

UC Berkeley

UC Berkeley Previously Published Works

Title

The Science of Precision Prevention: Research Opportunities and Clinical Applications to Reduce Cardiovascular Health Disparities.

Permalink

<https://escholarship.org/uc/item/8qd7m0zg>

Journal

JACC: Advances, 3(1)

Authors

Pearson, Thomas

Vitalis, Debbie

Pratt, Charlotte

et al.

Publication Date

2024

DOI

10.1016/j.jacadv.2023.100759

Peer reviewed



HHS Public Access

Author manuscript

JACC Adv. Author manuscript; available in PMC 2024 February 19.

Published in final edited form as:

JACC Adv. 2024 January ; 3(1): . doi:10.1016/j.jacadv.2023.100759.

The Science of Precision Prevention:

Research Opportunities and Clinical Applications to Reduce Cardiovascular Health Disparities

Thomas A. Pearson, MD, PhD, MPH^a, Debbie Vitalis, MPH, PhD^b, Charlotte Pratt, MSC, PhD^b, Rebecca Campo, PhD^b, Antonis A. Armoundas, PhD^c, David Au, MD, MS^d, Bettina Beech, DrPH, MPH^e, Olga Brazhnik, PhD^b, Christopher G. Chute, MD, DrPH, MPH^f, Karina W. Davidson, PhD^{g,h}, Ana V. Diez-Roux, MD, PhD, MPHⁱ, Lawrence J. Fine, MD, DrPH^b, Davera Gabriel, RN^j, Peter Groenveld, MD, MS^k, Jaclyn Hall, PhD^l, Alison B. Hamilton, PhD, MPH^m, Hui Hu, PhDⁿ, Heng Ji, MA, MS, PhD^o, Amy Kind, MD, PhD^p, William E. Kraus, MD^q, Harlan Krumholz, MD, SM^r, George A. Mensah, MD^s, Raina M. Merchant, MD, MSHP^k, Dariush Mozaffarian, MD, MPH, DrPH^t, David M. Murray, PhD^u, Dianne Neumark-Sztainer, PhD, MPH^v, Maya Petersen, MD, PhD^{w,x}, David Goff, MD, PhD^b

^aCollege of Medicine and College of Public Health and Health Professions, University of Florida Health Science Center, Gainesville, Florida, USA;

^bDivision of Cardiovascular Sciences, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland, USA;

^cCardiovascular Research Center, Massachusetts General Hospital and Broad Institute, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA;

^dCenter of Innovation for Veteran-Centered and Value-Driven Care, University of Washington, Seattle, Washington, USA;

^eUH Population Health, University of Houston, Houston, Texas, USA;

^fJohns Hopkins Medicine, Institute for Clinical and Translational Research, Baltimore, Maryland, USA;

^gInstitute of Health System Science, Feinstein Institutes for Medical Research, Northwell Health, New Hyde Park, New York, USA;

^hFeinstein Institutes for Medical Research, Northwell Health, Manhasset, New York, USA;

ⁱUrban Health Collaborative, Drexel Dornsife School of Public Health, Philadelphia, Pennsylvania, USA;

^jBiomedical Informatics and Data Science Section, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA;

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

ADDRESS FOR CORRESPONDENCE: Dr Thomas A. Pearson, Department of Epidemiology, University of Florida, 2004 Mowry Road, Gainesville, Florida 32610-0231, USA. tapearson@ufl.edu.

The authors attest they are in compliance with human studies committees and animal welfare regulations of the authors' institutions and Food and Drug Administration guidelines, including patient consent where appropriate. For more information, visit the [Author Center](#).

^kCenter for Health Care Transformation and Innovation, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA;

^lDepartment of Health Outcomes and Biomedical Informatics, Institute for Child Health Policy, College of Medicine, University of Florida, Gainesville, Florida, USA;

^mCenter for the Study of Healthcare Innovation, Implementation & Policy, VA Greater Los Angeles Healthcare System, Los Angeles, California, USA;

ⁿChanning Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA;

^oDepartment of Computer Science, University of Illinois Urbana-Champaign, Champaign, Illinois, USA;

^pCenter for Health Disparities Research (CHDR), University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin, USA;

^qDuke Molecular Physiology Institute, School of Medicine, Duke University, Durham, North Carolina, USA;

^rInstitute for Social and Policy Studies, of Investigative Medicine and of Public Health (Health Policy), Yale University, New Haven, Connecticut, USA;

^sCenter for Translation Research and Implementation Science, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland, USA;

^tFriedman School of Nutrition Science & Policy, Tufts University, Medford, Massachusetts, USA;

^uOffice of Disease Prevention, National Institutes of Health, Bethesda, Maryland, USA;

^vDivision of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, Minnesota, USA;

^wDivision of Biostatistics, and UCSF-UC Berkeley Program in Computational Precision Health, School of Public Health, University of California-Berkeley, Berkeley, California, USA;

^xUniversity of California-San Francisco, San Francisco, California, USA.

Abstract

Precision prevention embraces personalized prevention but includes broader factors such as social determinants of health to improve cardiovascular health. The quality, quantity, precision, and diversity of data relatable to individuals and communities continue to expand. New analytical methods can be applied to these data to create tools to attribute risk, which may allow a better understanding of cardiovascular health disparities. Interventions using these analytic tools should be evaluated to establish feasibility and efficacy for addressing cardiovascular disease disparities in diverse individuals and communities. Training in these approaches is important to create the next generation of scientists and practitioners in precision prevention. This state-of-the-art review is based on a workshop convened to identify current gaps in knowledge and methods used in precision prevention intervention research, discuss opportunities to expand trials of implementation science to close the health equity gaps, and expand the education and training of a diverse precision prevention workforce.

Keywords

data science; health equity; health promotion; implementation science; personalized medicine; precision analytics

The development of new types of data and data analytic tools promises new opportunities across the spectrum from preventive to curative medicine. Precision prevention has emerged with the aim to use data analytics to identify and provide the right preventive intervention for the right person or population at the right time and place. This approach moves from strategies based on average effects and risks to individual effects and risks.

As applied to the prevention of cardiovascular disease (CVD), there are many opportunities and challenges with the analysis of new types of individual and community data using qualitative and quantitative methods to develop more precise tools and strategies to prevent CVD. Although advances in cardiovascular prevention and care have led to a previous decline in CVD, the toll remains high. Reductions in age-standardized CVD mortality rates have stalled and even reversed in younger age groups, related to stagnant or increasing prevalences of cardiometabolic diseases such as obesity, diabetes, and hypertension with persistent geographic, economic, racial, and ethnic health disparities.¹ Data science can potentially personalize prevention approaches in ways that could accelerate progress in addressing these and other health disparities. Thus, there is a timely need to test the efficacy, effectiveness, and implementation of precision prevention for health and health equity.^{2,3}

These opportunities and challenges were the focus of a workshop on “*The Science of Precision Prevention: Research Opportunities to Reduce Disparities in Cardiovascular Health*” convened virtually by the National Heart, Lung, and Blood Institute (NHLBI) on December 5 to 7, 2022.⁴ The theme of the workshop aligns with several objectives in the NHLBI Strategic Vision.³ To develop this theme, the Workshop had 5 objectives:

- Convene investigators with diverse expertise in biological (basic, behavioral, population), clinical (ethics, disparities research, clinical trials, and implementation), and qualitative and quantitative (artificial intelligence (AI), biostatistics) sciences.
- Identify gaps in knowledge and methods as foci for future research and training.
- Discuss research opportunities and challenges that could potentially expand the number of trials that develop and test the application of precision data science tools.
- Expand ways to use precision prevention approaches to close the health equity gap.
- Explore ways to enhance the acquisition of precision prevention knowledge and skills in the training of the next generation of diverse researchers and other clinical and public health professionals.

This state-of-the-art review summarizes the information and discussions presented at the 3-day workshop, which consisted of the following sessions/themes: Advances in Data

Science; Development and Validation of Tools and Their Applications; Application of Data Science in Interventions for Diverse Individuals and Communities; and Education and Skills Development and Training in Precision Prevention for a Diverse Workforce. The workshop also emphasized the convergence of data science and implementation science and the need to go beyond traditional approaches to embrace equity-focused, clinical- and community-engaged implementation science. Each session included presentations from experts in the field and break-out/discussion sessions. A major goal was to define the state of precision prevention science and its best practices in research and training. Gaps in knowledge were identified as well as opportunities in precision prevention science research to address them. Trials employing precision prevention tools were discussed, along with study designs to address unmet burdens of CVD and the associated disparities in these burdens across communities.

DEVELOPMENT OF THE PRECISION PREVENTION CONCEPT

The focus on precision prevention builds on a prior *JACC* state-of-the-art review on precision health analytics⁵ which identified new sources and types of data from individuals and the application of analytical methods to create tools to support decisions by patients and clinicians in the well-established field of Precision Medicine. The review also identified new data resources and data science methods to create surveillance and health education tools for communities in the relatively newer field of Precision Public Health. Figure 1 presents the convergence of the multiple dimensions of Data Science, Precision Analytics, and Implementation Science, which includes health promotion, and system science approaches to create a more comprehensive body of knowledge, Precision Health. In this model, health promotion and disease prevention for conditions such as heart, lung, blood, and sleep disorders might be addressed at the individual, community, or a combination of levels.

This integration of individual- and community-level data, precision analytics, and implementation science supports current definitions of precision prevention.⁶⁻⁸ These expand the data types used in precision prevention to include “biologic, behavioral, socioeconomic, and epidemiologic data to describe and implement strategies tailored to reducing health conditions in specific individuals or populations.”⁶ precision prevention should be adapted to individuals’ or families’ differing economic or cultural circumstances.⁷ Precision Prevention integrates public health with health professionals’ care by considering socioeconomic, psychological/behavioral, and environmental/cultural factors in creating programs to improve the quality of life of individuals and populations.⁸ The inclusion of social determinants of health (SDoH) in this framework recognizes the powerful effects of structural factors that influence health, such as related to geography, income, education, and structural racism. This integration⁸ underscores the opportunity for precision prevention to mitigate health disparities in individuals and communities. Figure 2 illustrates this integration of individual- and community-level data analytics, with implementation selecting the right intervention for the right person or community at the right time.

There are 3 integrating domains: Data Science, Analytic Sciences, and Implementation Science and Interventions (including health promotion), to develop a Conceptual Framework for the Workshop and for this Review (Figure 2). The Framework identifies recent advances

and opportunities in each of the 3 domains. Data Science provides a better understanding of the structure, function, meaning, and utilization of data and new data types such as SDoH and measures of diversity and disparities. Quantitative and qualitative methods offer many new methods to create tools to identify and quantify disease and risk.⁹ Implementation Science provides opportunities to define the feasibility, efficacy, and cost-effectiveness of new interventions, including their ability to promote health and reduce disparities in CVD risk factors in individuals and communities.¹⁰ A crosscutting theme is the education and training of the research and practice workforce to develop and apply advances in precision prevention. Opportunities likely remain in the interactions among domains, due to many investigators currently performing research limited to a single domain. New types of data and precision analytics should allow the creation of new intervention tools for implementation scientists, or the identification of the need for new data types, as opportunities for research expansion. The convening of experts in these 3 domains was an important goal of this workshop.

