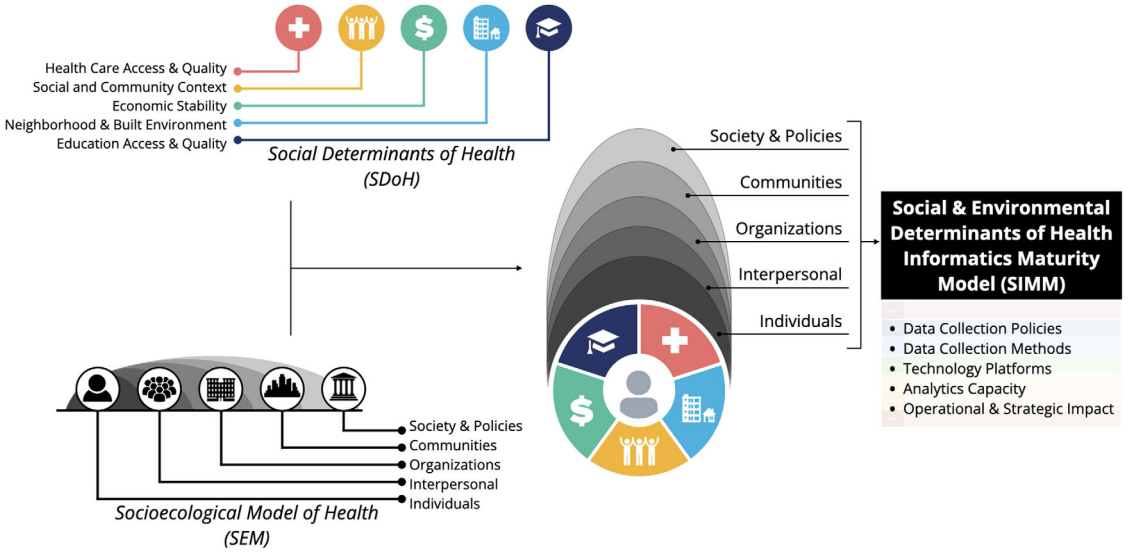

Figure 2. Conceptual model for social and environmental determinants of health informatics maturity model. Colors and icons adapted from Healthy People 2030, US Department of Health and Human Services, Office of Disease and Health Promotion. SDoH = social determinants of health, SEM = socioecological model of health.

validated and standardized tools, and validated tools augmented with items specific to the population of interest.

Evaluation phase

We performed two rounds of evaluation of the SIMM with respondents recruited from the CTSA network and CTSA-related activities, such as the Informatics Enterprise Committee (IEC), National COVID Cohort Collaborative (N3C), and CD2H.

First round

To evaluate the relevance of the identified domain axes and the established levels of maturity, the first version of the SIMM (v0.9) was reviewed at the CD2H Informatics Maturity and Best Practices community meeting in May 2020 by 15 informatics experts. Data were collected using Zoom's polling feature. The goal of the review meeting was to discuss and solicit comments on the candidate maturity model. To select a level for each domain axis, the attendees were asked to think about the highest level in the SIMM at which their organization consistently and comprehensively operates on a daily basis. The recommendations of the working group were compiled, and modifications were made to better describe the levels of maturity (SIMM v1.0) as well as to provide more context for improvement opportunities. In the case of conflicting recommendations, the same consensus approach described above was used.

Second round

We gathered data to evaluate SIMM v1.0 across three different opportunities. SIMM (v1.0) was presented at the N3C SDoH

Table 1. Generic description of each maturity level in the social and environmental determinants of health maturity model (SIMM).

Level	Description
Level 1: Absent	At this level, no data collection policies, methods, technology platforms, analytics capacity, and operational and strategic impact exist.
Level 2: Ad Hoc	Ad hoc, unstructured, chaotic, minimal, or no coordination.
Level 3: Emerging	It is characteristic at this level that some SEDoH processes are repeatable, possibly with consistent results. This level could be considered as an originating/developmental stage.
Level 4: Coordinated	This level is characterized by coordination and communication between the different groups. Leadership is typically supportive, but does not provide designated resources for support. Success is still mostly a function of a few individual groups' passion for making SEDoH happen.
Level 5: Supported	At this level, institutional support services related to SEDoH are established with support by leadership. There is operational and technical support as well as formal support for the piloting and launch of new services.
Level 6: Integrated	SEDoH data are readily available across the institution and information decision-making.
Level 7: Transformative	At the top maturity level, organizations start to learn from the collected SEDoH data. They are fully leveraging SEDoH, designing and developing their own SEDoH solutions to meet their innovative needs. The organization is continuously reviewing and improving its capabilities.

SEDoH = the social and environmental determinants of health.

workgroup meeting in July 2020 and the University of Rochester's Center for Leading Innovation and Collaboration IEC meeting in February 2021. During these virtual meetings, an overview of the need, rationale, and applications of the model was presented, and the attendees were asked for feedback. Then, they were sent a Google Form survey via e-mail that included the self-assessment and an opportunity to provide additional feedback on the tool itself. Comments were received from 21 respondents (12 IEC and nine N3C). The survey was anonymous and did not gather information about the individual or their home institution, but it was only distributed to individuals who are part of the CTSA Network.

