The Sociobehavioral Phenotype: Applying a Precision Medicine Framework to Social Determinants of Health

Ravi B. Parikh, MD, MPP; Sachin H. Jain, MD, MBA; and Amol S. Navathe, MD, PhD

Commonly used risk algorithms in healthcare, including indices to predict hospital readmission, define "high-risk" patients by disease-based (eg, comorbidity index) or utilization-based (eg, number of emergency department [ED] visits) factors.[1,2] Clinicians and systems can target interventions to such high-risk groups—so-called "precision delivery."[3] However, current programs that target high-risk individuals, including high-risk case management programs, do not uniformly reduce utilization or adverse outcomes.[4]

Why might such interventions be ineffective? Perhaps the problem is in our conception of risk, rather than the interventions.

An individual's health phenotype has far-reaching connections with social, behavioral, and environmental factors, such as housing instability and food insecurity. Nevertheless, few risk models include variables associated with social determinants of health, such as income level or educational attainment, and traditional risk models often identify such individuals when it is too late to intervene upon socioeconomic risk factors.[5] To complement risk-stratification models based on clinical severity, we introduce the "sociobehavioral phenotype"—actionable risk profiles based on empirically derived social, economic, and behavioral determinants of health.

Defining the Sociobehavioral Phenotype: Lessons From Precision Medicine

To build and implement sociobehavioral phenotypes, clinicians, researchers, and systems leaders can draw lessons from the field of precision medicine. Consider the analogy of metastatic non–small cell lung cancer (NSCLC): Prior to the advent of targeted and immune-based therapy, cytotoxic chemotherapy was the standard of care. Genomic sequencing of tumor cells identified pathologic mutations; researchers later targeted such mutations in drug development. Subsequent development of targeted inhibitors and evaluation in randomized trials led to many more therapeutic options becoming available for NSCLC. As new therapies targeting pathologic "driver" mutations became available, oncologists integrated screening for such mutations into the workup of NSCLC.

Just as many driver mutations exist in lung cancer, driver socioeconomic risk factors (eg, poor social support, housing instability) or behaviors (eg, medication nonadherence) underlie many individuals' health outcomes, such as mortality or acute care utilization. Clinicians and researchers can analyze large observational data sets to classify 1 or more of these common risk factors into distinct sociobehavioral phenotypes and subsequently test interventions targeted toward those phenotypes.

The Added Value of the Sociobehavioral Phenotype

The sociobehavioral phenotype aligns well with broader calls for providers to screen for social determinants of health. In resource-constrained environments, it is of paramount importance to target interventions to the individuals most likely to derive benefit from them. However, current interventions may not serve the majority of patients with predominantly sociobehavioral contributors to poor health. Take, for example, a telephonic case management program targeted to individuals with high likelihoods of hospital readmission. Although such a program may benefit some subtypes of patients, it is unlikely that a telephone-based program would succeed for a homeless individual who does not have a reliable address or telephone number. An unknown phone number may also go unanswered due to fear or suspicion. Furthermore, telephonic or remote case management programs may fail to lower health expenditures because certain high-risk patients rely on frequent in-person contact.[4] Sociobehavioral phenotyping based on electronic health record (EHR) data may identify driver risk factors to develop and refine effective interventions for such individuals.

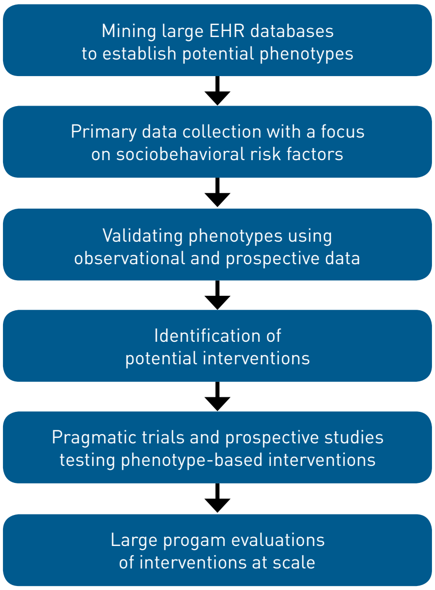
A Coordinated Research Agenda

Learning from the precision medicine example, a coordinated research agenda with parallel and iterative data mining and experimentation could identify relevant sociobehavioral phenotypes that ameliorate social determinants of health (**Figure**). Mining large databases from EHRs and observational studies are necessary first steps to identify sociobehavioral phenotypes. There are examples of this: Using a large sample of clinical notes from the Veterans Health

TAKEAWAY POINTS

- Social determinants of health lead to increased healthcare utilization and poor outcomes, but efforts to risk-stratify patients have largely ignored these risk factors.
- Advanced analytics applied to electronic health records and primary survey collection can reveal distinct phenotypes of individuals with actionable sociobehavioral contributors to health.
- Healthcare organizations can target these sociobehavioral phenotypes with evidence-based interventions.
- We propose a research agenda to identify, validate, and apply sociobehavioral phenotypes in healthcare.