An example of precision prevention that was discussed at the workshop is Precision Nutrition.¹¹ Large numbers of deaths from chronic diseases have dietary origins, and similarly large proportions of U.S. children and adults are exposed to and affected by poor nutrition with accompanying persistent or increasing nutritional and health disparities.^{11,12} At the same time, there is strong scientific evidence for current dietary recommendations and rich emerging science on new dietary approaches. This provides opportunities for data collection, integration, and interpretation for the development not only of personalized recommendations. This evidence also informs broad public health recommendations addressing environmental determinants of poor nutrition such as accessibility and affordability of healthy diets, and social and cultural preferences to increase food and nutrition security in order to reduce diet-related diseases in individuals, families, communities, and nations.¹³

This convergence of sciences (data, analytics, and intervention implementation) also provides challenges and opportunities to address CVD disparities known to exist in individuals and communities in the United States and worldwide. These advances must be viewed and implemented through a “Health Equity Lens” that recognizes that access to knowledge and information is not enough: health-related behaviors and risk factors are influenced by a complex, larger network of social, cultural, neighborhood, economic, and policy forces. People, providers, and communities need information that they can understand and use, and that is culturally and linguistically appropriate.^{14,15} Reducing health disparities requires questions to be asked when data science and implementation science are considered in creating a precision prevention program: *“Who is represented in the data sets? What, when, and for how long are data captured in data sets? How might differences in data quality contribute to bias? Which assumptions are used in modeling and harmonizing data sets and how free are they of implicit and explicit bias? How do findings inform clinical and public health strategies? Who is reached and served and for what outcome?”* To help address these important questions, this review provides opportunities to advance precision prevention and crosscutting themes of diversity, equity, inclusion, and accessibility. Below, we discuss the 4 opportunities: Opportunities in Data Science to Advance Precision Prevention; Quantitative and Qualitative Analytics; Development, Validation, Implementation, and Application of

Data Science in Precision Prevention Interventions (including health promotion); and Education and Training of Investigators and Consumers in Precision Prevention Research.

OPPORTUNITIES IN DATA SCIENCE TO ADVANCE PRECISION PREVENTION

Recent developments in data sciences have made available new types and sources of data, improved data quality, and generated larger sample sizes inclusive of more diverse individuals and communities to better describe the exposome.¹⁶ At the same time, these new types and sources of data often have increasing dissimilarities, potential for bias, and lack of comparability and compatibility. In contrast to the genome, a measure of inherited and unchanging DNA variation, the exposome may be defined as the sum of exposures experienced by an individual over a life-time. A key objective of precision prevention is to understand and quantify how these exposures relate to health. The data types with the potential to describe an individual's or community's exposome might be classified as a combination of environmental, sociocultural, psychosocial, behavioral, and biological exposures, with subclassifications of these broad categories¹⁷ (Figure 3).

An accurate and detailed measurement of the characteristics of places (eg, where a person lives, works, shops, and plays) can contribute greatly to an understanding of the exposome. Place provides a critical context that predicts and can independently impact health beyond individual-level determinants. Concerning health disparities, community characteristics describe physical, social, economic, and policy environments and their many intermediaries affecting health (Figure 4). The understanding of gene-environment, epigenetic-environment, microbiome-environment, and other environmental interactions and effects is at an early stage but should be considered when constructing models of determinants of and intervention effects on health outcomes. Exposome measures can serve as independent risk factors over and above the individual-level factors. Descriptors of the exposome can be constructed from data on residents or characteristics related to the place of interest, such as access to health care, healthy or unhealthy food sources, use of land including greenspace, crime, and safety, redlining and other measures of structural racism, and/or levels of air pollution. These data may be available from personal monitors, administrative sources, surveys, or other observations of the place of interest.

Many factors (eg, environmental, sociocultural, behavioral) are external to the individual's unique biology. To be impactful, these exposome measures must meet the FAIR criteria: Findable, Accessible, Interpretable, and Reusable.¹⁸ Cloud computing platforms such as the NHLBI BioData Catalyst¹⁹ accelerates the process of finding and accessing data and enable interrogating large-scale data resources in extensive collaborations.

The interoperability of these data requires the harmonization of data syntax (acceptable data structure and quality standards) and data semantics (understanding of the meaning of the data). The consistency and comparability of each data type allow standardization for their inclusion in local or centralized databases referent to diverse communities. Multidomain data harmonization has been demonstrated to be practical and scalable²⁰ when a single "common data model" is chosen as the harmonization target, such as the OHDSI OMOP

data model.²¹ The many applications include the linkage and layering of data and data tools to assess health and risk for disease within and between individuals and communities over time.

A successful example of the amalgamation of data linked to place and time to quantify the social exposome is the Area Deprivation Index (ADI).²² This social exposome metric uses 17 measures of the SDoH available to rank exposures of neighborhood disadvantage across a small discrete geography (ie, a census block-group) in a single time period. The ADI is harmonized for the full U.S., allowing comparisons of exposures between localities or regions. This assignment of an exposome measure linked to an individual's residence in a specific area at a specific time has been associated with physiologic and pre-clinical measures, health care costs, and disease risk.²³ The Neighborhood Atlas²² which displays differences in the exposome is an example of data democratization, allowing groups of persons defined by place and time to identify and compare their risks.²⁴ The value of such comparisons to assess health equity is many fold,¹⁴ with 1 practical application using it to adjust for economic challenges affecting access to high quality health care.²⁵

As illustrated by the Neighborhood Atlas and many other studies, place and time are critically important data for the linkage and layering of components and levels of the exposome. Spatial information has long been used to report data, especially in the U.S., at the county or 5-digit zip code levels. However, these geopolitical or mail delivery areas are not standard in size and are often quite large and heterogenous. More geographically precise measures of social exposome provide better accuracy, prediction of health outcomes, and overall utility for real-world resource targeting and applications. These more geoprecise measures may require additional linkage techniques including geocoding and 9-digit zip code cross-referencing. The precision of social exposure measurements may also be improved by collection of geolocation data related to where people actually spend their time (rather than their home address). The residents' perception of the quality of their place of residence may also assist in the assessment of its contribution to their health. Increasingly, a wide variety of tools including software for AI or Machine Learning are available to link addresses or 9-digit zip codes with health data (eg, Electronic Health Records), SDoH (including race/ethnicity, socioeconomic status, and population density, transportation), and access to health services (eg, rurality, provider density, etc). Likewise, many environmental exposomes are now available by geographic location (eg, air pollution [PM 2.5], agricultural exposures/pesticides, water pollution). Similar to the ADI, these geo-linked measures cannot only characterize the community, but also be used as an exposome measure at the individual level. Geospatial data should be scrutinized for biases in data selection, missing data, or data collection locations favoring well-resourced communities or population subgroups.

The workshop identified several research opportunities. 1) Considerable data remain in forms limiting their inclusion/analysis in large, inclusive datasets, or not viable for characterization of the exposome or definition of health outcomes. The risks in analyzing heterogeneous or unharmonized data such as clinical and genomic data should be better understood. Synergies between federated and centralized data repositories should be explored. 2) Current data science efforts to convert unstructured scientific, clinical, and sociocultural data to structural knowledge might use natural language processing, machine

learning, and other methods.⁹ The development of these data science methods must meet the syntactic and semantic harmonization standards to allow their applications to both individuals and communities.¹⁸ 3) The influences of SDoH and demographic factors on data collection methods and systems need to be examined to assure health equity in precision prevention research and interventions (see Development, validation, implementation, and application of data science in precision prevention interventions section).

QUANTITATIVE AND QUALITATIVE ANALYTICS: OPPORTUNITIES FOR PRECISION PREVENTION

The rapid improvement in data sources, ie, data quality, new types of data, the inclusion of larger and more diverse populations, and linkage in space and time, has in turn created opportunities for new precision analytic tools (Figure 2). In parallel with these advances in data sources, precision analytics have involved quantitative and qualitative methods to develop diagnostic and prognostic tools.⁵ Large databases, such as Electronic Health Records, have been analyzed to objectively identify phenotypes (risk factors, case definitions) in many persons receiving healthcare services.

The availability of multiple longitudinal cohorts with similar baseline and outcome data have been especially strong contributors to the development of predictive models and risk scores in both individuals and communities. AI and Machine Learning have been used to improve risk prediction, for causal analysis to identify potentially intervenable determinants of risk, and for phenotyping for broad/heterogeneous and poorly understood diagnoses.^{9,10} Many other mathematical methods can be applied to large datasets with exposome and outcome data. One example of potential importance in disparity research would be mediation analysis, in which racial, ethnic, and geographic disparities could be analyzed as to which and how much specific factors account for the disparity, thereby defining and prioritizing targets for possible intervention. Some datasets may benefit from a systems science approach to its analysis. One product of these analyses would be tools which could be incorporated into interventions applicable to individuals and populations, and tested for efficacy, safety, and cost in implementation research studies.

With the availability of raw population-level data, orders of magnitude larger and more diverse than classic case-control or cohort research studies, various computer-driven statistical/machine learning methods can be applied to identify individuals with phenotypes. These computerized phenotypes can then be used either as risk factors/risk markers for a disease or for the definition of cases of disease (Figure 5).²⁶ The larger size of the populations potentially accessed allows subgrouping of the population by age, sex, race, and other diversity characteristics as well as clusters of phenotypes that could be used with sufficient statistical power for precision prevention and meaningful risk comparisons. An important quality control feature of computerized phenotyping is identifying selection bias in subgroups, which could affect their inclusion as a case or an at-risk individual. Some phenotypes require sensitivity and specificity analysis against known cases and noncases before inclusion in larger databases.

Predictive modeling has been another application for multiple kinds of data. Instead of using cross-sectional samples, these analyses have used data from longitudinal cohorts to create risk scores in individuals and in communities. The best-known example of a risk score is the Framingham Risk Equation from the NHLBI-supported Framingham Heart Study, pioneering the application of multiple regression analyses to include demographic (eg, age, sex), physical measurements (systolic blood pressure), laboratory measurements (total cholesterol, HDL cholesterol), and behavioral factors (cigarette smoking).^{27,28} A limitation of the Framingham Heart Study is its lack of a racially and ethnically diverse sample. The Framingham risk equations have been refined by adding data from several other cohort studies, creating the ACC/AHA Pooled Cohort Equations.^{29,30} These epidemiologic tools have been incorporated with clinical data such as imaging coronary artery calcium and with genomic data to provide a more comprehensive modeling tool for the assessment of CHD risk.³¹

A more recent variation of a predictive risk score illustrating the combination of data from diverse sources predicts health rather than disease: the CVH Score (AHA Life's Essential 8).^{32,33} In addition to its primary outcomes of the absence of CVD, the CVH Score accessed data from 3 prospective cohort studies (recruited 1976–1989) and the 2005 to 2016 National Health and Nutrition Education Surveys, which provided CVH metrics measured with the same questions. The risk score uses demographic and 8 CVH measures: physical measurements (blood pressure, body mass index), laboratory data (hemoglobin A1C, blood lipids), and health behaviors (nicotine exposure, physical activity, diet, and sleep) to create an overall CVH score. This model performs well in internal validation, but less well in external validation in different study populations using differing methods for measurement of CVH. Some retraining and validation are needed for specific research or clinical uses, but generally it has value in estimating overall CVH risk.³⁴

Such predictive risk equations using aggregation of individuals' data might also be used to assess the risk of communities, especially after customization to take into consideration various aspects of diversity (eg, rural vs suburban vs urban). Community-wide data such as level of SDoH or ADI, could be added to risk prediction equations calculating an individual's risk score. This collective characterization and quantification of digital data can be used to assess the structure, function, and dynamics of an individual or population (Figure 6).³⁵ The customization of tools for diverse communities is of importance for community engagement. This may be facilitated by diversity in the team and its leadership who are designing the tools, as well as a community advisory board consisting of stakeholders. This advisory board may identify additional data sources that may be used to provide local adjustments of the tool, with an intentional focus on equity with attention paid to avoiding possible biases in the use of tools. The advisory board may also assess the level of technical sophistication required, allowing it to be adjusted to meet individual users' needs.

