Correspondence

American Life in Realtime: a benchmark registry of health data for equitable precision health

Check for updates

Social, structural and environmental determinants of health, such as food or housing insecurity, systemic racism or chronic stress, account for 60–80% of the modifiable risk of disparities in marginalized populations[1,2]. Such determinants have been difficult to address systematically because of their complexity, multidimensionality and heterogeneity (Fig. 1). Emerging precision health methods use large-scale person-generated health data from smartphones and wearables to better characterize and, ultimately, improve health and well-being through strategies customized to individual context and need[3,4]. Applying artificial intelligence and machine learning to person-generated health data allows unprecedented assessment of recursive, networked and latent associations between everyday life and health, including social, structural and environmental exposures, behaviors, biometrics, and health outcomes. Thus, precision health provides an important opportunity for reducing health disparities among minoritized racial or ethnic groups, or those who are under-resourced.

Despite the potential for improving health equity, the research community lacks benchmark training datasets of person-generated health data, which limits the ability to develop precision health models that are equally effective across diverse populations. Both the validity and the generalizability of an artificial intelligence or machine learning system are intrinsically tied to the underlying training data. The ideal benchmark dataset should feature high-quality, well-characterized data that comprehensively represent the target population in order to instill the highest standards of scientific transparency and rigor to model development, validation and evaluation.

Person-generated health data cohorts in the US National Institutes of Health's All of Us research program, UK Biobank, the Framingham Heart Study and the majority of commercial studies rely on convenience sampling and/or 'bring your own device' designs. Consequently, those who lack access to digital technologies (who tend to be older, Black, Latino, Indigenous, poorer and sicker) are systematically under-represented[5,6]. The National Health and Nutrition Examination Survey is representative, but it uses a cross-sectional design and a 1-week accelerometer measurement period, which limits its ability to assess temporal effects or account for seasonality. The absence of a benchmark dataset risks the introduction of systemic bias, exacerbation of health disparities, and causing of additional patient harm in already marginalized groups[7,8].

To bridge this critical gap, we created American Life in Realtime (ALiR), a publicly available benchmark dataset, cohort and research infrastructure for person-generated health data. ALiR has four primary objectives that advance equitable precision health: promoting inclusive representation; encouraging methodological rigor in artificial intelligence and machine learning; fostering interdisciplinary collaboration and transparency; and facilitating comprehensive exploration of the dynamic interplay between everyday life and health. Here we highlight several design choices for achieving these objectives, as well as precision health uses for the data and infrastructure.

We used several strategies to ensure that the enrolled ALiR cohort ($n = 1,038$) would achieve inclusive representation of the adult US population across demographic, socioeconomic and health factors[9]. Participants were invited from the Understanding America Study[10], an established, probability-based survey panel whose members are randomly sampled from all US addresses. To reduce digital inclusion barriers, we provided a Fitbit Inspire 2 tracker to all participants as a study incentive, and a 4G Samsung Galaxy Tablet to those who would not otherwise have internet access. We designed the participant study app to be compatible with a wide range of mobile devices and operating systems, and maintained a help-desk for technical support. As factors beyond digital inclusion, such as mistrust or privacy concerns, could lower participation among historically marginalized groups, we also oversampled people who were Black, American Indian, Alaska Native, Hawaiian, Pacific
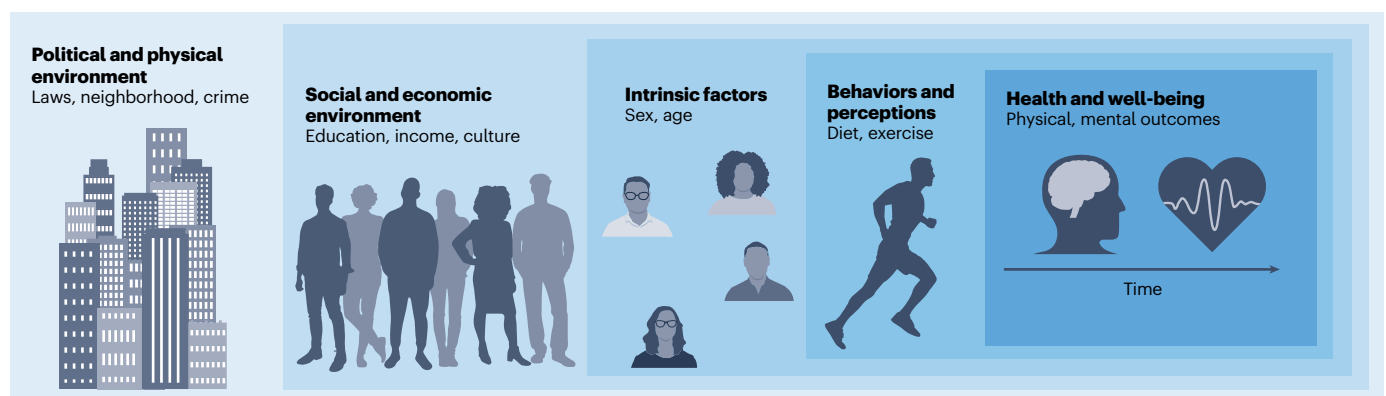
Fig. 1 | A framework for social, structural and environmental determinants of health.