The third evaluation opportunity in Round 2 was a publicly available online self-assessment. SIMM v1.0 was built using REDcap MariaDB SQL instance and added to the University of Washington's Maturity Model Self-Assessment Toolkit, along with other models such as the Research Informatics & Open Science Maturity model [25] (the toolkit is now being managed by Northwestern University). Users accessed the SIMM v1.0 assessment within the Maturity Model Resource Portal, an HTML front-end interface designed on top of a react.io framework with the REDCap server after institutional identity and access management steps. Users could login with their portal credentials, or complete any assessment anonymously as a guest. In total, 14 respondents representing 11 institutions (one from the South, seven from the Midwest, and three from the West Coast) completed the self-assessment between July 2020 and February 2022.

Ethical considerations

This project does not meet the definition of human subjects research and was deemed exempt by the Children's Hospital Los Angeles IRB, IRB# CHLA-20-00249.

Results

We developed a SIMM comprising seven advancing levels of maturity across five domain areas of data collection policies, data collection methods and technologies, technology platforms for analysis and visualization, analytics capacity, and operational and strategic impact (Fig. 3). The model also includes an assessment of data sources for both person-level and contextual SEDoH data.

SIMM Assessment Results – First Round (v0.9)

As seen in Figure 4, fifteen informaticists participated in the first round to evaluate SIMM v0.9. Over half of the respondents reported that their organizations had no standardized policies in place for collecting SEDoH data across their organization. Our data suggest that it was mostly ad hoc or some individuals followed consistent data collection practices. For the data collection methods and technologies domain, a majority of the respondents endorsed level 4 as the lowest maturity level, indicating that although SEDoH data was collected electronically using third-party platforms (e.g., Purple Binder), it was not integrated with the EHR. The maturity of technology platforms for analysis and visualization ranged from level 1 to level 6, with the majority centered around the middle levels. Similarly, with regard to SEDoH data, analytics capacity spanned from the lowest to the advanced maturity stages.

SIMM Assessment Results – Second Round (v1.0)

After revising the SIMM, 35 participants evaluated SIMM v1.0 in the second round. The maturity of the data collection policies and the data collection methods domain was centered around the middle levels. The majority of organizations indicated a higher maturity for the technology platforms for analysis and visualization domain. Analytics capacity dimension ranged from the least advanced to the most advanced maturity level and a similar trend was witnessed for the operational and strategic impact domain.

Discussion

SEDoH informatics maturity model

While the concept of maturity models is not new to the healthcare field, its application to the SEDoH domain has not been extensively researched. In the present study, we defined a seven-level maturity model with a three-fold objective: first, to assist organizations in self-assessing their current level of informatics maturity in relation to SEDoH capability, second, to help organizations with derivation

Table 2. Maturity level descriptions for the five domains of the social and environmental determinants of health informatics maturity model (SIMM)

Level	Description
Domain 1: Data collection policies	
Level 1: Absent	Institutions are not collecting SEDoH data (collecting demographics, such as race and ethnicity, does not qualify as SEDoH data collection). No well-defined, operational data collection policies in the organization.
Level 2: Ad Hoc	Ad hoc. No policies exist yet but various groups in the organization are collecting SEDoH data inconsistently.
Level 3: Emerging	Individuals with consistent SEDoH data collection practices.
Level 4: Coordinated	Certain groups, like clinics within a division (e.g., labor and delivery clinic), have standardized policy for SEDoH data collection.
Level 5: Supported	Standard policy for SEDoH data collection exists at the division level of an organization (e.g., Division of Obstetrics and Gynecology).
Level 6: Integrated	Few departments/ sections within a division of the organization (e.g., Department of Medicine) have a standard policy for SEDoH data collection.
Level 7: Transformative	Standardized data collection policy for SEDoH exists at an organizational or institutional level.
Domain 2: Data collection methods and technologies	
Level 1: Absent	Paper-based SEDoH data collection.
Level 2: Ad Hoc	Electronic data collection that is not part of the EHR or it is being done for research only.
Level 3: Emerging	SEDoH data are unstructured and free-texted within the EHR.
Level 4: Coordinated	Electronic, structured SEDoH data collection using a third-party platform such as Purple Binder or native solutions such as Qualtrics. Data collection is not integrated into clinical workflow.
Level 5: Supported	EHR-linked, structured data collection. For example, using the SDOH Wheel in Epic or Cerner Determinants of Health to intervene at the point of care using evidence-based screening tools and gain a better understanding of the social risk factors impacting your patient population.
Level 6: Integrated	Integrating historical data from external sources (e.g., USDA, EPA, ACS)**
Level 7: Transformative	Dynamic query of external sources** to bring in real-time or near-real-time actionable information.
Domain 3: Technology platforms for analysis and visualization	
Level 1: Absent	No specific technology platforms are used for SEDoH data analysis and visualization.
Level 2: Ad Hoc	Non-EHR, basic software is used for data visualization and performing analysis (e.g., Excel)
Level 3: Emerging	Non-EHR, interactive visual analytics platform in use(e.g., Power BI or Tableau)
Level 4: Coordinated	Analytics and visualizations are embedded in the EHR, that is, EHR-linked (HL7 or FHIR integration) analytics platform. This is non-actionable, read-only information, like a PDF where you are presenting the information to the user but there is no specific pathway that is triggered by that information.
Level 5: Supported	EHR-linked integrated and interactive analytics capability. This means that the information is actionable (e.g., it can trigger a CDS pathway or an order for a social work consult) or it creates a structured data field, such as a risk profile, that can be used to identify risk and group patients.
Level 6: Integrated	EHR-based integrated analytics and population health platform in use to support specific care management goals.
Level 7: Transformative	EHR-based integrated analytics and population health platform, enriched with external feed, such as ACS data or USDA data.
Domain 4: Analytics capacity	
Level 1: Absent	Data within the technology platforms is not being used for analysis.
Level 2: Ad Hoc	Standardized terminology is used to represent SEDoH concepts and data elements across the institution.
Level 3: Emerging	Internal reporting of SEDoH data, that is, reports and dashboards are created for internal consumption of the institution.
Level 4: Coordinated	Creating reports and dashboards for external agencies (for community benefits requirement, for federal agencies, etc.)
Level 5: Supported	SEDoH data are being used for waste and care variability reduction.
Level 6: Integrated	SEDoH data are being used for population health management, for example, identifying high-risk patients for targeted interventions.
Level 7: Transformative	SEDoH data inform precision medicine and prescriptive analytics initiatives resulting in a learning healthcare system.
Domain 5: Operational and strategic impact	
Level 1: Absent	Leadership is not committed to SEDoH data initiatives.
Level 2: Ad Hoc	SEDoH data are presented for research or projects only.