AI has been used to analyze large, heterogeneous datasets for application to a wide range of research questions. Implementing AI tools requires developing trust in AI solutions (Figure 7). Potential stakeholders would benefit from impartial, transparent, accountable, private/secure, and reliable AI.¹⁰ Validation, regulatory approval, and stakeholder acceptance of AI technology will be necessary before its implementation in clinical or community settings.¹⁰

Several recently developed mathematical applications show great promise for contributions to precision prevention. Mediation analysis, which is a subset of causal methods can be applied to identify and quantify the extent to which a single factor may be causally related to disease outcomes.^{36,37} Mediation analysis should clarify strategies to reduce confounding, the need and use of interaction terms, and assumptions of the analysis' final model. Of particular interest in precision prevention is the description, rank-ordering of the relative strengths, and independence of individual factors in a multivariate risk model when applied to the disparity. The risk associated with race, ethnicity or place of residence can be broken down into the proportions mediated by potential intermediating factors. Individual SDoH, behaviors, comorbidities, and geospatial factors identified as causal mechanisms contributing substantially to disease risk could then be prioritized for their inclusion in individual or community-wide interventions.³⁸ In addition, tools are needed to understand intervention effect heterogeneity and optimal assignment of interventions based on an individual or community characteristics. AI methods may be used for causal and mediation analyses of intervention effects and to identify the "right intervention for the right person at the right time." Such methods could identify intervention effect heterogeneity and optimal assignment of prevention interventions in response to different forms of risk prediction.

DEVELOPMENT, VALIDATION, IMPLEMENTATION, AND APPLICATION OF DATA SCIENCE IN PRECISION PREVENTION INTERVENTIONS

Implementation Science, the study of methods to promote the adoption and integration of evidence-based practices, interventions, and policies into routine health care and communities, and its convergence with data science is long overdue. Importantly, all implementation science requires a careful assessment of the community's research priorities, needs, and organizational readiness to change. The thorough identification of community needs, prioritization, and willingness to engage using the principles of community-based participatory research is an important way to assure that prevention intervention implementation will be successful. It is also useful to address the targeted health issue as part of the intervention generation to identify barriers and social factors affecting adoption, adherence, sustainability, and satisfaction with the proposed intervention.

Precision prevention tools, including computerized phenotypes, risk equations/scores, AI output, and mediation analyses are the products of quantitative and qualitative analyses to optimize their functionality in the context of the data used to develop them. Yet, each tool requires further validation as feasible/acceptable and safe/efficacious when applied to real-world settings, including engagement of diverse individuals and communities (Figure 2). Implementation science examines how to intervene to implement best practices.³⁹ Implementation typically targets one or more levels at which the best practices will be used, including individuals (patients or family members, or health care providers) or groups (clinics, hospitals/health systems). Community and/or societal factors, whose presence is suggested by variable implementation rates of the same guidelines, should be better understood as influences on guideline implementation across communities.³⁹ Implementation research may vary by the tool or intervention being assessed, the characteristics of the implementation team, and the nature of the implementation process

itself (Figure 8). The measures of implementation are usually the behavior change of the targeted group to elicit more healthy outcomes. It is emphasized that the target is not the condition or outcomes of the diseases under care, as already established as best practices by clinical trials. Rather, the target of analysis in implementation research is often the practices rather than the health outcomes, although hybrid designs are often used to study the two in tandem.⁴⁰ Various measures of efficacy might be collected, including patient- or community-level behavioral measures of adherence/compliance, or process measures of care provision and outcomes of disease risk. The cost of the implementation, often related to its feasibility, might also be considered, and can also be an important product of an implementation research study.

The NIH Stage Model for Behavior Intervention Development describes the process of intervention generation and validation for implementing prevention interventions (Figure 9). Stage 0 consists of the basic science basis for the intervention, best described in a conceptual model of molecular pathways or mechanisms involved in health/wellbeing and disease/risk. For example, the personalization of a cardiovascular risk reduction plan might involve consideration of genetic, pharmacologic, and lifestyle behavior medicine (Figure 10). In the model, the efficacy of behavioral intervention may be influenced by the interaction with genetic predispositions, pharmacologic options, or the interaction of the two (Figure 10). At the community level, environment and SDoH may provide a similar model to interact with healthy lifestyles. This process involves mapping out behavioral, social, socioeconomic, and environmental exposomes. An assessment of community or organizational readiness in tandem with community members to address the targeted health issue is part of the intervention generation to identify barriers and social factors affecting adoption, adherence, sustainability, and satisfaction with the proposed intervention. The next stages at the individual or community levels would then require some testing of effectiveness, usually in clinical or community research settings. The intervention then would need individual or community engagement, through qualitative research in partnership with diverse community or organizational representatives or subpopulations of patients whose participation will be crucial to tailor the intervention for lifelong prevention intervention implementation.

Implementation studies have several important considerations, including to whom the intervention is being delivered (single participants, multiple participants individually, groups of participants, or organizations). The intervention could be delivered electronically or face-to-face, by a single agent (provider, therapist, etc), by a team of agents, or by different agents over time. Some of the questions to consider include: *Will participants be individually assigned to study arms, or in a group (such as communities)? Will the intervention be delivered independently or nonindependently to participants?* The responses to these questions may favor 1 study design over the other (Table 1).

The first group of study designs in Table 1 entails individual assignment and independent delivery. Individual assignment requires random or nonrandom assignment of individual participants to study arms. Independent delivery requires that participants receive the intervention individually rather than in a group format; it also requires that participants do not interact with other participants or with a shared intervention agent. The best-known design that includes these features is the randomized clinical trials (RCT) as in drug

trials where participants receive their medication individually, do not interact with other participants, and do not interact with a shared intervention agent.

The Multiphase Optimization Strategy studies are useful for building an intervention that has these features with several phases in which individual intervention components are identified. A conceptual model is developed for the mechanisms by which the intervention components are expected to affect the outcome.⁴¹ The components are tested individually and in combination to identify the most promising components, which are then incorporated into an optimized intervention to be tested with an RCT.

The Sequential, Multiple Assignment, Randomized Trials design uses adaptive interventions which has a sequence of prespecified rules and decision points to guide modifications of the intervention during the trial, usually with 2 or more stages of randomization at each decision point.⁴² This design is useful to identify a sequence of steps leading to the optimal outcome.

The Regression Discontinuity Design is helpful where randomization is not possible. It assigns participants according to a cut-off in a quantitative score and models the relationship between the assignment variable and the outcome.⁴³ If analyzed correctly, it can provide evidence as strong as from an RCT.⁴⁴

N-of-1 Trials randomize 2 or more interventions to single participants in a multiple cross-over design.⁴⁵ The design can be used to compare several interventions to identify that individual's optimal intervention. All these designs involve independent assignment of participants and independent delivery of the intervention. When independent assignment of participants and independent delivery of the intervention is not possible, the parallel group or cluster-randomized trial design is optimal.⁴⁶ Groups or clusters such as communities, worksites, schools, or clinics, are randomized to study arms. Observations in members of the same group or clusters will not be independent and will have a positive intraclass correlation. This should be considered during planning to adjust power and sample size.⁴⁶

The Individually Randomized Group-Treatment Trials randomize participants to interventions but entail interaction among participants or with a shared intervention agent post randomization.⁴⁶ Behaviors observed among members who interact with each other, or a shared intervention agent will not be independent and that needs to be considered during planning. Many study designs involving individuals (Multiphase Optimization Strategy, SMART, Regression Discontinuity Design) can be adapted to the assignment of groups or clusters.^{40,47}

Finally, the Stepped Wedge Group Cluster-Randomized Trial starts with groups in the control condition with cross-over to the intervention in random order and stepped schedule, with all groups receiving the intervention before the end of the study. Observations from members of the same group will not be independent.⁴⁶ Because of several difficulties with this design, it is recommended to be used only if a parallel group-randomized trial is not possible.

For implementation in widely diverse groups, it is important to confirm the generalizability and heterogeneity of intervention and implementation effectiveness through

community engagement. Prior analyses of the implementation of cardiovascular guidelines acknowledged that each intervention at the patient, provider, or clinic level is likely to be implemented in a unique setting which may impact its effectiveness even if implemented identically across settings.³⁹ This requirement for diversity/equity/inclusion to demonstrate feasibility and efficacy may mitigate disparities in the later dissemination of guidelines. The replication of similar interventions in diverse settings would provide greater assurance that the benefits of the intervention are not biased or uneven, depending on who receives them.

The poor performance of a tool may elicit a search for additional data sources to improve the efficacy of the tool. An important question includes: *Can precision prevention reduce disparities in cardiovascular risk in real world settings?* Study design, data source selection, subject recruitment and participation, data collection, and statistical analysis should address critical questions in the planning stages. *Who is represented in the data set to be used? What data are captured and for how long? What are the assumptions of the models and do they minimize bias? Who are reached and served, and for what outcome? How do findings inform clinical and public health practice and prevention?* For individuals and communities, large and ongoing social inequities, as well as discrimination, probably contribute to the largest health disparities.

The uneven implementation and efficacy of preventive strategies and technologies likely also contribute. Data science has the potential to link patients more closely to health care providers via telemedicine and data portals, improving health behaviors via messaging, electronic coaching, and “nudges” which counteract implicit bias. Unique risk factors or a greater attributable risk of known factors in disadvantaged groups might be identified by AI or mediation analysis.

Can information technologies such as clinical decision support play a role in health disparity reduction by facilitating the implementation of evidence-based practices and policies in underserved individuals and communities?⁴⁸ Likely the tools most widely implemented and validated in clinical care are the predictive models and risk scores, including those which incorporate clinical, imaging, and genomic data.^{30,31} An example of disparity mitigation in hypertension control is persistent hypertension identification from electronic health records.⁴⁹ Another example of precision prevention addressing health disparities is the Sustainable East African Research in Community Health Program (SEARCH),^{50,51} which addressed the issue that the people most in need of cardiovascular health services (eg hypertension detection and management) in rural East Africa were not engaging in clinic-based services as a one-size-fits-all model. The SEARCH approach expanded health campaigns and health worker services for HIV and other diseases to include hypertension and demonstrated a 21% reduction in population-level hypertension mortality in a cluster randomized trial.^{50,51} SEARCH is now evaluating a data-driven machine learning approach to tailoring hypertension services for individual needs leveraging community health workers. Several other analytic tools, notably the ADI and Natural Language Processing (a form of AI), are examples of clinical applications with promise to reduce health disparities, but require validation of efficacy and cost effectiveness.

EDUCATION AND TRAINING OF INVESTIGATORS AND CONSUMERS IN PRECISION PREVENTION RESEARCH

An underlying premise of this Workshop is that precision prevention as a field must continue its evolution from observations/predictions to intervention and implementation to reduce the burden of diseases and related disparities. The Workshop's Conceptual Model implies that expanding quality, quantity, and types of data along with the rigorous development of quantitative and qualitative analytical methods will set the stage for new investigators to construct and validate interventions that they can implement in real-world settings (Figure 2).