Political and physical environment
Laws, neighborhood, crime

Social and economic environment
Education, income, culture

Intrinsic factors
Sex, age

Behaviors and perceptions
Diet, exercise

Health and well-being
Physical, mental outcomes

Time

Correspondence

Table 1 | ALiR's individual-level data

		Electronic surveys (repeated every 1–2 years)	Electronic surveys (repeated every 1–3 months)	Fitbit (continuous)	Public datasets (continuous)
Physical environment	Location		Zip code and/or census region		Weather
			Urbanicity		Air and water quality
	Neighborhood	Features and built environment	Perceived safety		Crime
		Transportation and walkability	Police presence		
	Housing	Type and characteristics			
Social environment	Network	Social network	Household		
			Marital status and family structure		
			Social support (PROMIS)		
	Context	Culture, religiosity and politics	Hope (PhenX)		
	Stressors	Racism and discrimination	Loneliness		
		Childhood trauma (ACEs)	Stress (PSS-4)		
Economic environment	Education and competencies	Education		Fitbit engagement	
		Financial literacy		Fitbit proficiency	
		Digital literacy			
		Health literacy			
	Work and Finance	Wealth and debt (HRS)	Labor force status and occupation		
		Financial well-being and readiness	Income		
		Housing stability	Food security (USDA)		
	Healthcare	Insurance	Access		
			Quality and satisfaction		
Intrinsic factors	Demographics and personality	Personality (Big-5)	Age		
		Decision-making	Race and ethnicity		
		Risk tolerance	Sex and gender identity		
		Mood and affect	Sexual orientation		
		Life satisfaction	Citizenship		
			Immigration generation		
			Birthplace		
			Primary language		
			Health motivations		
Health behaviors	Exercise	Physical activity (HRS)		Active and sedentary time	
				Step count	
				Distance covered	
				Walking speed	
	Diet		Diet (BRFSS)		
			Sugar consumption		
	Sleep		Sleep hygiene (PROMIS)	Time in bed	
				Sleep and wake onset	
				Napping	
	Substance use		Drugs and alcohol (ASSIST)		
	Healthcare	Prevention and screening (HRS)	Utilization (HRS)		
		Medical mistrust (PhenX)			
		Vaccine hesitancy (PhenX)			

Table 1 (continued) | ALiR's individual-level data

		Electronic surveys (repeated every 1–2 years)	Electronic surveys (repeated every 1–3 months)	Fitbit (continuous)	Public datasets (continuous)
Health outcomes	Health status	Self-reported health (HRS)	General health (BRFSS)		
	Physical health	Chronic condition diagnoses (HRS)	Functional ability (iADL and ADL)	Physical health	
		Pregnancy (PhenX)	Body mass index	Resting and active heart rate	
			Influenza-like Illness (PhenX)	Circadian rhythm	
	Sleep health	Sleep disorders (HRS)	Sleep disturbance (PROMIS)	Sleep duration	
				Sleep efficiency and latency	
				Sleep Stages	
	Mental health	Mental Illness Diagnoses (HRS)	Depression (PHQ-8)		
			Anxiety (GAD-7)		
			Substance Abuse (ASSIST)		
	Cognitive health	Cognitive status and ability			
		Episodic memory test scores			
		Mental function test scores			

The list presented here is not exhaustive. Core Understanding America Study survey measures, repeated quarterly for sociodemographics and at least biennially for other variables, are available from 2014. For ALiR, survey measures are repeated monthly, and minute-level data from wrist-worn Fitbits are collected continuously. ALiR activities occur for at least 1 year, although participants may subsequently choose to contribute indefinitely. Contextual data from public datasets are matched geospatially and temporally to participant timelines. Specific instruments are in parentheses: PROMIS, Patient Reported Outcomes Measurement Information System; PhenX, US National Institutes of Health's consensus measures for Phenotypes and eXposures; ACEs, Adverse Childhood Events Scale; PSS-4, four-item Perceived Stress Scale 4; HRS, University of Michigan's Health and Retirement Study; USDA, US Department of Agriculture; Big-5, Big Five Personality Test; BRFSS, Behavioral Risk Factor Surveillance System; ASSIST, World Health Organization's Alcohol, Smoking, and Substance Involvement Screening Test; ADL, Activities of Daily Living; iADL, instrumental ADL; PHQ-8, eight-item Patient Health Questionnaire; GAD-7, seven-item Generalized Anxiety Disorder scale.

Islander, mixed race, Hispanic or Latino, and people whose education was lower than a bachelor's degree.