(*Continued*)

Table 2. (*Continued*)

Level	Description
Level 3: Emerging	Presenting SEDoH data at point of care to help clinicians make clinical decisions.
Level 4: Coordinated	SEDoH data inform department or division or clinic decisions within an organization (e.g., use of heat maps to assess high-prevalence asthma areas for establishing outpatient clinics).
Level 5: Supported	SEDoH services center (infrastructure and operational unit) that operates to support patient-specific SEDoH needs and delivery of services to address SEDoH.
Level 6: Integrated	SEDoH data guide programmatic implementation, that is, all organizational decisions (not just specific SEDoH initiatives) take into consideration the SEDoH data about the target population. For example, hiring Spanish staff for a population with 30% Spanish patients.
Level 7: Transformative	SEDoH data influence strategic and financial decisions at the highest organizational level. Organizational leadership is committed and makes decisions based on SEDoH data, including allocation of significant financial resources.

ACS = American community survey; = clinical decision support; CDS PDF = portable document format; EPA = Environmental Protection Agency; FHIR = fast healthcare interoperability resources; HER = electronic health record; HL7 = health level 7; SEDoH = social and environmental determinants of health; USDA = United States department of agriculture.
Data sources for social and environmental determinants of health (SEDoH) informatics maturity model include Person-level data (SEDoH data about an individual, ideally self-reported, and often collected using standardized and validated instruments) and Contextual data (Data about the environments in which a patient lives). All of the levels and descriptions outlined below apply to both personal and contextual data, except where noted.
***Primarily for contextual data, but also personal data where relevant;, patients are providing personal-level SDoH data to some other entity, and the health system in question is integrating it into their data.*