The convergence of data science and implementation science should be included in the training of researchers and clinicians. The cross-cutting theme of Education and Training identifies at least 2 workforce requirements for the application of data implementation sciences, including health promotion interventions, to occur. First, there needs to be integration of the 3 domains (Figure 2) leading to the expansion of a research workforce producing novel and evidence-based methods easily utilized by clinicians, community leaders, and policy makers. The second need is the education and training of the clinicians, community leaders, and policy makers as consumers who understand the utility and potential impact of precision prevention interventions as applied to their patients and communities.

The strong demand for all types of data scientists has persisted for many years, with recognized workforce shortages in biostatisticians, informaticists, epidemiologists, computer scientists, applied mathematicians, interventionists, and those with clinical trial expertise. In addition, for those disciplines traditionally not intersecting with data science such as applied social scientists and implementation scientists, data science training will provide immense opportunities for precision prevention interventions supported with data science tools.

For all these disciplines, the current didactic education and mentored research training in each discipline may not include the knowledge, skills, and attitudes which would equip the trainee to contribute to the implementation of precision prevention science as part of their research and/or clinical career. An optimistic trend is the establishment of implementation science research units, certificates in clinical trials, and graduate degree programs. An important question is the extent to which these programs have didactic courses and mentored research in precision prevention science or can provide the level of quantitative or data science training needed to support a precision prevention paradigm.

Furthermore, addressing health disparities in all training efforts is crucial to ensure that interventions reach individuals and populations at greatest need. While it is crucial to address the broad SDoH factors (eg, poverty, racism), AI tools such as Large Language Models may help to identify and address the specific needs and goals of individuals and communities. Although AI tools may help determine the right treatment at the right time for individuals, it should be employed with vigilance to avoiding unintended harms for vulnerable populations. Additional tailoring of interventions to individuals and population groups with greater needs, may improve outcomes and reduce disparities.

Given that engagement of community stakeholders can greatly increase the likelihood of developing and implementing interventions that meet the needs of communities, training on how to engage appropriate stakeholders in designing interventions, interpreting findings, and disseminating to larger population groups will be important. While some principles should be included across training for different groups, other aspects of training will need to differ in accordance with who is getting trained. For example, training programs for researchers interested in engaging in precision prevention research will need to provide skill development in areas such as multilevel observational research, intervention research, and appropriate analyses and implementation study designs. Considering the pace of this area of study, a specific focus on training that is adaptive and incorporates fundamental skills as well as insights from areas which are rapidly evolving will be important.

Training of clinicians interested in precision prevention should address SDoH, implementation science, clinical trial designs, biostatistics, and grant writing. Training of community stakeholders and policymakers should focus on democratizing precision prevention, clarifying the meaning of precision prevention vs precision medicine, and gaining community trust. Training needs in precision prevention will become clearer as this area develops further to improve the health of the public, and particularly those from historically marginalized backgrounds and those whose needs have not been adequately met, being the driving force.

The opportunities afforded by precision prevention training programs are considerable. This training emphasizes the benefits of multilevel approaches to understand broad SDoH and individual, genetic, and behavioral factors. Training for both observational (prediction) and interventional (disease prevention and health promotion) studies could be obtained. The inclusion of community engagement in training the workforce would favor more holistic research inclusive of all community segments, rather than those with greater resources and fewer needs, and recognize the heterogeneity and inequities within populations. The training should allow learning about and implementing novel analytic methods, and effective study designs, especially if tailored to individual trainees' research projects. Such study designs would involve collaborations with investigators and trainees from diverse sociocultural backgrounds and a broader range of disciplines and methodologies. This newly acquired knowledge and skills should contribute to career development and opportunities for external funding, including NIH and nongovernmental sources.

Training in emerging and cross-disciplinary sciences also brings challenges. The discipline of precision prevention needs to have a consensus on what it entails, as attempted in Figure 2. The tension between individual vs community approaches is a healthy one, but broadens the knowledge and skills required. Privacy and other ethical issues should also be considered as the precision of data increases. There is a need for continuing education of established investigators in novel study designs and data analytic methods. The uncertainties of funding for precision prevention persist, requiring considerable development time and effort before the funding and initiation of study protocols. The costs of precision prevention knowledge and skill development for investigators and practitioners has not been explored, but the didactic and experiential learning of health care providers should be considered in light of the healthcare costs potentially reduced by mitigation of health disparities.

Knowledge, skills, and abilities required for investigators and practitioners will be necessary to develop curricula and training programs in precision prevention. The 3 domains proposed in Figure 2 recommend the inclusion of data science, population sciences (ie, epidemiology, demography, exposome measurements, qualitative tools, sociology with recognition of health inequities and SDoH), analytical sciences (ie, biostatistics, computer science, clinical trial methods, machine learning and AI, applied mathematics), implementation science (ie, study design, intervention development, engagement of participating individuals and communities, implementation evaluation, outcomes assessment), as well as Research Ethics/Protection of Human Subjects.

The coursework should emphasize the integration of domains to produce unique data sets, effective clinical and community-engagement tools, as well as culturally appropriate applications in real world settings. Additionally, engaging and training a diverse research workforce is paramount to moving precision prevention science forward as this will allow creative and diverse solutions to scientific challenges, foster scientific innovation, and enhance public trust necessary for working with vulnerable populations.

Another challenge is the education and training of healthcare providers, community leaders and change agents in applying precision prevention research. This group also includes policymakers who might provide legislative support for precision prevention interventions. The means to identify individuals and communities with elevated risks begins the precision process, ie, the identification of the right recipient for the right intervention at the right time.

The identification of data sources should include what each has to offer. The different approaches to prevention in individual vs communities should be understood. The basic tenets of diversity, inclusion, and equity should be operative. The tailoring of the preventive approaches should be emphasized for high-risk individuals and communities with different needs, cultures, and resources. The feasibility and effectiveness of precision prevention research studies should be relatable to other high-risk individuals or communities. Basic skills to successfully build partnerships with community leaders and members to promote community participation in research should be part of the training of all research team members. Sources of data, expertise, and financial resources should be identifiable in the local setting. Workshops and other forms of curricula should be tested and made accessible to stakeholders.

SYNTHESIS OF OPPORTUNITIES FOR PRECISION PREVENTION

The Central Illustration depicts the various components of precision prevention for Precision Health, relevant to both Precision Medicine in individuals and Precision Public Health in communities.¹ Crosscutting themes identified in this Workshop include processes to integrate data and implementation sciences, and interventions to address health disparities.

The Workshop also addressed gaps in data sources, quantitative and qualitative analysis methods, validated interventions, and study methods and design for studies on intervention feasibility and effectiveness. A list of research training opportunities were also identified in the Workshop.⁴

Finally, what would an ambitious, high-risk/high reward precision prevention research program (aka Moonshot Trial) look like? Given the considerations of this workshop, the conditions most worthy of deliberation would be those with poor outcomes and those that are amenable to prevention. While the underlying disease burden may be a factor, an important characteristic would be demonstrated disparities in subpopulations and/or communities. Also for consideration are the 3 domains from the precision prevention Conceptual Model (Figure 2). Research questions for “moonshot trials” would include the following: *Can data be accessed? What qualitative and/or quantitative analytic methods can be used to identify unique risk predictors or outcomes? Can AI methods be used in novel databases to create better measures of risk or outcomes? Can a precise intervention be developed? Can implementation science methods be used to demonstrate the feasibility of the intervention and its effectiveness in reducing the burden and disparity of the condition?* The trial components would likely begin with identification and participation of a population, or clusters of communities at increased risk for the target CVD condition.

The study population should have an increased CVD risk (eg, certain dietary, other adverse lifestyle, epigenetic, social determinant, and/or other exposome factors). Additionally, analytic methods could be used to help identify communities or individuals at high risk. Analytic methods may be available or developed with or without validation to assess risk or change in risk over time, for application as a decision-making tool to build an intervention in the high-risk individual or community, particularly those with health disparities.

The implementation science study would have several goals: implementation of the intervention inclusive of most individuals or communities, improvement of measures of intermediary risk reduction over the short term, reduction of CVD outcomes in the individuals and/or communities participating in the intervention and demonstrated reduction or elimination of the disparities of CVD outcomes as described before the intervention. The specific study design will vary widely depending on the research questions, condition, individual vs community-level participation, the intervention selected for testing, and the outcomes measured. Studies with the same design and similar participants performed in different settings (eg, clinical vs community, different target populations, with various SDoH, etc) might be necessary to fully understand the generalizability, feasibility, and effectiveness of the intervention on the condition.

CONCLUSIONS

This report of a Workshop⁴ entitled: *“The Science of precision prevention: research Opportunities and Clinical Applications to Reduce Disparities in Cardiovascular Health”* reviews the state of the science and best practices for Precision Prevention Research. Investigators with diverse expertise in biological, clinical, quantitative, and public health sciences identified expanding opportunities to describe better the exposome in large populations in a manner that recognizes and advances diversity, equity, and inclusiveness.

The wealth of new data that are increasingly becoming available can be used in qualitative and quantitative analyses to create new estimates of individual or community risk and other tools that support clinical decisions, community policies, or other interventions.

These data and tools must advance the state-of-the-art beyond prediction to prevention, in which interventions can demonstrate changes in individuals' or communities' behaviors and burdens of disease. Implementation science methods and clinical trials are needed to validate that prevention goals and outcomes can be attained, and that they can be implemented in diverse groups and lead to sustainable improvement in health and reductions in health disparities. A larger, well-trained, and diverse scientific workforce with knowledge and skills to expand precision prevention research is needed. Likewise, practitioners in clinical medicine and in public health require knowledge and skills to co-create and apply appropriate, effective precision prevention practices at the right time in the right individuals, families, and communities (Figure 1).

ACKNOWLEDGEMENTS

The authors would like to thank their Precision Prevention Workshop Planning Committee members (Gabriel Anaya, MD, MSc; Vanessa Barnes, BS; Alison Brown, PhD, MS; Dave Clark, DrPH; Sam Feudo, BS; Laurie Friedman Donze, PhD; Bramaramba Kowtha, MS, RDN; Nicole Redmond, MD, PhD, MPH; Eric Shiroma, ScD; Deborah Young-Hyman, PhD) and the NHLBI Workshop Development Support staff.

FUNDING SUPPORT AND AUTHOR DISCLOSURES

The views expressed are those of the authors and do not necessarily reflect the views of the National Heart, Lung, and Blood Institute; the National Institutes of Health; or the U.S. Department of Health and Human Services. The authors have reported that they have no relationships relevant to the contents of this paper to disclose.