These ALiR features offer several advantages. Probability sampling improves the accuracy and validity of population-level inference, such as generalizable predictions of health outcomes in response to population-level stressors like current events, systemic racism, natural disasters or surges in cases of SARS-CoV-2 infection. The provision of hardware, in our experience, eliminated sociodemographic disparities in participation rates. Oversampling resulted in a proportionally larger sample of historically under-represented and marginalized populations, providing the statistical power to detect subgroup-specific differences, such as heterogeneity in outcomes experienced by Black and Latino people (weights that rebalance the sample's demographic composition to match the US population are also provided).

To encourage the methodological rigor of artificial intelligence and machine learning, we designed a comprehensive data-collection strategy with validated, longitudinal measures (labels) of participant exposures, behaviors and outcomes over long time scales (Table 1). The measures were derived from high-quality consensus instruments such as the University of Michigan's Health and Retirement Study, the Patient Reported Outcomes Measurement Information System, and the US National Institutes of Health's consensus measures for Phenotypes and eXposures. A custom mobile app facilitates Fitbit integration, fields electronic surveys, deploys push notification announcements and reminders, and incentivizes long-term engagement through earned points that are redeemable as monetary compensation.

We created a flexible and sustainable infrastructure to foster growth, agility, collaboration and transparency. ALiR is both a data resource and a research platform, which can encourage agile and cost-efficient community collaboration. Data collected through ALiR will be available to registered users through the Understanding America Study website after curation of each cohort-year of data, with year 1 anticipated in mid-2023. With appropriate privacy and data-security safeguards, making data freely available encourages transparency, reproducibility and explainability of outputs from statistical, artificial intelligence and machine learning analyses.

The research platform can accommodate additional features, including the following: participants; application programming interfaces from wearables, medical devices, the Internet of Things, genomics and biomarkers; survey designs, including preloaded information and skip logic, randomization, experiments, and ecological momentary assessments; and interactive communications such as notifications and visual dashboards. ALiR seeks to achieve a truly large-scale sample through leveraging of the ongoing expansion of the Understanding America Study, which is expected to reach 20,000 participants by 2025, and by incorporating special populations such those with specific diseases, which are managed by academic and industry partners. Accordingly, the code will be open-sourced to encourage harmonizable data-collection efforts by others.

The ultimate aim of ALiR is to facilitate broad, multidisciplinary and equitable investigation of precision health. Engineers may leverage the platform to test the performance of new hardware or sensors in diverse populations. Methodologists may characterize factors that drive or amplify selection biases, such as social and structural patterning of study participation and attrition, data quality and 'missingness', as well as developing and testing solutions to minimize their impact, such as

Correspondence

incentive designs and imputation techniques. Social scientists may investigate the clustering and importance of social determinants in various populations to prioritize public health investments. Behavioral researchers may develop just-in-time interventions where deviations from individual-specific baselines trigger automated 'nudges' and/or suggestions, such as passive detection of influenza-like symptoms via Fitbit data, which triggers a recommendation for SARS-CoV-2 testing. Operations researchers may evaluate utility of caseworker or health system integrations.

Collectively, ALiR is a model for achieving diversity, equity, inclusion, transparency and multi-disciplinary collaboration in precision health.

Ritika R. Chaturvedi ◉[1]✉, Marco Angrisani[2], Wendy M. Troxel[3], Tania Gutsche[2], Eva Ortega[2], Monika Jain[4], Adrien Boch[4] & Arie Kapteyn[2]

[1]Leonard D. Schaeffer Center for Health Policy & Economics, University of Southern California, Los Angeles, CA, USA. [2]Center for Economic and Social Research, University of Southern California, Los Angeles, CA, USA. [3]RAND Corporation, Division of Social and Economic Wellbeing, Pittsburgh, PA, USA. [4]Evidation Health, San Mateo, CA, USA. ✉e-mail: ritikac@usc.edu

References

1. Hood, C. M. et al. *Am. J. Prev. Med.* **50**, 129–135 (2016).
2. Bradley, E. H. et al. *BMJ Qual. Saf. Health Care* **20**, 826–831 (2011).
3. Collins, F. S. et al. *N. Engl. J. Med.* **372**, 793–795 (2015).
4. Khoury, M. J. et al. *J. Am. Med. Assoc.* **316**, 1357–1358 (2016).
5. Cho, Peter Jaeho, et al. *JMIR mHealth uHealth* **10.4**, e29510 (2022).
6. Pratap, A. et al. *NPJ Digit. Med.* **3**, 21 (2020).
7. Zou, J. & Schiebinger, L. *Nature* **559**, 324–326 (2018).
8. Obermeyer, Z. et al. *Science* **366**, 447–453 (2019).
9. USC Schaeffer Center. https://healthpolicy.usc.edu/article/biomedical-expert-designs-health-dataset-representing-all-demographic-groups (accessed August 2022).
10. Alattar, L., Messel, M. & Rogofsky, D. *Soc. Secur. Bull.* **78**, 13–28 (2018).