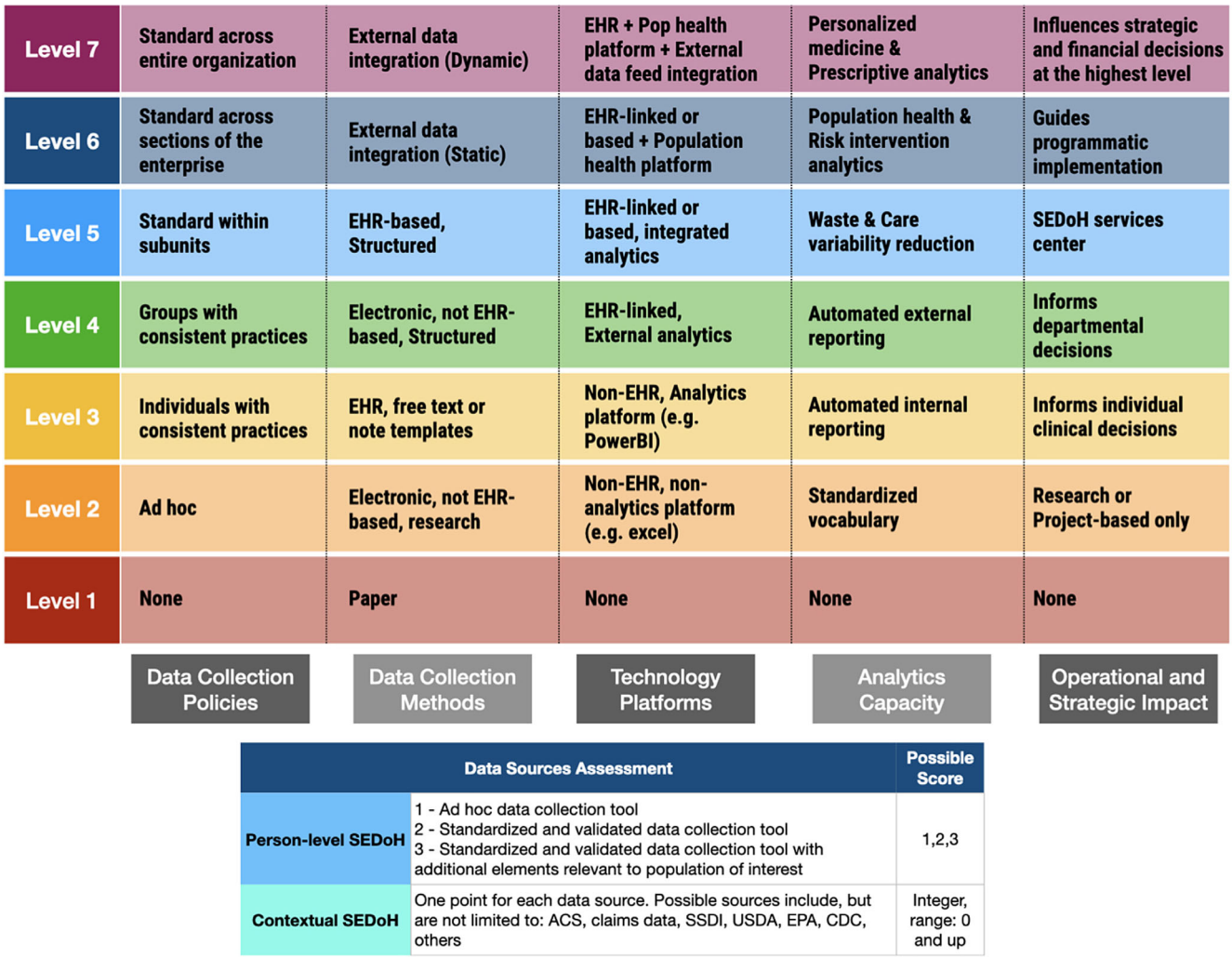

Figure 3. The social and environmental determinants of health informatics maturity model (SIMM) with Data Sources Assessment. EHR = electronic health record. SEDoH = social and environmental determinants of health, ACS = American Community Survey, SSDI = social security death index, USDA = United States Department of Agriculture, EPA = Environmental Protection Agency, CDC = Centers for Disease Control and Prevention.

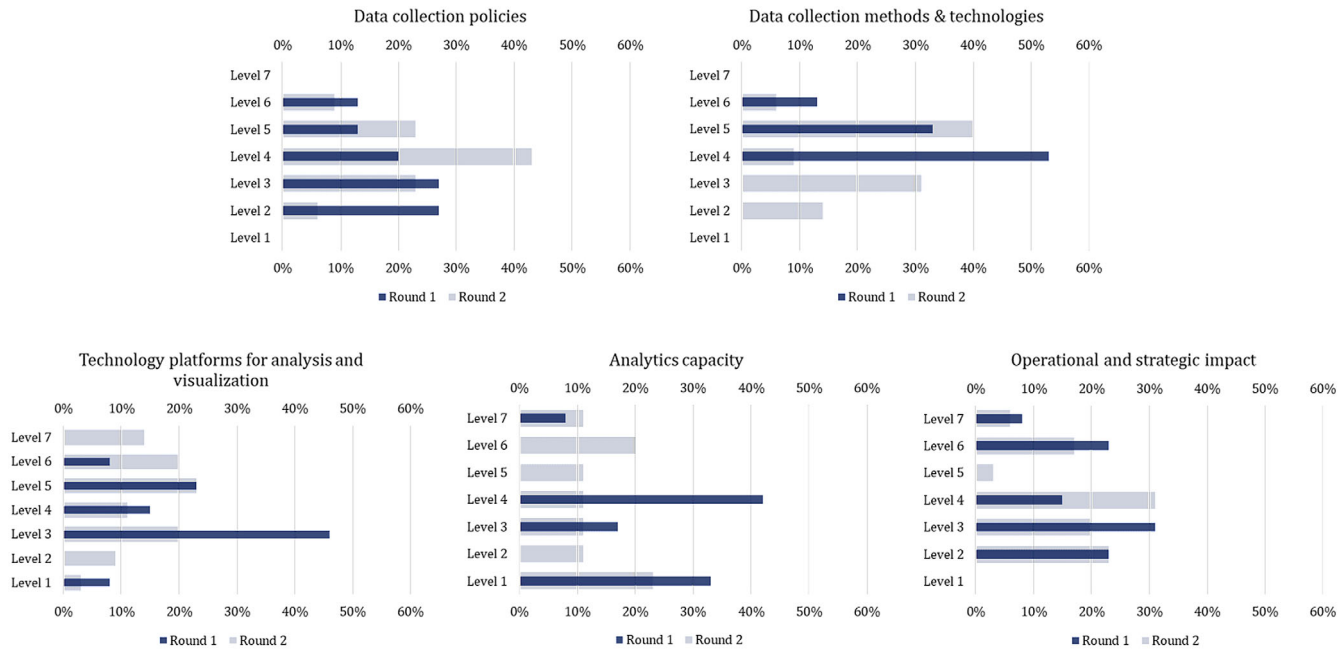

Figure 4. Comparison of responses to the social and environmental determinants of health informatics maturity model (SIMM) self-assessment survey between version 0.9 (First round, $n = 15$) and version 1.0 (Second round, $n = 35$).

of a gap analysis, and third, to outline steps that they can take to improve their current level of maturity.

The SEDoH maturity model consists of 35 maturity statements categorized under five domains of data collection policies, data collection methods and technologies, technology platforms for analysis and visualization, analytics capacity, and operational and strategic impact. Domain experts reviewed our SEDoH maturity model to evaluate the conceptual completeness of the model (i.e., appropriateness of the domains and levels, as well as usefulness and ease of use). Two rounds of assessment and feedback by experts at N3C, IEC, and across the CTSA network suggest reasonable real-world applicability.

The distributions of responses were broader and more even in all but one domain in the second evaluation round compared to the first round (Fig. 4), suggesting that the edits made to the SIMM tool for v1.0 increased its capacity to discriminate between stages of maturity. Our assessments reveal that a consistent finding across both stages of assessment was that none of the respondents suggested level 1 for three out of the five domains, (i.e., "Data collection policies," "Data collection methods and technologies," "Operational and strategic impact"). This could be seen as a manifestation of the initial commitment of many organizations to improve efforts related to SEDoH data. Another interpretation could be that existing maturity in other informatics domains might translate to a basic level of maturity for SEDoH data (no other informatics maturity assessments were performed, but this could be a reasonable assumption given that all respondents were from the CTSA network). Further, an important finding of the round one assessment was that organizations endorsed maturity levels 4, 5, and 6 for the "Data collection methods and technologies" dimension of the maturity model, with over half suggesting the middle-level maturity at level 4. This growing momentum likely reflects a confluence of several phenomena: first, an increasing recognition of the importance of SEDoH at a community, regional, and national level [29], especially since the COVID-19 pandemic, as well as recent initiatives leading to the adoption of solutions that connect healthcare and community partner stakeholders together to move the needle on social determinants. In addition, the recent federal reforms of health care and health information technology (IT) facilitate data collection by extending investment in electronic health records to healthcare providers that receive public funds [30], as well as incorporating social determinants into EHR to promote patient and population health [31]. Novel initiatives have emerged to address socioeconomic and non-medical health determinants within the framework of the healthcare delivery system. These encompass Medicaid efforts led by states or health plans, multipayer federal and state initiatives, and provider-level activities aimed at recognizing and resolving the social and non-medical needs of their patients [32].

Importance of SEDoH and SIMM: Now More Than Ever

Despite a shift in our understanding of health and its determinants over the past few decades and improvements in medical care and in disease prevention, population-level health inequalities in health-care result in $309 billion in losses to the economy annually, disproportionately affecting minoritized and underserved communities [33]. In recent years, public health leaders and researchers in the United States have increasingly recognized that medical care alone cannot adequately improve health overall or reduce health disparities without also addressing where and how people live [34]. Social, economic, and environmental factors influence health behaviors and are primary drivers of health outcomes. For instance, children born to parents who have not completed high school education are more likely to live in an environment that presents obstacles to healthy living, such as exposure to trash, substandard housing, unsafe living conditions, limited access to playgrounds, etc [35]. The COVID-19 pandemic highlighted on a national scale many of the existing inequities. According to the CDC, people of color had higher rates of infection, hospitalization, and death due to COVID-19, resulting from an increased risk of exposure to the virus due to living, working, and transportation

conditions, as well as witnessed increased barriers to treatment due to existing disparities in access to health care [36]. Capturing and collecting SEDoH data in clinical settings is essential to reduce health inequities, improve health, and control healthcare costs [37,38], according to a 2014 report from the Institute of Medicine [39]. For example, Hennepin County Medical Center standardized food insecurity screening at two outpatient clinics that used an EHR-based food resource referral system. It was observed that systematic screening increased Senior Care referrals by 1,450% ($P < .001$) and Pediatrics recorded a 275% referral increase ($P < .001$) [40]. Given the broad recognition of the increasing need to address SEDoH in healthcare, the SIMM is a timely and important tool to help identify areas where organizations are not operating optimally and allow them to determine strategies that can improve their operations and processes. Further, the impact of maturity model implementation can be learned from a recent mixed methods case study in Australia that indicated that higher digital health maturity was associated with better outcomes, including maintaining a patient health record, tracking patient experience data, tracking the patient journey, and mitigating the clinical risk [41].

Challenges in accessing and analyzing SEDoH data

While it is being increasingly accepted that SEDoH data should be routinely incorporated into clinical decision-making to improve patient outcomes and operational efficiency, collecting and analyzing this data remain inherently challenging. It is unclear whether or not SDoH are frequently recorded as part of the standard clinical care [42]. Our study findings also demonstrate that standardized data collection policies were endorsed only by individuals or small organizational subunits. Data harmonization across clinical, public health, and administrative datasets continues to be a significant challenge [43]. Moreover, while it is easy to free-text this information within the electronic health record (EHR), extracting unstructured data from narrative notes requires time-consuming manual chart reviews or the use of advanced natural language processing [44]. In addition, the healthcare industry primarily relies on administrative claims data to evaluate SDoH [45,]. However, it is important to realize that the documentation of this data within the EHR is inconsistent and spotty due to lack of standards for capturing this data in a structured format. To address these unmet needs, SEDoH extraction research from unstructured clinic notes has been increasing [46]. A variety of SEDoH, such as substance use, work, housing situation, environmental factors, physical activity, sexual factors, transportation, education, and language, have all been annotated in clinical corpora [47–50]. The i2b2 NLP Smoking Challenge introduced a corpus of 502 notes with tobacco use status labels [51]. As a tool, the SIMM can help organizations understand their current state as it relates to SEDoH data and develop a systematic approach to improvement, but they will still need to look to the literature, best practices, and guidelines for specific solutions, such as protocols to improve race and ethnicity data collection [52,53]. Ultimately, higher levels of SEDoH informatics maturity will increase the quality and timeliness of SEDoH data, thus allowing clinicians to make more informed, targeted medical decisions that deliver more equitable, high-quality care.

Taking into account the growing significance of SEDoH, informatics maturity models, and the pressing need to integrate SEDoH data into health informatics, our work describes a multidimensional approach to measure organizational SEDoH capability, including items for data collection policies, data collection methods, technology and analytics capability, and for operational and strategic impact. There are some limitations to our work. First, the participants in our evaluation rounds are not representative of all healthcare organizations in the US; they were primarily academics affiliated with CTSAs and/or had a pre-existing interest in SEDoH (such as the members of the N3C SDoH workgroup). Second, this work was conducted at a period of time when there was increased knowledge, awareness, and acceptance of the concept of SEDoH data and organizational maturity. Finally, our work did not include non-academic healthcare organizations and community organizations. The validity and utility of the SIMM outside of academic medical centers will need to be evaluated in future studies.

Conclusion

Clinical and economic motivations have encouraged healthcare providers, health insurance companies, local governments, and community-based organizations to take action to address SEDoH. The initial findings of this exploratory study suggest that the proposed SIMM is applicable for identifying an organization's SEDoH informatics capacity and providing a roadmap for moving to the next level. As a result of identifying the strong and weak points of the SEDoH data workflow, depending on the assessment findings, improvement opportunities can be identified by an organization. Additional work is still needed to evaluate the generalizability of the SEDoH maturity model and identify specific strategies and protocols to increase maturity.

Acknowledgments. We are grateful for the time and contributions of the faculty and staff who regularly participate in N3C, CD2H, IEC, and other CTSA Network activities who provided invaluable insights and responded to the self-assessment surveys. Study data were collected and managed using REDCap electronic data capture tools (Harris et al. 2009) hosted at the Institute of Translational Health Sciences (ITHS).

Funding statement. This work was supported by grants UL1TR001855 (SC CTSI), UL1TR001422 (NUCATS), and UL1TR002319 (ITHS) from the National Center for Advancing Translational Science (NCATS) of the U.S. National Institutes of Health. This work also received support from the National Center for Data to Health (CD2H) grant (NIH/NCATS U24TR002306) and supplemental funding from the National COVID Cohort Collaborative (NIH/NCATS U24TR002306-04S3). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Competing interests. JE is a paid consultant for Sanofi. The other authors declare no competing interests.

References

1. **World Health Organization**. Social determinants of health. (https://www.who.int/health-topics/social-determinants-of-health#tab=tab_1). Accessed March 22, 2023.
2. **Stringhini S, Carmeli C, Jokela M, *et al.*** Socioeconomic status and the 25 × 25 risk factors as determinants of premature mortality: a multicohort study and meta-analysis of 1·7 million men and women. *Lancet*. 2017;**389**(10075):1229–1237. doi: 10.1016/S0140-6736(16)32380-7.
3. **Hood CM, Gennuso KP, Swain GR, Catlin BB.** County health rankings: relationships between determinant factors and health outcomes. *Am J Prev Med*. 2016;**50**(2):129–135. doi: 10.1016/j.amepre.2015.08.024.
4. **Office of Disease Prevention and Health Promotion**. Healthy People 2030. Social Determinants of Health. (https://health.gov/healthypeople/priority-areas/social-determinants-health). Accessed March 20, 2023.

5. **Peretz PJ, Islam N, Matiz LA.** Community health workers and covid-19 - addressing social determinants of health in times of crisis and beyond. *N Engl J Med*. 2020;**383**(19):e108. doi: 10.1056/NEJMp2022641.
6. **Rollston R, Galea S.** COVID-19 and the social determinants of health. *Am J Health Promot*. 2020;**34**(6):687–689. doi: 10.1177/0890117120930536b.
7. **Shah GH, Shankar P, Schwind JS, Sittaramane V.** The detrimental impact of the COVID-19 crisis on health equity and social determinants of health. *J Public Health Manag Pract*. 2020;**26**(4):317–319. doi: 10.1097/PHH.0000000000001200.
8. **Ramírez IJ, Lee J.** COVID-19 emergence and social and health determinants in Colorado: a rapid spatial analysis. *Int J Environ Res Public Health*. 2020;**17**(11):3856. doi: 10.3390/ijerph17113856 Published 2020 May 29.
9. **Phuong J, Zampino E, Dobbins N**, ***et al.*** Extracting patient-level social determinants of health into the OMOP common data model. *AMIA Annu Symp Proc*. 2022;**2021**:1989–1998.
10. **Kingsbury P, Abajian H, Abajian M**, ***et al.*** SEnDAE: a resource for expanding research into social and environmental determinants of health. *Comput Methods Programs Biomed*. 2023;**238**:107542. doi: 10.1016/j.cmpb.2023.107542.
11. **Chen M, Tan X, Padman R.** Social determinants of health in electronic health records and their impact on analysis and risk prediction: a systematic review. *J Am Med Inform Assoc*. 2020;**27**(11):1764–1773. doi: 10.1093/jamia/ocaa143.
12. **Carvalho JV, Rocha Á., Abreu A.** Maturity models of healthcare information systems and technologies: a literature review. *J Med Syst*. 2016;**40**(6):131. doi: 10.1007/s10916-016-0486-5.
13. **Electronic Medical Record Adoption Model (EMRAM)**. HIMSS. (https://www.himss.org/what-we-do-solutions/digital-health-transformation/maturity-models/electronic-medical-record-adoption-model-emram). Accessed March 24, 2023.
14. **Continuity of Care Maturity Model (CCMM)**. HIMSS. (https://www.himss.org/what-we-do-solutions/digital-health-transformation/maturity-models/continuity-care-maturity-model-ccmm). Accessed March 28, 2023.
15. **Sharma, B.**, Electronic Healthcare Maturity Model (eHMM). White Paper. Quintegra Solutions Limited, 2008. (http://www.quintegrasolutions.com/eHMM%20White%20Paper.pdf) Accessed March 27, 2023.
16. **Knosp BM, Barnett WK, Anderson NR, Embi PJ.** Research IT maturity models for academic health centers: early development and initial evaluation. *J Clin Transl Sci*. 2018;**3**(5):289–294. doi: 10.1017/cts.2018.339.
17. **Knosp BM, Dorr DA, Campion TR.** Maturity in enterprise data warehouses for research operations: analysis of a pilot study. *J Clin Transl Sci*. 2023;**7**(1):e70. doi: 10.1017/cts.2023.23.
18. **Champieux R, Solomonides A, Conte M**, ***et al.*** Ten simple rules for organizations to support research data sharing. *PLoS Comput Biol*. 2023;**19**(6):e1011136. doi: 10.1371/journal.pcbi.1011136.
19. **van de Wetering R, Batenburg R.** A PACS maturity model: a systematic meta-analytic review on maturation and evolvability of PACS in the hospital enterprise. *Int J Med Inform*. 2009;**78**(2):127–140. doi: 10.1016/j.ijmedinf.2008.06.010.
20. **Sanders, D**. The Healthcare Analytics Adoption Model: A Roadmap to Analytic Maturity (white paper). 2020. (https://www.healthcatalyst.com/learn/white-papers/the-healthcare-analytics-adoption-model-a-roadmap-to-analytic-maturity) Accessed March 18, 2023.
21. **Miff S**. THE TIME IS NOW FOR HEALTH SYSTEMS TO GET SERIOUS ABOUT SOCIAL DETERMINANTS OF HEALTH. (https://pccinnovation.org/the-time-is-now-for-health-systems-to-get-serious-about-social-determinants-of-health/) Accessed Febraury 24, 2023.
22. **Phuong J, Hong S, Palchuk MB**, ***et al.*** Advancing interoperability of patient-level social determinants of health data to support COVID-19 research. *AMIA Jt Summits Transl Sci Proc*. 2022;**2022**:396–405.
23. **Madlock-Brown C, Wilkens K, Weiskopf N**, ***et al.*** Clinical, social, and policy factors in COVID-19 cases and deaths: methodological considerations for feature selection and modeling in county-level analyses. *BMC Public Health*. 2022;**22**(1):747. doi: 10.1186/s12889-022-13168-y.
24. **Phuong J, Riches NO, Madlock-Brown C**, ***et al.*** Social determinants of health factors for gene-environment COVID-19 research: challenges and opportunities. *Adv Genet (Hoboken)*. 2022;**3**(2):2100056. doi: 10.1002/ggn2.202100056.
25. **Advancing the Application of Maturity Models**. (https://cd2h.org/node/156) Accessed June 27, 2023.
26. **Feldstein AC, Glasgow RE.** A practical, robust implementation and sustainability model (PRISM) for integrating research findings into practice. *Jt Comm J Qual Patient Saf*. 2008;**34**(4):228–243. doi: 10.1016/s1553-7250(08)34030-6.
27. **Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J.** The updated consolidated framework for implementation research based on user feedback. *Implement Sci*. 2022;**17**(1):75. doi: 10.1186/s13012-022-01245-0 Published 2022 Oct 29.
28. **United States Census Bureau**. American Community Survey (ACS). (https://www.census.gov/programs-surveys/acs). Accessed October 25, 2023.
29. **Fuchs VR.** Social determinants of health: caveats and nuances. *JAMA*. 2017;**317**(1):25–26. doi: 10.1001/jama.2016.17335.
30. **Siegel B, Nolan L.** Leveling the field–ensuring equity through national health care reform. *N Engl J Med*. 2009;**361**(25):2401–2403. doi: 10.1056/NEJMp0909323.
31. **Bazemore AW, Cottrell EK, Gold R**, ***et al.*** Community vital signs': incorporating geocoded social determinants into electronic records to promote patient and population health. *J Am Med Inform Assoc*. 2016;**23**(2):407–412. doi: 10.1093/jamia/ocv088.
32. **Alley DE, Asomugha CN, Conway PH, Sanghavi DM**. Accountable health communities—addressing social needs through medicare and medicaid. *N Engl J Med*. 2016;**374**(1):8–11.
33. **Ubri P, Artiga A**. Disparities in health and health care: five key questions. The Henry J. Kaiser Family Foundation. (http://kff.org/disparities-policy/issue-brief/disparities-in-health-and-health-care-five-key-questions-and-answers) Accessed March 27, 2022.
34. **Braveman P, Egerter S, Williams DR.** The social determinants of health: coming of age. *Annu Rev Public Health*. 2011;**32**(1):381–398. doi: 10.1146/annurev-publhealth-031210-101218.
35. **Singh GK, Siahpush M, Kogan MD.** Neighborhood socioeconomic conditions, built environments, and childhood obesity. *Health Aff (Millwood)*. 2010;**29**(3):503–512. doi: 10.1377/hlthaff.2009.0730.
36. **CDC, Risk for COVID-19 Infection, Hospitalization, and Death By Race/Ethnicity**. (https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-discovery/hospitalization-death-by-race-ethnicity.html) Accessed April 11, 2023.
37. **Singh GK, Daus GP, Allender M**, ***et al.*** Social determinants of health in the united states: addressing major health inequality trends for the nation, 1935-2016. *Int J MCH AIDS*. 2017;**6**(2):139–164. doi: 10.21106/ijma.236.
38. **Blizinsky KD, Bonham VL.** Leveraging the learning health care model to improve equity in the age of genomic medicine. *Learn Health Syst*. 2018;**2**(1):e10046. doi: 10.1002/lrh2.10046.
39. **Committee on the Recommended Social and Behavioral Domains and Measures for Electronic Health Records; Board on Population Health and Public Health Practice; Institute of Medicine**. *Capturing Social and Behavioral Domains and Measures in Electronic Health Records: Phase 2*. Washington (DC): National Academies Press (US); 2015,https://www.ncbi.nlm.nih.gov/books/NBK268995/ doi: 10.17226/18951.
40. **Hager K, De Kesel Lofthus A, Balan B, Cutts D.** Electronic medical record-based referrals to community nutritional assistance for food-insecure patients. *Ann Fam Med*. 2020;**18**(3):278–278. doi: 10.1370/afm.2530.
41. **Woods L, Dendere R, Eden R**, ***et al.*** Perceived impact of digital health maturity on patient experience, population health, health care costs, and provider experience: mixed methods case study. *J Med Internet Res*. 2023;**25**:e45868. doi: 10.2196/45868.
42. **Garg A, Boynton-Jarrett R, Dworkin PH.** Avoiding the unintended consequences of screening for social determinants of health. *JAMA*. 2016;**316**(8):813–814. doi: 10.1001/jama.2016.9282.
43. **Camacho-Rivera Marlene, Jessica Y, *et al.*** Social determinants of health during the COVID-19 pandemic in the US: precision through context. In: Hsueh PYS, Wetter T, Zhu X, eds. *Personal Health Informatics. Cognitive Informatics in Biomedicine and Healthcare*. Cham: Springer; 2022:397–425.
44. **Dorr D, Bejan CA, Pizzimenti C, Singh S, Storer M, Quinones A.** Identifying patients with significant problems related to social