ABBREVIATIONS AND ACRONYMS

| | |
|---------------|---|
| ADI | Area Deprivation Index |
| AI | artificial intelligence |
| CVD | cardiovascular disease |
| NHLBI | National Heart, Lung, and Blood Institute |
| RCT | randomized controlled trials |
| SDoH | Social Determinants of Health |
| SEARCH | Sustainable East African Research in Community Health |

REFERENCES

1. Sidney S, Quesenberry CP Jr, Jaffe MG, Sorel M, Go AS, Rana JS. Heterogeneity in national U.S. mortality trends within heart disease subgroups, 2000–2015. *BMC Cardiovasc Disord.* 2017;17(1):192. [PubMed: 28720073]
2. Grant MG, Pratt C, Wong RP, et al. Implementing the National Heart, Lung, and Blood Institute's strategic vision in the Division of Cardiovascular Sciences-2022 update. *Circ Res.* 2022;131(8):713–724. [PubMed: 36173825]
3. Writing Group for the Division of Cardiovascular Sciences' Strategic Vision Implementation Plan, Goff DC Jr, Buxton DB, et al. Implementing the National Heart, Lung, and Blood Institute's strategic vision in the Division of Cardiovascular Sciences. *Circ Res.* 2019;124(4):491–497. [PubMed: 31031412]

4. NHLBI. The science of precision prevention to reduce disparities in cardiovascular health: NHLBI. 2022. Accessed December 5, 2022. <https://www.nhlbi.nih.gov/events/2022/science-precision-prevention-reduce-disparities-cardiovascular-health>
5. Pearson TA, Califf RM, Roper R, et al. Precision health analytics with predictive analytics and implementation research: JACC state-of-the-art review. *J Am Coll Cardiol*. 2020;76(3):306–320. [PubMed: 32674794]
6. Rebbeck TR. Precision prevention of cancer. *Cancer Epidemiol Biomarkers Prev*. 2014;23(12):2713–2715. [PubMed: 25362191]
7. Gillman MW, Hammond RA. Precision treatment and precision prevention: integrating “below and above the skin”. *JAMA Pediatr*. 2016;170(1):9–10. [PubMed: 26595371]
8. Biro K, Dombradi V, Jani A, Boruzs K, Gray M. Creating a common language: defining individualized, personalized and precision prevention in public health. *J Public Health*. 2018;40(4):e552–e559.
9. Sevakula RK, Au-Yeung WM, Singh JP, Heist EK, Isselbacher EM, Armoundas AA. State-of-the-art machine learning techniques aiming to improve patient outcomes pertaining to the cardiovascular system. *J Am Heart Assoc*. 2020;9(4):e013924. [PubMed: 32067584]
10. Bazoukis G, Hall J, Loscalzo J, Antman EM, Fuster V, Armoundas AA. The inclusion of augmented intelligence in medicine: a framework for successful implementation. *Cell Rep Med*. 2022;3(1):100485. [PubMed: 35106506]
11. Pratt CA, Brown AGM, Dixit S, et al. Perspectives: on precision nutrition research in heart, lung, and blood diseases and sleep disorders. *Adv Nutr*. 2022;13(5):1402–1414. [PubMed: 35561742]
12. Mozaffarian D. Dietary and policy priorities for cardiovascular disease, diabetes, and obesity: a comprehensive review. *Circulation*. 2016;133(2): 187–225. [PubMed: 26746178]
13. Berciano S, Figueiredo J, Brisbois TD, et al. Precision nutrition: maintaining scientific integrity while realizing market potential. *Front Nutr*. 2022;9:979665. [PubMed: 36118748]
14. Westfall JM, Mensah GA. T4 Translational moonshot: making cardiovascular discoveries work for everyone. *Circ Res*. 2018;122(2):210–212. [PubMed: 29348250]
15. CDC. Using a health equity lens: Centers for Disease Control and Prevention. 2022. Accessed December 5, 2022. https://www.cdc.gov/healthcommunication/Health_Equity_Lens.html
16. Vermeulen R, Schymanski EL, Barabasi AL, Miller GW. The exposome and health: where chemistry meets biology. *Science*. 2020;367(6476): 392–396. [PubMed: 31974245]
17. NIMHD. NIMHD research framework. 2018. Accessed December 5, 2022. <https://nimhd.nih.gov/researchFramework>
18. NCATS. The CTSA National Center for Data to Health (CD2H): NCATS. 2022. Accessed December 5, 2022. <https://cd2h.org/mission>
19. Ahalt S, Avillach P, Boyles R, et al. Building a collaborative cloud platform to accelerate heart, lung, blood, and sleep research. *J Am Med Inform Assoc*. 2023;30(7):1293–1300. [PubMed: 37192819]
20. Haendel MA, Chute CG, Bennett TD, et al. The National COVID Cohort Collaborative (N3C): rationale, design, infrastructure, and deployment. *J Am Med Inform Assoc*. 2021;28(3):427–443. [PubMed: 32805036]
21. Hripcsak G, Duke JD, Shah NH, et al. Observational Health Data Sciences and Informatics (OHDSI): opportunities for observational researchers. *Stud Health Technol Inform*. 2015;216: 574–578. [PubMed: 26262116]
22. Kind AJH, Buckingham WR. Making neighborhood-disadvantage metrics accessible - the neighborhood atlas. *N Engl J Med*. 2018;378(26): 2456–2458. [PubMed: 29949490]
23. Powell WR, Buckingham WR, Larson JL, et al. Association of neighborhood-level disadvantage with Alzheimer disease neuropathology. *JAMA Netw Open*. 2020;3(6):e207559. [PubMed: 32525547]
24. Woelfle M, Olliaro P, Todd MH. Open science is a research accelerator. *Nat Chem*. 2011;3(10):745–748. [PubMed: 21941234]
25. ACO. Realizing Equity, Access, and Community Health (REACH) model: Accountable Care Organization (ACO). 2022. Accessed December

- 5, 2022. <https://www.cms.gov/newsroom/fact-sheets/accountable-care-organization-aco-realizing-equity-access-and-community-health-reach-model>
26. Hu H, Liu X, Zheng Y, et al. Methodological challenges in spatial and contextual exposome-health studies. *Crit Rev Environ Sci Technol*. 2023;53(7):827–846. [PubMed: 37138645]
 27. FHS. Cardiovascular disease (10-year risk), recommended: 2018 ACSVD 10-year risk calculator. 2023. Accessed December 5, 2022. <https://www.framinghamheartstudy.org/fhs-risk-functions/cardiovascular-disease-10-year-risk/>
 28. D’Agostino RB Sr, Vasan RS, Pencina MJ, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. *Circulation*. 2008;117(6):743–753. [PubMed: 18212285]
 29. Goff DC Jr, Lloyd-Jones DM, Bennett G, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice guidelines. *J Am Coll Cardiol*. 2014;63(25 Pt B):2935–2959. [PubMed: 24239921]
 30. Karmali KN, Goff DC Jr, Ning H, Lloyd-Jones DM. A systematic examination of the 2013 ACC/AHA pooled cohort risk assessment tool for atherosclerotic cardiovascular disease. *J Am Coll Cardiol*. 2014;64(10):959–968. [PubMed: 25190228]
 31. Khan SS, Post WS, Guo X, et al. Coronary artery calcium score and polygenic risk score for the prediction of coronary heart disease events. *JAMA*. 2023;329(20):1768–1777. [PubMed: 37219552]
 32. Lloyd-Jones DM, Hong Y, Labarthe D, et al. Defining and setting national goals for cardiovascular health promotion and disease reduction: the American Heart Association’s strategic impact goal through 2020 and beyond. *Circulation*. 2010;121(4):586–613. [PubMed: 20089546]
 33. Lloyd-Jones DM, Allen NB, Anderson CAM, et al. Life’s essential 8: updating and enhancing the American Heart Association’s construct of cardiovascular health: a presidential advisory from the American Heart Association. *Circulation*. 2022;146(5):e18–e43. [PubMed: 35766027]
 34. Zheng Y, Huang T, Guasch-Ferre M, et al. Estimation of life’s essential 8 score with incomplete data of individual metrics. *medRxiv*. 2023. 10.1101/2023.03.03.23286786
 35. Smith RJ, Merchant RM. Harnessing the crowd to accelerate molecular medicine research. *Trends Mol Med*. 2015;21(7):403–405. [PubMed: 26141797]
 36. Lee H, Herbert RD, McAuley JH. Mediation analysis. *JAMA*. 2019;321(7):697–698. [PubMed: 30681700]
 37. Lee H, Cashin AG, Lamb SE, et al. A guideline for reporting mediation analyses of randomized trials and observational studies: the AGReMA statement. *JAMA*. 2021;326(11):1045–1056. [PubMed: 34546296]
 38. Kosorok MR, Moodie EE. *Adaptive Treatment Strategies in Practice: Planning Trials and Analyzing Data for Personalized Medicine*. SIAM; 2015.
 39. Chan WV, Pearson TA, Bennett GC, et al. ACC/AHA special report: clinical practice guideline implementation strategies: a summary of systematic reviews by the NHLBI Implementation Science Work Group: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *J Am Coll Cardiol*. 2017;69(8):1076–1092. [PubMed: 28132746]
 40. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. *Med Care*. 2012;50(3):217–226. [PubMed: 22310560]
 41. Collins LM. *Optimization of Behavioral, Biobehavioral, and Biomedical Interventions: The Multiphase Optimization Strategy (MOST)*. Springer; 2018.
 42. Nahum-Shan I, Almirall D. *An introduction to adaptive interventions and SMART designs in education*: National Center for Special Education Research (NCSE). 2019. Accessed December 5, 2022. <https://ies.ed.gov/ncser/pubs/2020001/pdf/2020001.pdf>
 43. Hilton Boon M, Craig P, Thomson H, Campbell M, Moore L. Regression discontinuity designs in health: a systematic review. *Epidemiology*. 2021;32(1):87–93. [PubMed: 33196561]
 44. Rubin DB. Assignment to treatment group on the basis of a covariate. *J Educ Stat*. 1977;2(1):1–26.

45. Davidson KW, Silverstein M, Cheung K, Paluch RA, Epstein LH. Experimental designs to optimize treatments for individuals: personalized N-of-1 trials. *JAMA Pediatr.* 2021;175(4):404–409. [PubMed: 33587109]
46. Murray DM, Taljaard M, Turner EL, George SM. Essential ingredients and innovations in the design and analysis of group-randomized trials. *Annu Rev Public Health.* 2020;41:1–19. [PubMed: 31869281]
47. Pennell ML, Hade EM, Murray DM, Rhoda DA. Cutoff designs for community-based intervention studies. *Stat Med.* 2011;30(15):1865–1882. [PubMed: 21500240]
48. Lu Y, Melnick ER, Krumholz HM. Clinical decision support in cardiovascular medicine. *BMJ.* 2022;377:e059818. [PubMed: 35613721]
49. Lu Y, Huang C, Mahajan S, et al. Leveraging the electronic health records for population health: a case study of patients with Markedly elevated blood pressure. *J Am Heart Assoc.* 2020;9(7):e015033. [PubMed: 32200730]
50. Havlir DV, Balzer LB, Charlebois ED, et al. HIV testing and treatment with the use of a community health approach in rural Africa. *N Engl J Med.* 2019;381(3):219–229. [PubMed: 31314966]
51. Hickey MD, Ayieko J, Owaraganise A, et al. Effect of a patient-centered hypertension delivery strategy on all-cause mortality: secondary analysis of SEARCH, a community-randomized trial in rural Kenya and Uganda. *PLoS Med.* 2021;18(9):e1003803.

HIGHLIGHTS

- Cardiovascular risk assessment and prediction needs to progress to risk intervention and reduction.
- Opportunities for precision prevention are enhanced by new data sources, new analytic methods, and new applications such as implementation science and health promotion.
- Precision prevention can develop and validate interventions tailored to individuals or communities at high risk.
- Cardiovascular precision prevention implements the right intervention in the right individual/community at the right time.