FIGURE. A Research Agenda for Establishing and Generalizing Sociobehavioral Phenotypes



EHR indicates electronic health record.

Administration (VHA), researchers have developed and validated natural language processing lexicons to identify individuals who are homeless or at high risk of being homeless.[6] Additionally, neural network methods applied to data from the Health and Retirement Study have yielded potential sociobehavioral phenotypes, such as elderly patients with few children and thus little social support, that are predictive of poor health outcomes.[7] Clinicians and researchers will need to supplement these EHR-derived phenotypes with primary survey data collection focused on driver social and behavioral risk factors, perhaps as part of regular screens for socioeconomic determinants of health.

After defining and validating sociobehavioral phenotypes, clinicians and delivery scientists can identify which phenotypes are actionable by developing randomized trials of interventions in specific subgroups. Unlike the situation in metastatic lung cancer a decade ago, healthcare delivery scientists have an armamentarium of community-based interventions to use among selected high-risk individuals. These include home-based and visiting nurse care, telephonic case management, wireless technologies, community health workers, patient-centered medical homes, social service programs, and direct financial assistance. By examining the effect of a social service or patient engagement intervention stratified by a sociobehavioral phenotype, policy makers and clinicians can determine which subpopulations derive benefit from each intervention, which can facilitate more efficient enrollment into these interventions (**Table**).

Nascent examples of sociobehavioral phenotyping are a harbinger of broader success. Since 2013, the VHA has deployed a screening for homelessness, the Homelessness Screening Clinical Reminder, for all patients who access outpatient VHA care.[8] Applying machine learning models to VHA EHR data reliably identifies patient phenotypes at high risk (>90%) of reporting homelessness, a risk factor for acute care utilization.[9] For individuals at high risk of homelessness, the VHA deploys interventions, including a homelessness primary care action team and temporary housing, that are associated with reduced ED utilization among those with previously high utilization.[10]

Identifying Heterogeneity Within Phenotypes

As in a genomics-based precision medicine framework, there will still be heterogeneity within a sociobehavioral phenotype. For example, the aforementioned examples of sociobehavioral phenotyping define phenotypes by singular socioeconomic risk factors, such as homelessness. Although homelessness could contribute to healthcare utilization, reasons for homelessness are varied. The single male who faces chronic housing insecurity due to substance use and mental illness would likely have a different set of needs than the mother of 4 children who is fleeing domestic violence. Each of these reasons demands different interventions to address underlying causes of social instability. Unsupervised clustering algorithms applied to very large data sets could help identify heterogeneity within socioeconomic risk factor groups. Such algorithms may differentiate between the constellation of homelessness, substance abuse, and poor primary care follow-up versus a constellation of homelessness, food insecurity, and diabetes. By leveraging data across a large number of patients, delivery scientists and policy makers can identify combinations of social factors that result in a particular pattern of morbidity, utilization, and cost, then prioritize different interventions or sequences of interventions for such phenotypes.

Limitations

There are several challenges in defining and implementing sociobehavioral phenotypes. Patient-level socioeconomic data are often not readily available in structured data sets. Additionally, accounting for socioeconomic risk factors in individual patient care may increase information overload among overstretched clinicians—who may

TABLE. Illustrative Associations Between Phenotypes and Interventions in Precision Medicine and Precision Delivery Frameworks

	Precision Medicine	Precision Delivery			
Problem	Death from metastatic NSCLC	Hospitalization, readmission, or avoidable adverse outcomes due to socioeconomic risk factors			
Risk phenotype	"Driver" mutation in gene encoding EGFR	Inadequate access to transportation	Food insecurity	Medication nonadherence due to lack of health literacy	Housing instability due to low income
Intervention	EGFR inhibitor	Discounted fares for ride-sharing services	Enrollment in a home-delivered meals program	Telephonic case management program and/or visiting nurse program for medication reminders	Enrollment in low-income housing program

EGFR indicates epidermal growth factor receptor; NSCLC, non–small cell lung cancer.

have trouble managing the data to which they already have access. And certain sociobehavioral phenotypes, such as food insecurity, carry risks of stigma and further alienating patients. Patients who fall into certain sociobehavioral phenotypes will need protections to ensure that they are not excluded from appropriate care or subjected to bias. Partnerships between academia and patient and advocacy groups could aid in refining data-generated phenotypes and associated interventions to ensure maximal uptake and acceptance.