determinants of health with natural language processing. *Stud Health Technol Inform*. 2019;**264**:1456–1457. doi: 10.3233/SHTI190482.

45. **Hatef E, Rouhizadeh M, Tia I**, ***et al.*** Assessing the availability of data on social and behavioral determinants in structured and unstructured electronic health records: a retrospective analysis of a multilevel health care system. *JMIR Med Inform*. 2019;**7**(3):e13802. doi: 10.2196/13802.
46. **Patra BG, Sharma MM, Vekaria V**, ***et al.*** Extracting social determinants of health from electronic health records using natural language processing: a systematic review. *J Am Med Inform Assoc*. 2021;**28**(12):2716–2727. doi: 10.1093/jamia/ocab170.
47. **Lybarger K, Ostendorf M, Yetisgen M.** Annotating social determinants of health using active learning, and characterizing determinants using neural event extraction. *J Biomed Inform*. 2021;**113**:103631. doi: 10.1016/j.jbi.2020.103631.
48. **Chapman AB, Jones A, Kelley AT**, ***et al.*** ReHouSED: a novel measurement of veteran housing stability using natural language processing. *J Biomed Inform*. 2021;**122**:103903. doi: 10.1016/j.jbi.2021.103903.
49. **Lybarger K, Yetisgen M, Uzuner Ö.** The 2022 n2c2/UW shared task on extracting social determinants of health. *J Am Med Inform Assoc*. 2023;**30**(8):1367–1378. doi: 10.1093/jamia/ocad012.
50. **Han S, Zhang RF, Shi L**, ***et al.*** Classifying social determinants of health from unstructured electronic health records using deep learning-based natural language processing. *J Biomed Inform*. 2022;**127**:103984. doi: 10.1016/j.jbi.2021.103984.
51. **Uzuner O, Goldstein I, Luo Y, Kohane I.** Identifying patient smoking status from medical discharge records. *J Am Med Inform Assoc*. 2008;**15**(1):14–24. doi: 10.1197/jamia.M2408.
52. **Ulmer C, McFadden B, Nerenz DR.** Institute of medicine (US) subcommittee on standardized collection of race/ethnicity data for healthcare quality improvement. *Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement*. Washington (DC): National Academies Press (US); 2009.
53. **Vega Perez RD, Hayden L, Mesa J**, ***et al.*** Improving patient race and ethnicity data capture to address health disparities: a case study from a large urban health system. *Cureus*. 2022;**14**(1):e20973. doi: 10.7759/cureus.20973.