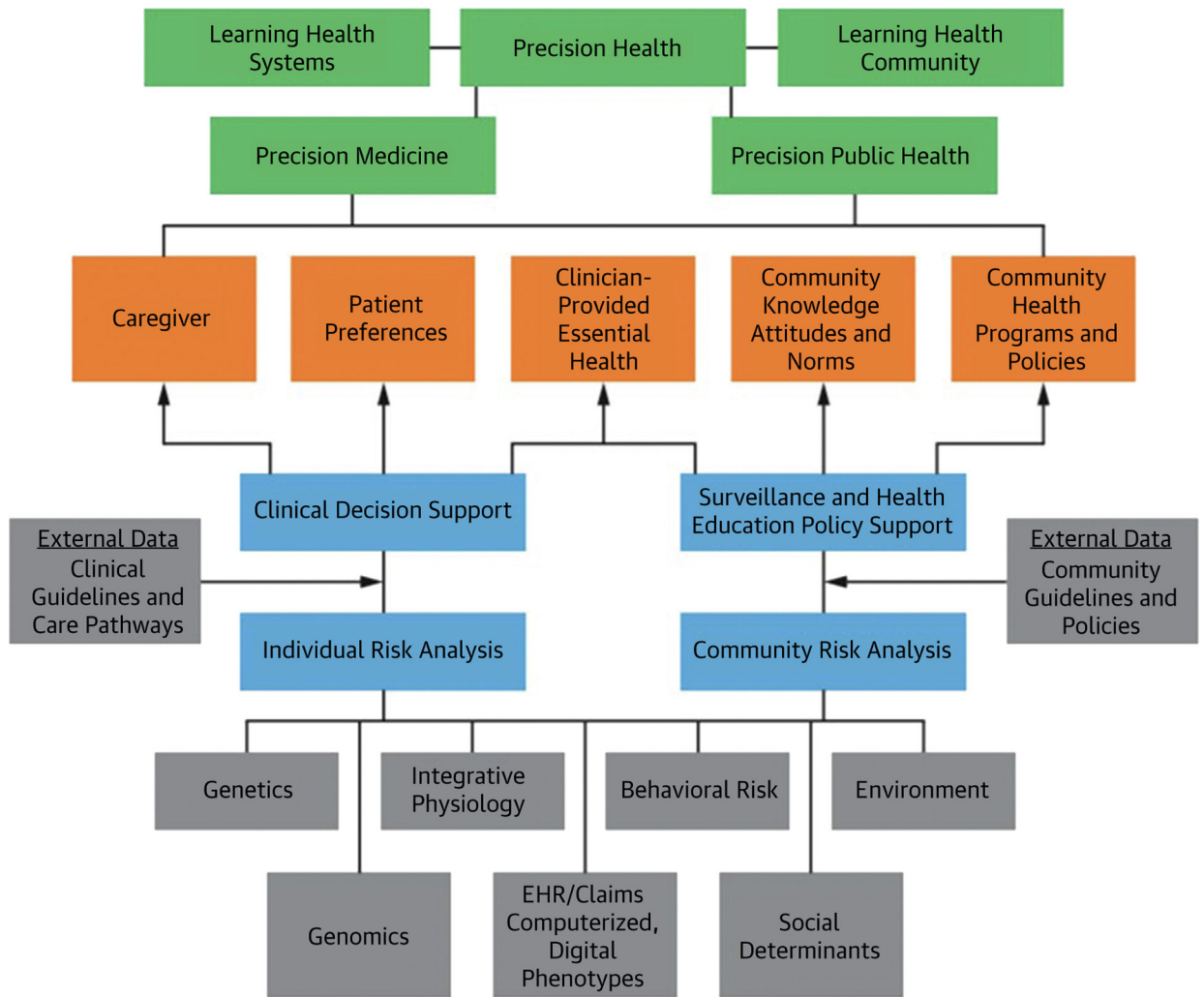


FIGURE 1. Precision Health as the Data-driven Merger of Precision Medicine and Precision Public Health

Conceptual framework to guide the precision prevention workshop

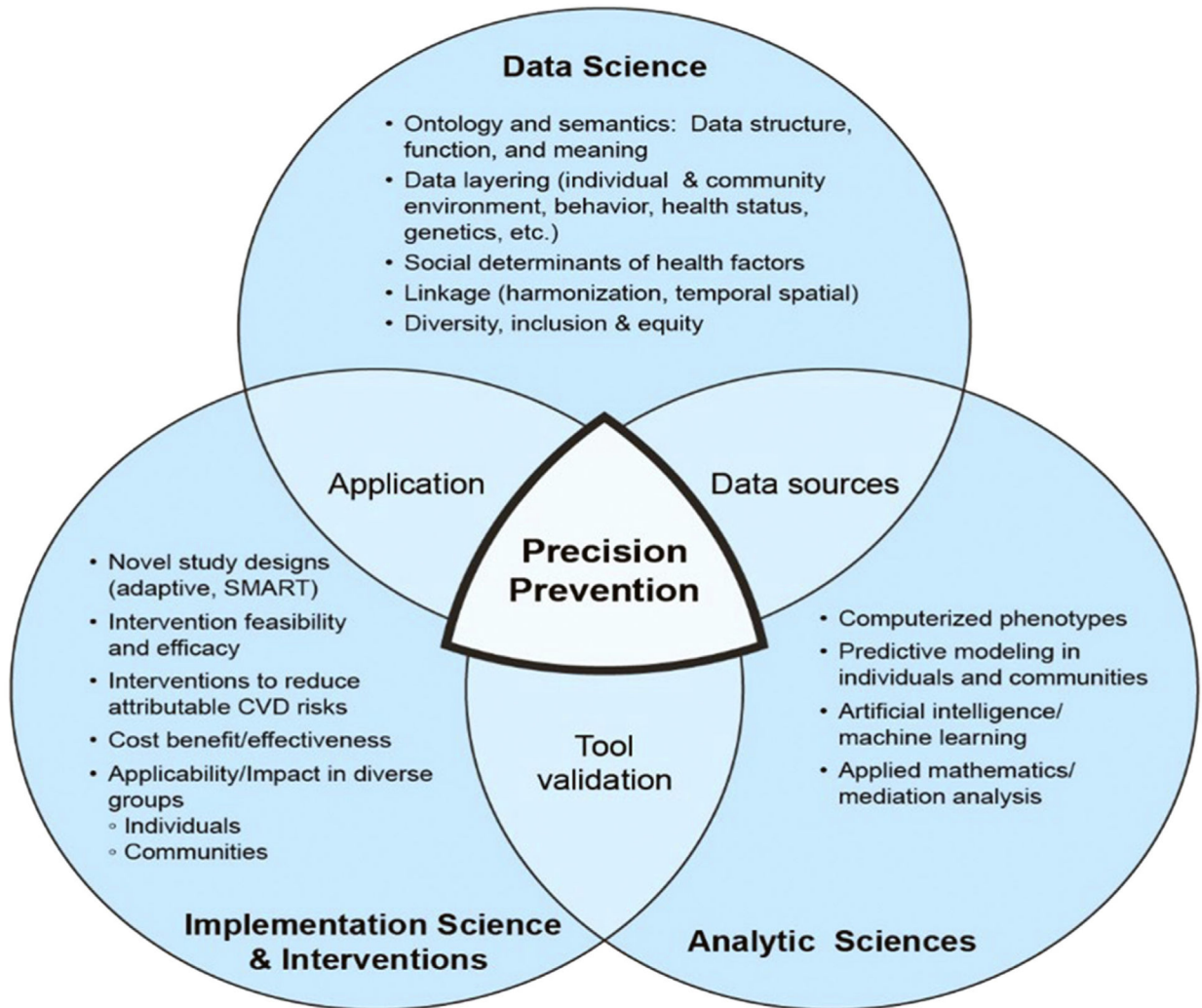


FIGURE 2. Intersection of 3-Dimensions as the Conceptual Framework for Precision Prevention





| | | Levels of Influence* | | | |
|--|----------------------------|---|---|---|--|
| | | Individual | Interpersonal | Community | Societal |
| Domains of Influence <i>(Over the Lifecourse)</i> | Biological | Biological Vulnerability and Mechanisms | Caregiver-Child Interaction Family Microbiome | Community Illness Exposure Herd Immunity | Sanitation Immunization Pathogen Exposure |
| | Behavioral | Health Behaviors Coping Strategies | Family Functioning School/Work Functioning | Community Functioning | Policies and Laws |
| | Physical/Built Environment | Personal Environment | Household Environment School/Work Environment | Community Environment Community Resources | Societal Structure |
| | Sociocultural Environment | Sociodemographics Limited English Cultural Identity Response to Discrimination | Social Networks Family/Peer Norms Interpersonal Discrimination | Community Norms Local Structural Discrimination | Social Norms Societal Structural Discrimination |
| | Health care System | Insurance Coverage Health Literacy Treatment Preferences | Patient-Clinician Relationship Medical Decision-Making | Availability of Services Safety Net Services | Quality of Care Health Care Policies |
| Health Outcomes | |  Individual Health |  Family/ Organizational Health |  Community Health |  Population Health |

FIGURE 3.
The National Institute of Minority Health Research Framework Including Domains and Levels of Influence

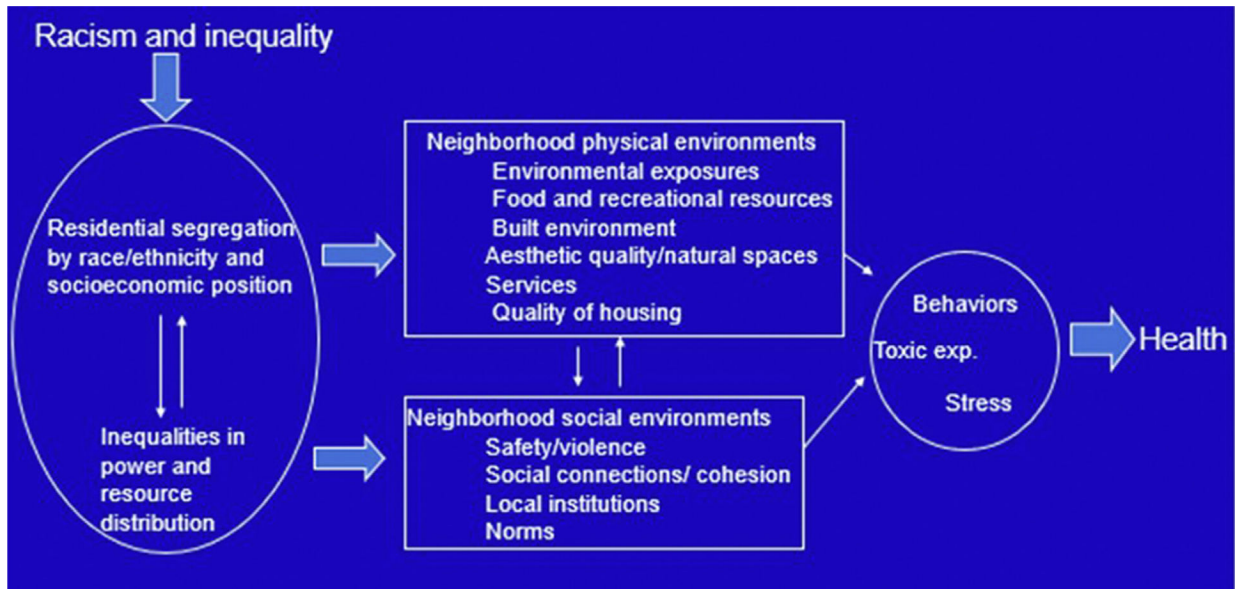


FIGURE 4. Conceptual Model of Pathways in Which Residential Segregation Can Lead to Poor Health Outcomes