CONCLUSIONS

Targeting interventions based on their ability to improve specific sociobehavioral risk factors is more likely to be successful than broad—and often blind—application of resource-constrained interventions. Just as oncologists in the precision medicine era screen for actionable mutations to guide use of targeted therapies, delivery scientists and clinicians can study and target interventions based on sociobehavioral phenotypes. Perhaps then we may experience a revolution in addressing socioeconomic determinants of health—similar to the precision medicine revolution of today. ■

Author Affiliations: Perelman School of Medicine and Leonard Davis Institute of Health Economics, University of Pennsylvania (RBP, ASN), Philadelphia, PA; Corporal Michael J. Cresencz VA Medical Center (RBP, ASN), Philadelphia, PA; CareMore Health System/Anthem Inc (SHJ), Cerritos, CA; Stanford University School of Medicine (SHJ), Palo Alto, CA.

Source of Funding: This work was funded in part by the Pennsylvania Commonwealth University Research Enhancement Program, the Leonard Davis Institute Policy Accelerator Program, and the Pennsylvania Department of Health. The department specifically disclaims responsibility for any analyses, interpretations, or conclusions.

Author Disclosures: Dr Jain is employed with CareMore Health. Dr Navathe is a noncompensated board member of Integrated Services, Inc; is a consultant or paid advisor for Navvis and Co, Agathos, and Navahealth; has received grants from Hawaii Medical Service Association, Oscar, Cigna, The Commonwealth Fund, and the Robert Wood Johnson Foundation; has received honoraria from Elsevier for his editorial role and from the National University Health System (Singapore) for his role as an advisor; has received speaker fees and travel from Cleveland Clinic; and is employed with University of Pennsylvania. Dr Parikh reports no relationship or financial interest with any entity that would pose a conflict of interest with the subject matter of this article.

Authorship Information: Concept and design (RBP, ASN); acquisition of data (RBP); analysis and interpretation of data (RBP); drafting of the manuscript (RBP, ASN); critical revision of the manuscript for important intellectual content (RBP, SHJ, ASN); obtaining funding (RBP); administrative, technical, or logistic support (RBP, SHJ, ASN); and supervision (SHJ, ASN).

Address Correspondence to: Amol S. Navathe, MD, PhD, Division of Health Policy, University of Pennsylvania, 1108 Blockley Hall, 423 Guardian Dr, Philadelphia, PA 19146. Email: amol@wharton.upenn.edu.

REFERENCES

1. van Walraven C, Dhalla IA, Bell C, et al. Derivation and validation of an index to predict early death or unplanned readmission after discharge from hospital to the community. *CMAJ*. 2010;182(6):551-557. doi: 10.1503/cmaj.091117.
2. Billings J, Blunt I, Steventon A, Georghiou T, Lewis G, Bardsley M. Development of a predictive model to identify inpatients at risk of re-admission within 30 days of discharge (PARR-30). *BMJ Open*. 2012;2(4):e001667. doi: 10.1136/bmjopen-2012-001667.
3. Parikh RB, Kakad M, Bates DW. Integrating predictive analytics into high-value care: the dawn of precision delivery. *JAMA*. 2016;315(7):651-652. doi: 10.1001/jama.2015.19417.
4. Peikes D, Chen A, Schore J, Brown R. Effects of care coordination on hospitalization, quality of care, and health care expenditures among Medicare beneficiaries: 15 randomized trials. *JAMA*. 2009;301(6):603-618. doi: 10.1001/jama.2009.126.
5. Adler NE, Cutler DM, Fielding JE, et al. Addressing social determinants of health and health disparities: a vital direction for health and health care. National Academy of Medicine website. nam.edu/addressing-social-determinants-of-health-and-health-disparities-a-vital-direction-for-health-and-health-care/. Published September 19, 2016. Accessed May 1, 2019.
6. Gundlapalli AV, Redd A, Carter M, et al. Validating a strategy for psychosocial phenotyping using a large corpus of clinical text. *J Am Med Inform Assoc*. 2013;20(e2):e355-e364. doi: 10.1136/amiajnl-2013-001946.
7. Seligman B, Tuljapurkar S, Rehkopf D. Machine learning approaches to the social determinants of health in the Health and Retirement Study. *SSM Popul Health*. 2017;4:95-99. doi: 10.1016/j.ssmph.2017.11.008.
8. Byrne T, Fargo JD, Montgomery AE, Roberts CB, Culhane DP, Kane V. Screening for homelessness in the Veterans Health Administration: monitoring housing stability through repeat screening. *Public Health Rep*. 2015;130(6):684-692. doi: 10.1177/003335491513000618.
9. Byrne T, Montgomery AE, Fargo JD. Predictive modeling of housing instability and homelessness in the Veterans Health Administration. *Health Serv Res*. 2019;54(1):75-85. doi: 10.1111/1475-6773.13050.
10. Gundlapalli AV, Redd A, Bolton D, et al. Patient-aligned care team engagement to connect veterans experiencing homelessness with appropriate health care. *Med Care*. 2017;55(suppl 9 suppl 2):S104-S110. doi: 10.1097/MLR.0000000000000770.

Visit **ajmc.com/link/4198** to download PDF