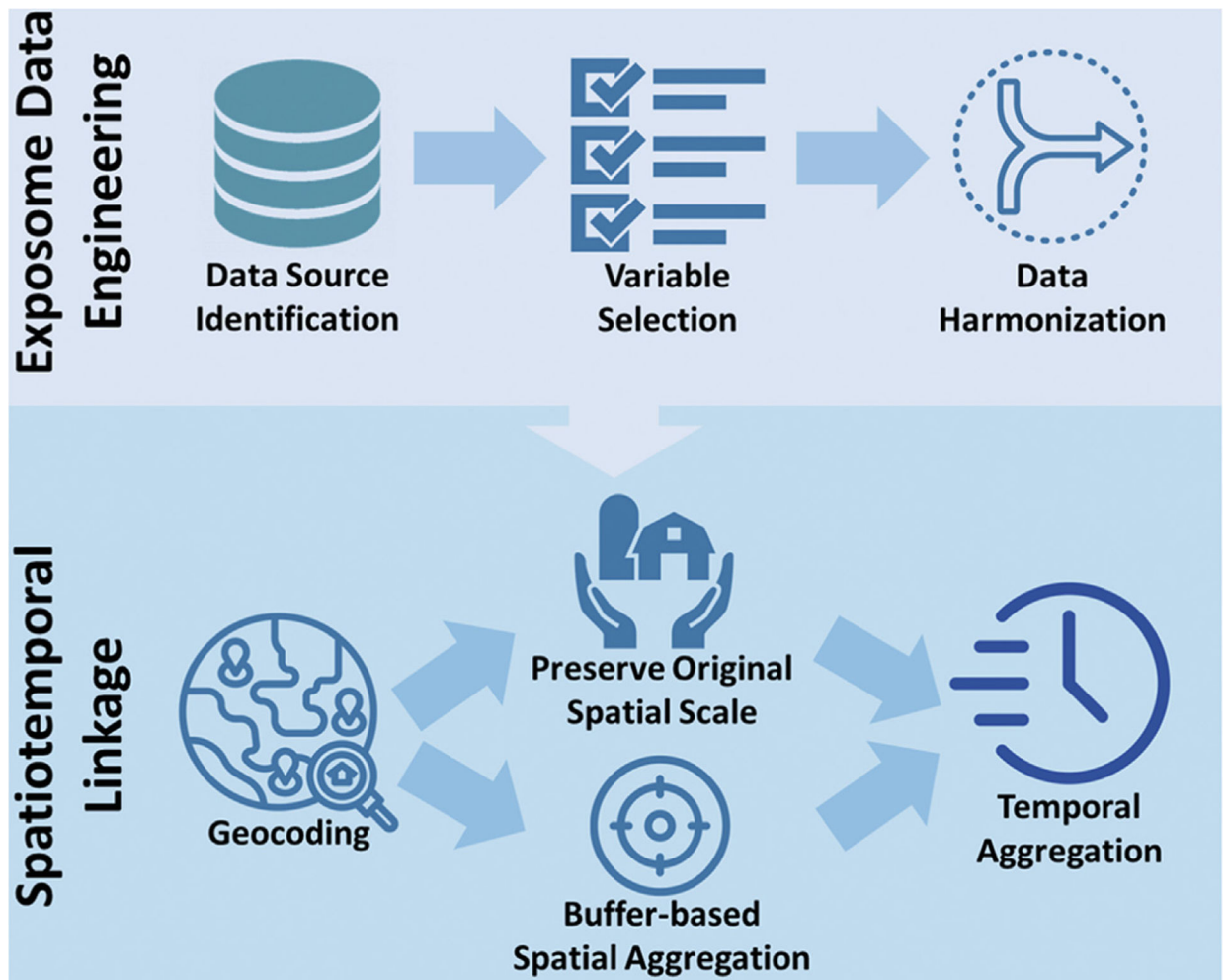


FIGURE 5. Methodological Challenges in Health Studies of Spatial and Contextual Exposomes

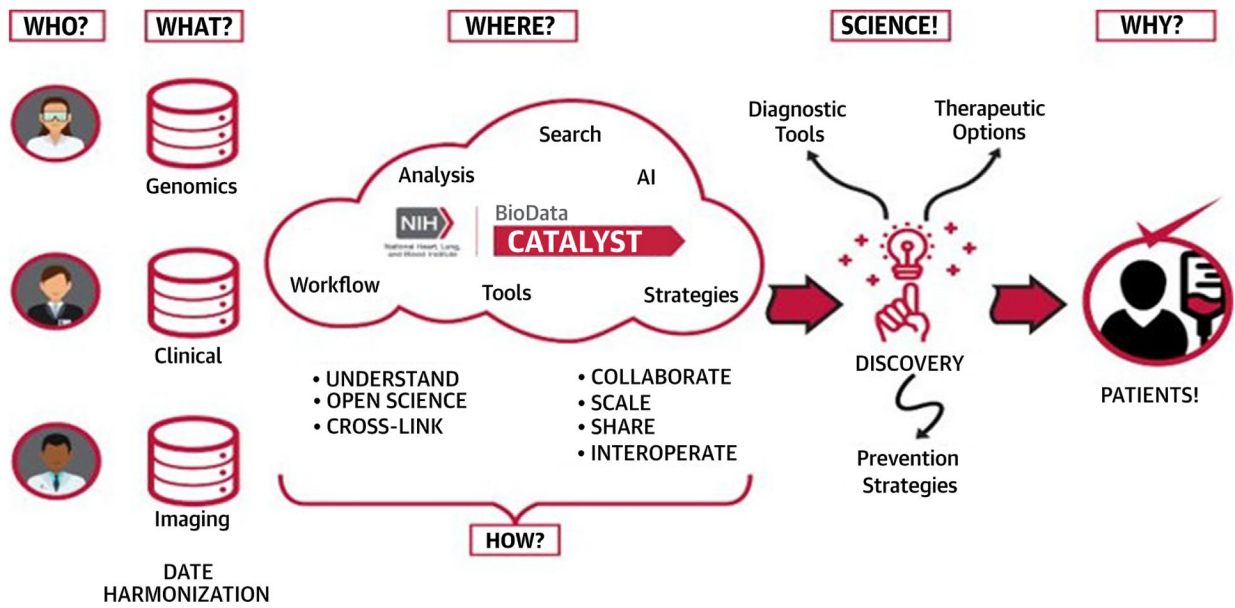


FIGURE 6. Role of NIH BioData Catalyst Consortium Using Data Harmonization to Discover Diagnostic Tools and Therapeutic Options

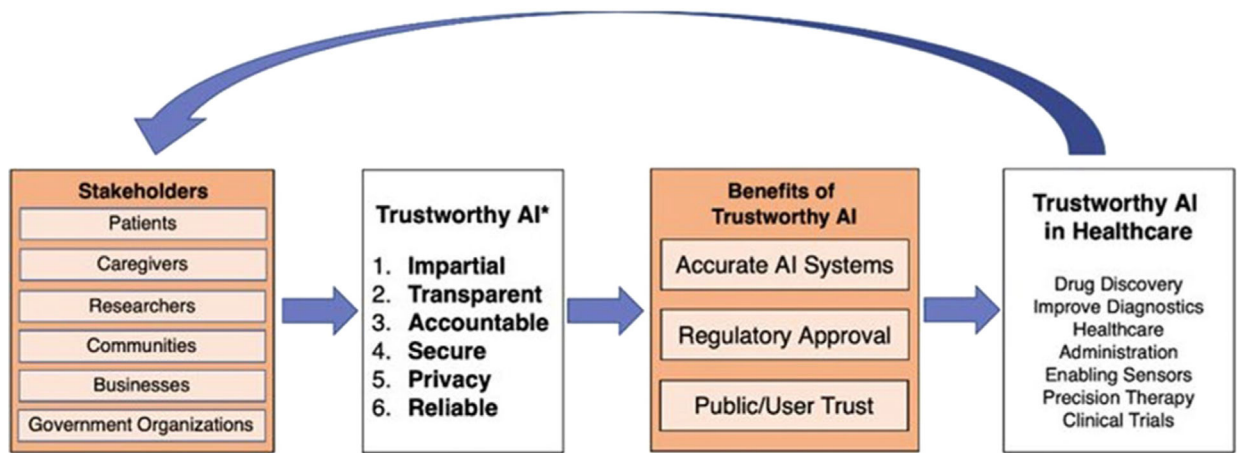


FIGURE 7. Stakeholders and Benefits of Trustworthy Artificial Intelligence/Machine Learning in Health Care

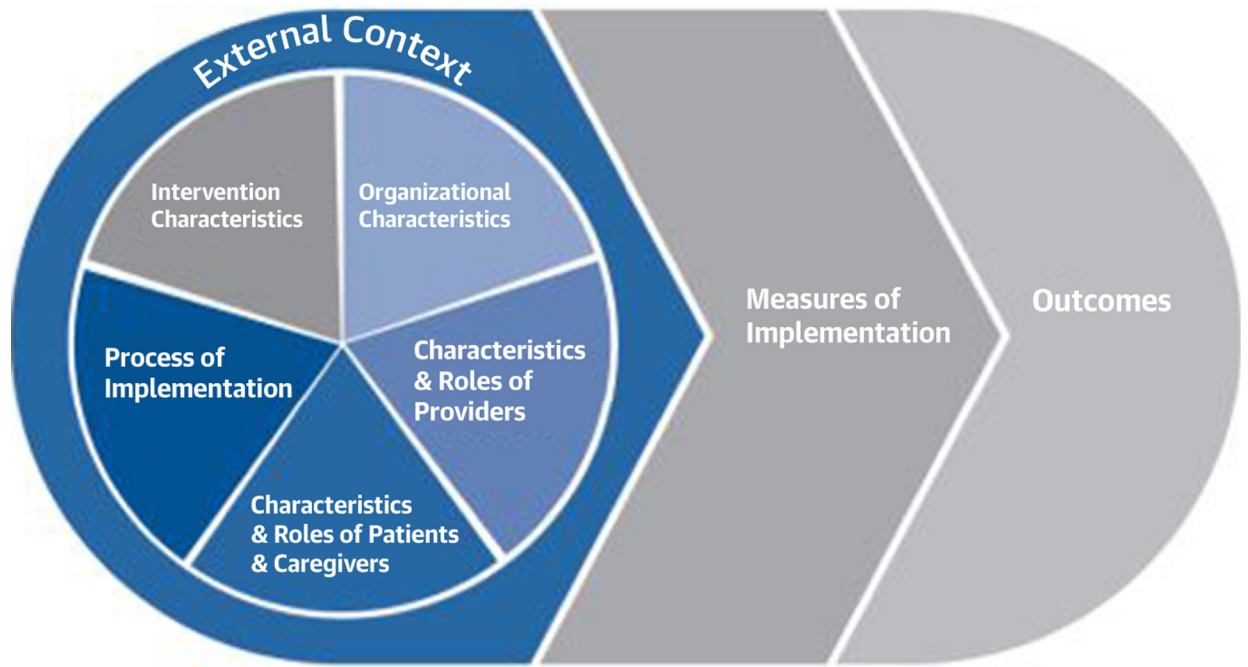


FIGURE 8. Modified Consolidated Framework for Implementation Research for Studying Process Design

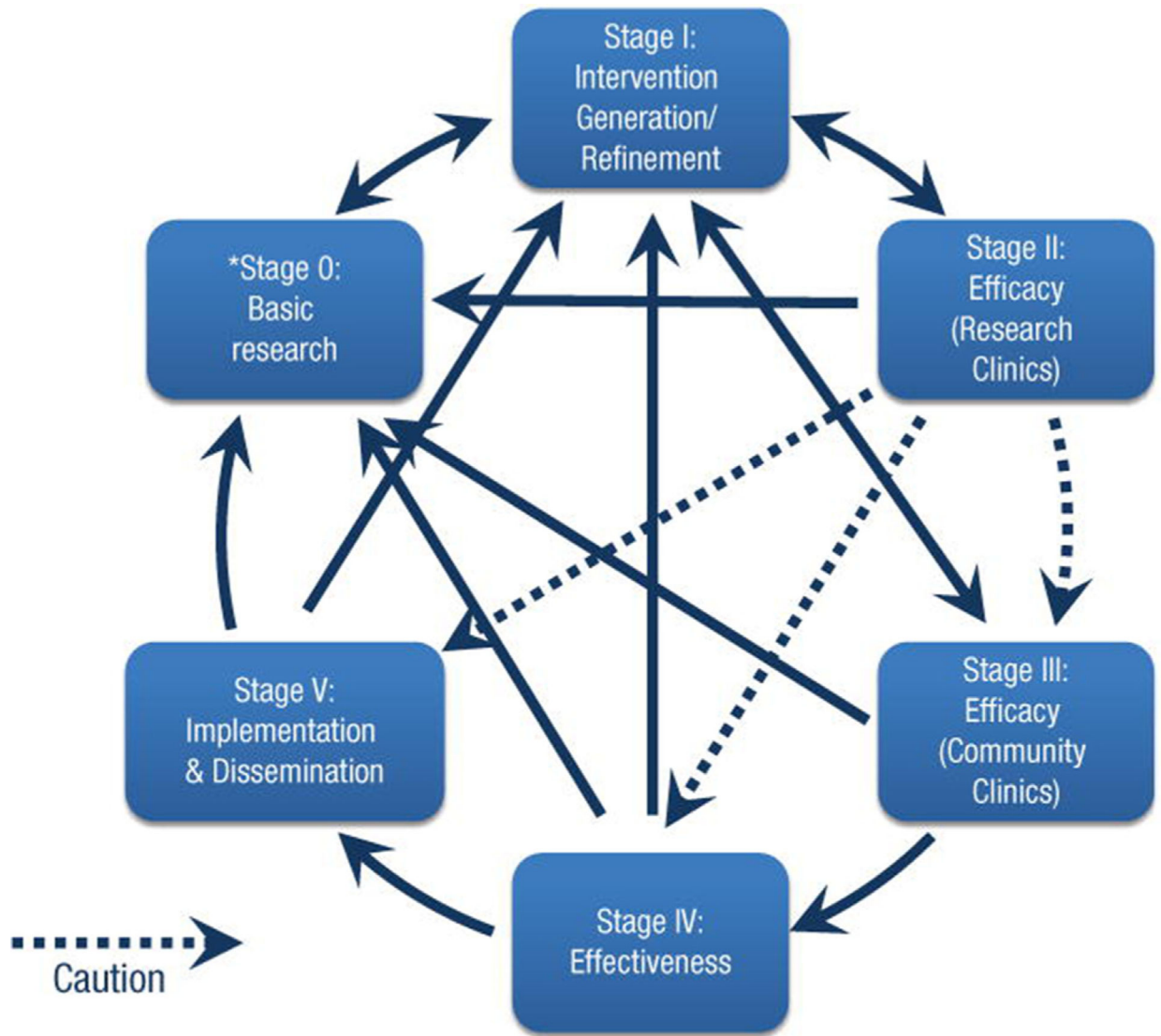


FIGURE 9.
The NIH Stages Model of Behavior Intervention Development

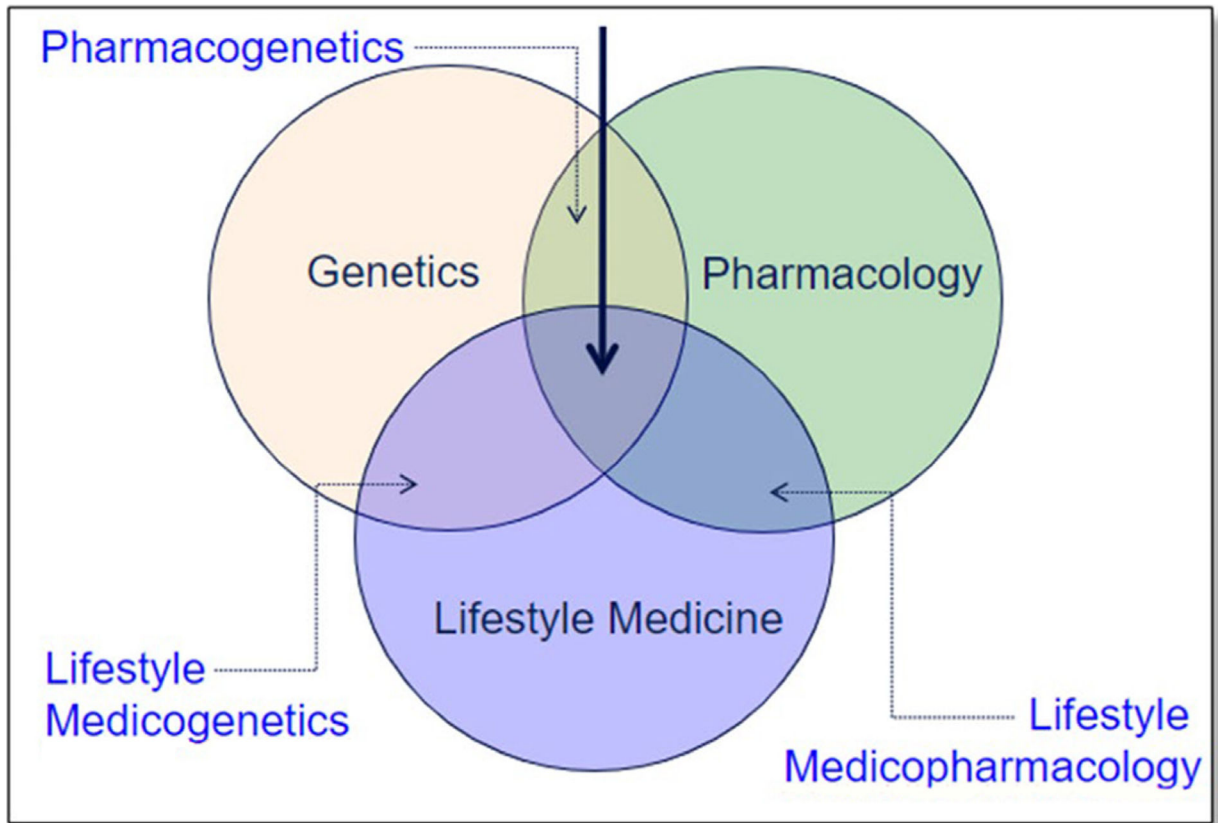
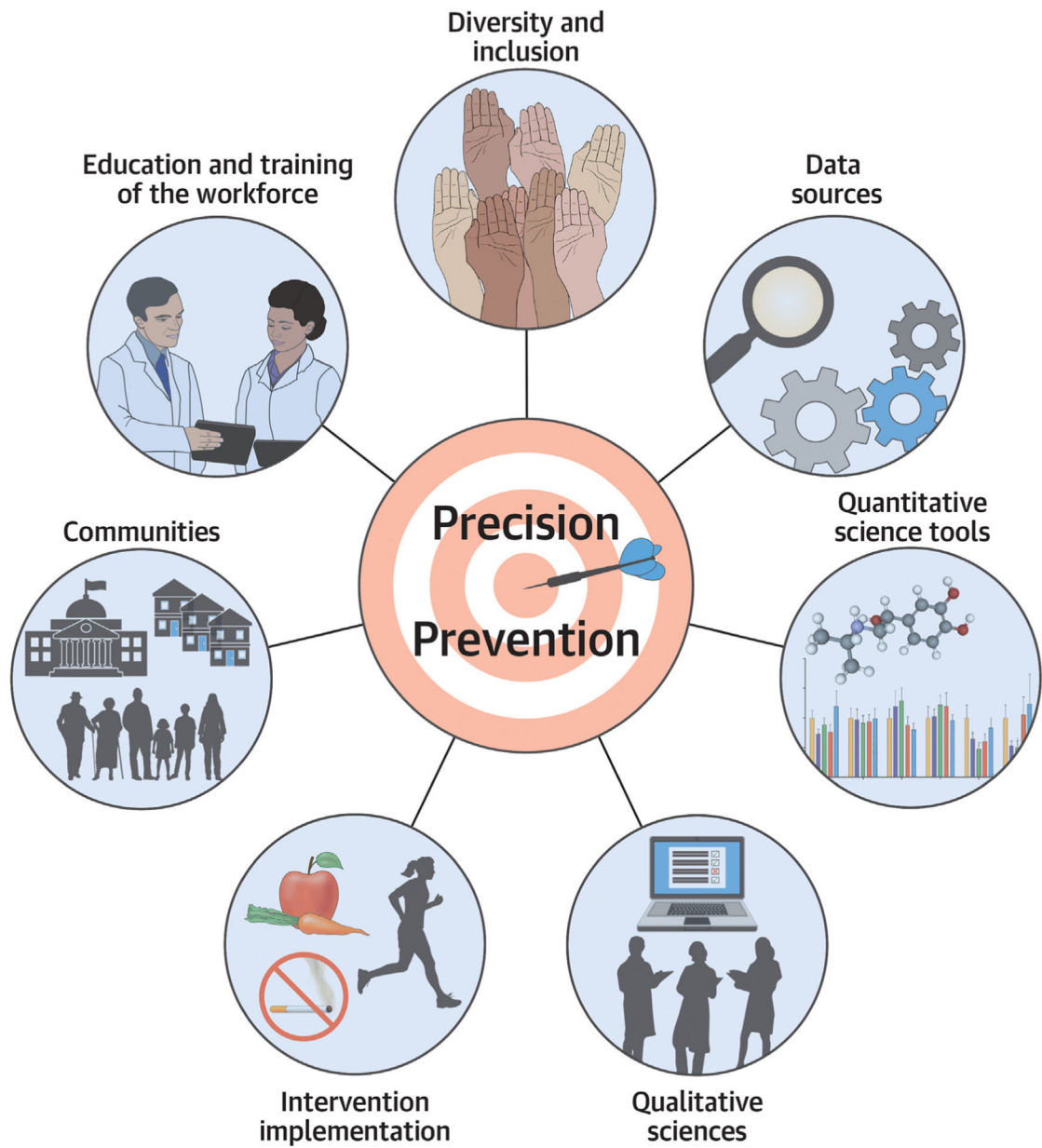


FIGURE 10.
Interaction of Genetics and Pharmacology in Personalized Lifestyle Medicine



CENTRAL ILLUSTRATION.
Dimensions of Precision Prevention

TABLE 1

Designs Applicable for Precision Prevention Intervention Studies

| |
|---|
| Individual assignment and independent delivery |
| Randomized clinical trials |
| Multiphase Optimization Strategy Trials (MOST) |
| Sequential Multiple Assignment, Randomized Trials (SMART) |
| Regression Discontinuity Design (RDO) |
| N-of-1 Trials |
| Assignment of groups or nonindependent delivery |
| Stepped wedge group-or-cluster randomized-trial |
| Individually randomized group-treatment trial Groups or cluster-based MOST, SMART, and RDD trials |

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript