Journal of the American Heart Association

BRIEF COMMUNICATION

Cardiovascular Health Disparities in Racial and Other Underrepresented Groups: Initial Results From the All of Us Research Program

Julián N. Acosta, MD*; Audrey C. Leasure, BS*; Cameron P. Both, BS; Natalia Szejko, MD, PhD; Stacy Brown, MD; Victor Torres-Lopez, BS; Safa Abdelhakim, MD; Joseph Schindler, MD; Nils Petersen, MD, PhD, MSc; Lauren Sansing, MD, MMS; Thomas M. Gill, MD; Kevin N. Sheth, MD† Guido J. Falcone, MD, ScD, MPH†

BACKGROUND: All of Us is a novel research program that aims to accelerate research in populations traditionally underrepresented in biomedical research. Our objective was to evaluate the burden of cardiovascular disease (CVD) in broadly defined underrepresented groups.

METHODS AND RESULTS: We evaluated the latest data release of All of Us. We conducted a cross-sectional analysis combining survey and electronic health record data to estimate the prevalence of CVD upon enrollment in underrepresented groups defined by race, ethnicity, age (>75 years), disability (not able to carry out everyday physical activities), sexual orientation and gender identity lesbian, gay, bisexual, transgender, queer, intersex, and asexual (LGBTQIA+), income (annual household income <$35 000 US dollars) and education (less than a high school degree). We used multivariate logistic regression to estimate the adjusted odds ratio (OR) and product terms to test for interaction. The latest All of Us data release includes 315 297 participants. Of these, 230 577 (73%) had information on CVD and 17 958 had CVD (overall prevalence, 7.8%; 95% CI, 7.7–7.9). Multivariate analyses adjusted by hypertension, hyperlipidemia, type 2 diabetes mellitus, body mass index, and smoking indicated that, compared with White participants, Black participants had a higher adjusted odds of CVD (OR, 1.21; 95% CI, 1.16–1.27). Higher adjusted odds of CVD were also observed in underrepresented groups defined by other factors, including age >75 years (OR, 1.90; 95% CI, 1.81–1.99), disability (OR, 1.60; 95% CI, 1.53–1.68), and income <$35 000 US dollars (OR, 1.22; 95% CI, 1.17–1.27). Sex significantly modified the odds of CVD in several of the evaluated groups.

CONCLUSIONS: Among participants enrolled in All of Us, underrepresented groups defined based on race, ethnicity and other factors have a disproportionately high burden of CVD. The All of Us research program constitutes a powerful platform to accelerate research focused on individuals in underrepresented groups.

Key Words: All of Us ■ cardiovascular disease ■ disparities research ■ myocardial infarction ■ stroke

Cardiovascular disease (CVD) is a well-established determinant of morbidity and mortality across the lifespan.[1] Mounting evidence indicates that racial and ethnic underrepresented groups carry a disproportionate burden of CVD.[2] It is increasingly recognized that underrepresented

Correspondence to: Guido J. Falcone, MD, ScD, MPH, 15 York Street, LLCI Room 1004D, P.O. Box 208018, New Haven, CT 06520. E-mail: guido.falcone@yale.edu

Supplementary Material for this article is available at https://www.ahajournals.org/doi/suppl/10.1161/JAHA.121.021724

*J. N. Acosta and A. C. Leasure are co-first authors.

†K. N. Sheth and G. J. Falcone jointly supervised this work.

For Sources of Funding and Disclosures, see page 7.

JAHA is available at: www.ahajournals.org/journal/jaha

groups defined by features other than race and ethnicity may also be at higher risk of CVD.[3] As an example, although CVD prevalence has decreased consistently in high-resource groups, this is not true for the remainder of the population, where CVD prevalence has decreased less consistently, remained stable, or even increased.[4] Another relevant example pertains to transgender people, for whom the prevalence of myocardial infarction is higher than in the general population.[5] The recently introduced All of Us Research Program seeks to accelerate precision medicine research by acquiring and publicly sharing health data from 1 million Americans.[6] Because one of its goals is to reduce health disparities across underrepresented groups, All of Us provides an updated framework to define underrepresented groups based not only on race and ethnicity but also on age, disability, education, income, and gender identity and sexual orientation. We used the latest release of All of Us data to evaluate cardiovascular health disparities in underrepresented groups defined by this novel framework.

METHODS

Data Availability

All data are publicly available at www.allofus.nih.gov. All analyses were conducted within a secure informatic workspace provided by the National Institutes of Health that allows users to access and analyze a centralized version of the All of Us data. All study participants provided written informed consent. We used release version number 4, which comprises data from all participants who enrolled from the beginning of the program on May 30, 2017, to August 1, 2020.

Study Design

The All of Us Research Program protocol has been previously published.[7] The All of Us protocol and materials have been approved by a dedicated institutional review board, the All of Us Institutional Review Board. Briefly, All of Us aims to enroll at least 1 million Americans who agree to share their electronic health record data, donate biospecimens, respond to surveys, and have standardized physical measurements taken. Inclusion criteria are age ≥18 years and capacity to provide consent.

Baseline Data

Participants are given several baseline health surveys and undergo an evaluation for physical measurements. All of Us uses several means to collect longitudinal health data, including continuous abstraction of electronic health record data in the form of billing codes, laboratory and medication data, radiology reports, and narrative content and linkage with other data sources.[8] Because "date of consent," required to calculate age, was missing for a substantial portion of the study participants, it was imputed using the age at the start of the study (May 30, 2017) when missing.

Identification of Underrepresented Populations

Given the wide variation in the definition of underrepresented groups, we followed the framework and definitions provided by All of Us (Figure S1). We used information from the baseline survey to identify self-reported race and ethnicity, which comprised the following categories: Hispanic/Latino/a/x participants (regardless of race), non-Hispanic White participants, non-Hispanic Black participants, non-Hispanic Asian participants, non-Hispanic other participants (including participants reporting another single population or none of these), and non-Hispanic >1 race selected. Other underrepresented groups were identified based on age (older adults, as defined by age >75 years), physical disability (those who answer that they cannot carry out every day physical activities at all or only a little), education (less than a high school degree), income level (<35 000 US dollars of yearly household income), and sex at birth (nonbinary), sexual orientation (participants who identify as asexual, bisexual, gay, or lesbian) or gender identity (participants who identify as something else than their sex at birth).

Ascertainment of CVD

We defined CVD as a composite of coronary artery disease or stroke (both ischemic and hemorrhagic). Coronary heart disease was identified from electronic health records using the *International Classification of Diseases, Ninth Revision* (*ICD-9*) diagnostic codes, *ICD-9* procedure codes, *International Classification of Diseases, Tenth Revision* (*ICD-10*) diagnostic codes, and *ICD-10* procedure codes following the MidSouth Clinical Data Research Network Coronary Heart Disease Algorithm (Table S1) or by answering affirmatively to either of the following questions: "Has a doctor or health care provider ever told you that you had a heart attack?" "Has a doctor or health care provider ever told you that you have coronary artery/coronary heart disease?" Stroke cases were identified using previously validated *ICD-9* and *ICD-10* codes (Table S1) or by answering affirmatively to the question, "Has a doctor or health care provider ever told you that you had a stroke?"[9–11] We used a similar approach combining self-reported responses to the past medical

history survey and data from diagnosis codes in the electronic health record data to ascertain the presence of prominent vascular risk factors, including hypertension (Observational Medical Outcomes Partnership code 316866), hyperlipidemia (Observational Medical Outcomes Partnership code 432867), and type 2 diabetes mellitus (Observational Medical Outcomes Partnership code 201826), and used self-reported data from the lifestyle survey to ascertain smoking status. In addition, we used data from physical measurements to calculate the body mass index.

Statistical Analysis

We used chi-squared tests and the *t* test, ANOVA, or Mann–Whitney *U* test for unadjusted comparisons of discrete and continuous variables, respectively. We estimated the prevalence of CVD as the number of study participants with CVD divided by the total number of study participants with available data and calculated the 95% CI for the proportion using the formula $p \pm 1.96 \times \sqrt{\frac{p \times (1-p)}{n}}$ where p is the proportion estimate in the sample, and n is the sample size. We used univariable and multivariable logistic regression to estimate the odds ratios (ORs) of CVD for each underrepresented group. Multivariable models were adjusted for sex, age, and vascular risk factors. Our primary analysis was a complete case analysis. In sensitivity analysis, we imputed missing data using predictive mean matching for continuous data, logistic regression imputation for dichotomous data, and multinomial regression imputation for discrete data with >2 categories. We used product terms to test for interaction between sex at birth and underrepresented groups. A 2-tailed *P* value of <0.05 was considered to be statistically significant. Analysis was conducted using R version 4.0.5 in a Jupyter Notebook environment.

RESULTS

The current release of All of Us includes data from 315 297 participants. Of these, 230 577 (73%) had information on CVD. Age at consent was 51.217 (meanSD) years, and 141 896 (61%) were women (Table 1). The overall prevalence of CVD was 7.8% (n=17 958; 95% CI, 7.7–7.9), varying widely across several of the underrepresented groups (Figure). Underrepresented groups with a higher prevalence of CVD also had worse cardiovascular health based on the evaluation of 5 prominent risk factors (Table 1 and Table S2).

Multivariate analyses adjusting for age, sex, and vascular risk factors indicated that, compared with White participants, Black participants had a higher adjusted odds of CVD (OR, 1.21; 95% CI, 1.16–1.27). Higher adjusted odds of CVD were also observed in underrepresented groups defined by factors other than race and ethnicity (Table 2), including age >75 years (OR, 1.90; 95% CI, 1.81–1.99), disability (OR, 1.60; 95% CI, 1.53–1.68), and income <$35 000 US dollars (OR, 1.22; 95% CI, 1.17–1.27). In contrast, Hispanic/Latino/a/x participants (OR, 0.84; 95% CI, 0.79–0.89), Asian participants (OR, 0.85; 95% CI, 0.74–0.98), and people with less than a high school degree (OR, 0.90; 95% CI, 0.85–0.96) had lower adjusted odds of CVD. Full results are displayed in Table 2. Sensitivity analysis imputing missing data yielded consistent results (data not shown).

Sex significantly modified the odds of CVD identified in several of the evaluated groups (Table S3). The higher odds of CVD in Black participants versus White participants was driven by women (OR, 1.40; 95% CI, 1.32–1.49; interaction, *P*<0.001), without a significant difference in men (OR, 1.02; 95% CI, 0.95–1.10). In underrepresented groups defined by factors other than race, ethnicity, older age and male sex synergistically increased the odds of CVD (interaction, *P*<0.001), with older men versus women having significantly higher estimates (OR, 2.11 [95% CI, 1.98–2.25] versus OR, 1.69 [95% CI, 1.57–1.81]). Similarly, disability and female sex synergistically increased the odds of CVD (interaction, *P*<0.001), with women who were disabled versus women who were not disabled having significantly higher estimates (OR, 1.82 [95% CI, 1.71–1.94] versus OR, 1.37 [95% CI, 1.27–1.48]).

DISCUSSION

We report the results of a cross-sectional study aimed to evaluate cardiovascular health disparities across different underrepresented groups enrolled in All of Us, the largest population-based open-access study implemented in the United States to date. Following a novel approach proposed by the All of Us program, we studied underrepresented groups defined not only by race and ethnicity but also based on age, disability, education, income, and gender identity and sexual orientation.[3] We found that several of these underrepresented groups had higher burdens of CVD.

This study provides important evidence confirming the scientific consistency of the first data release of All of Us. Several large observational studies have found that Black participants have a higher prevalence of CVD.[12–15] Although the unadjusted CVD prevalence was lower in Black participants in our analysis, multivariate analysis showed higher adjusted odds in this group. In addition, we found a lower CVD prevalence in Hispanic/Latino/a/x participants, in concordance with previous work showing lower CVD prevalence and mortality in this population, a phenomenon termed “the Hispanic paradox,”

Table 1. Underrepresented Populations in the All of Us Cohort

Underrepresented groups	All enrolled people (n=315 297)	No PMH/EHR data (n=84 720)	With PMH/EHR data (n=230 577)					
			All*	Hypertension	Hyperlipidemia	Type 2 DM	Ever smoked	BMI
Race/ethnicity, n (%)								
White	162 330 (51.5)	36 006 (42.5)	126 324 (54.8)	54 367 (43.0)	58 687 (46.5)	17 059 (13.5)	52 629 (42.5)	29.19 (7.18)
Black	66 954 (21.2)	23 584 (27.8)	43 370 (18.8)	23 044 (53.1)	14 370 (33.1)	9893 (22.8)	20 372 (48.9)	31.62 (8.80)
Hispanic/Latino/a/x	59 283 (18.8)	16 887 (19.9)	42 396 (18.4)	15 317 (36.1)	12 756 (30.1)	8386 (19.8)	12 227 (29.6)	30.72 (7.32)
Asian	10 276 (3.3)	3500 (4.1)	6776 (2.9)	1682 (24.8)	2000 (29.5)	759 (11.2)	1088 (16.4)	25.31 (5.21)
Other†	5470 (1.7)	1546 (1.8)	3924 (1.7)	1532 (39.0)	1453 (37.0)	666 (17.0)	1631 (43.1)	29.53 (7.54)
>1	4950 (1.6)	1348 (1.6)	3602 (1.6)	1128 (31.3)	1017 (28.2)	398 (11.0)	1364 (39.1)	29.39 (8.00)
Did not answer	6034 (1.9)	1849 (2.2)	4185 (1.8)	2007 (48.0)	1708 (40.8)	821 (19.6)	1729 (48.9)	29.66 (7.42)
Age, n (%)								
<75 y	297 030 (94.2)	81 577 (96.3)	215 453 (93.4)	88 048 (40.9)	80 880 (37.5)	34 570 (16.0)	83 612 (39.9)	29.99 (7.75)
>75 y	18 267 (5.8)	3143 (3.7)	15 124 (6.6)	11 029 (72.9)	11 111 (73.5)	3412 (22.6)	7428 (50.2)	28.17 (5.60)
Disability, n (%)								
Without disability	272 982 (86.6)	68 896 (81.3)	204 086 (88.5)	83 105 (40.7)	79 545 (39.0)	29 738 (14.6)	78 283 (39.2)	29.52 (7.33)
With disability	30 670 (9.7)	7777 (9.2)	22 893 (9.9)	14 325 (62.6)	11 075 (48.4)	7506 (32.8)	11 644 (52.4)	33.03 (9.45)
Did not answer	11 645 (3.7)	8047 (9.5)	3598 (1.6)	1647 (45.8)	1371 (38.1)	738 (20.5)	1113 (48.3)	29.56 (7.24)
Sex/gender, n (%)								
LGBTQIA+, no	272 870 (86.5)	72 065 (85.1)	200 805 (87.1)	87 372 (43.5)	81 884 (40.8)	33 241 (16.6)	78 085 (39.8)	29.85 (7.58)
LGBTQIA+, yes	42 427 (13.5)	12 655 (14.9)	29 772 (12.9)	11 705 (39.3)	10 107 (33.9)	4741 (15.9)	12 955 (46.0)	30.01 (8.02)
Education, n (%)								
High school completed	275 881 (87.5)	71 649 (84.6)	204 232 (88.6)	86 432 (42.3)	82 695 (40.5)	31 390 (15.4)	78 098 (39.2)	29.78 (7.61)
Less than high school	31 984 (10.1)	10 414 (12.3)	21 570 (9.4)	10 320 (47.8)	7589 (35.2)	5605 (26.0)	10 858 (52.1)	30.72 (7.90)
Did not answer	7432 (2.4)	2657 (3.1)	4775 (2.1)	2325 (48.7)	1707 (35.7)	987 (20.7)	2084 (48.4)	29.69 (7.60)
Income, n (%)								
Household income >35 000 US dollars	141 199 (44.8)	32 363 (38.2)	108 836 (47.2)	43 115 (39.6)	48 101 (44.2)	13 158 (12.1)	36 446 (34.1)	29.06 (6.86)
Household income ≤35 000 US dollars	111 266 (35.3)	33 835 (39.9)	77 431 (33.6)	35 719 (46.1)	27 245 (35.2)	16 055 (20.7)	38 570 (51.3)	30.99 (8.48)
Did not answer	62 832 (19.9)	18 522 (21.9)	44 310 (19.2)	20 243 (45.7)	16 645 (37.6)	8769 (19.8)	16 024 (37.9)	29.78 (7.52)

Columns for risk factors present the prevalence of each risk factor in each subgroup. BMI indicates body mass index; CVD, cardiovascular disease; DM, diabetes mellitus; EHR, electronic health record; > 1, non-Hispanic >1 race selected; HSD, high school degree; LGBTQIA+, lesbian, gay, bisexual, transgender, queer, intersex, and asexual; and PMH, past medical history.

*All participants with available EHR or PMH data (primary analytic sample).

†Participants who did not self-report as "Hispanic, Latino or Spanish," the "other" category comprises the following two categories from All of Us questionnaires. Another single population: participants self-reporting either Middle Eastern or North African or Native Hawaiian or other Pacific Islander (please note All of Us does not provide disaggregated data on these yet). None of these populations: participants self-reporting "None of these fully describe me" (options are White, Black, African American, or African, Asian, Middle Eastern or North African, Native Hawaiian or other Pacific Islander).

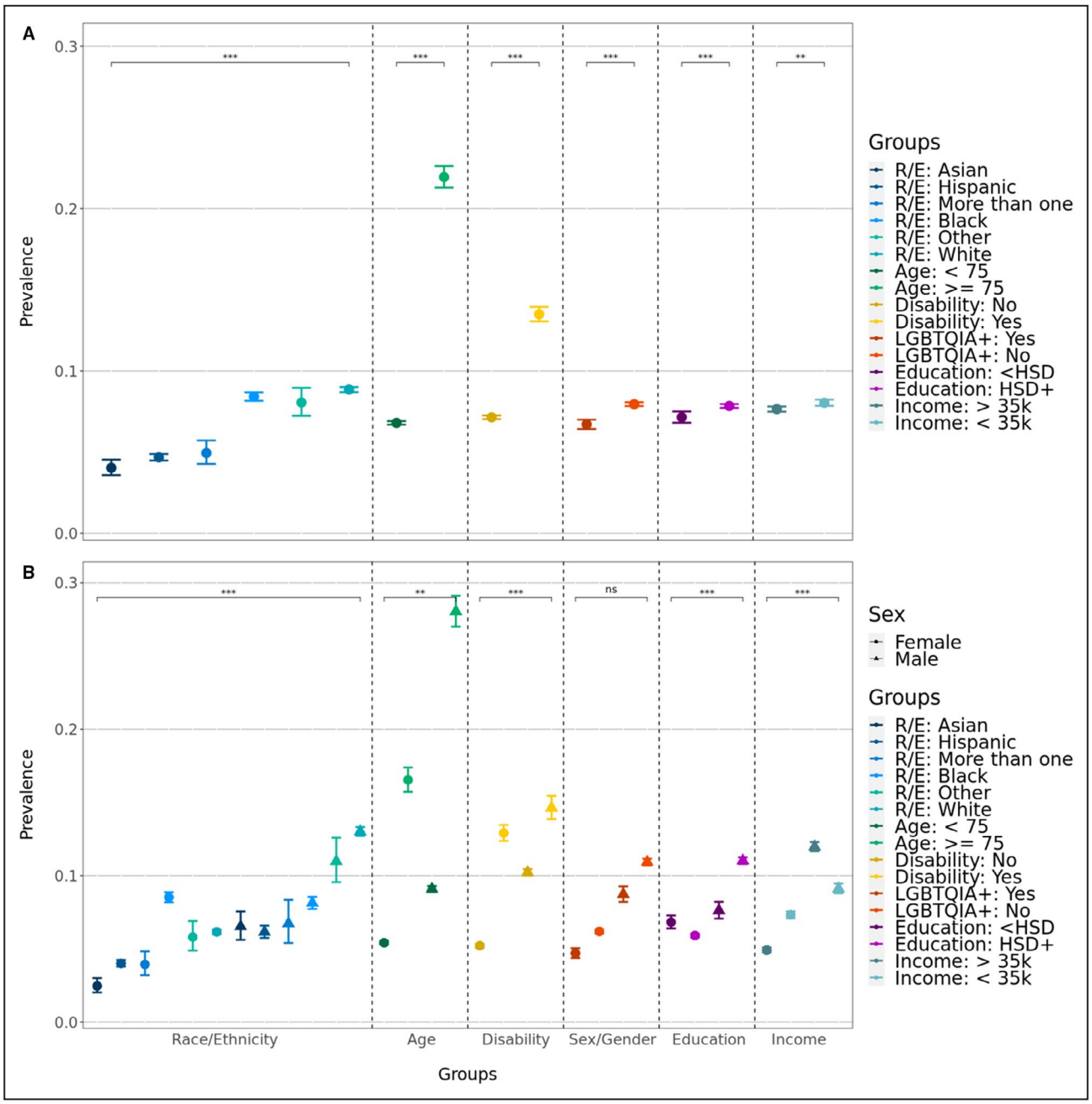

Figure. Prevalence of cardiovascular disease in underrepresented groups enrolled in All of Us.
A, Point estimates and 95% CIs for cardiovascular disease prevalence across underrepresented groups enrolled in the All of Us research program. **B**, Same analyses after stratifying by gender. Error bars correspond to 95% CIs. ns=nonsignificant. ***P*<0.05. ****P*<0.001. All *P* values correspond to univariable analyses. HSD indicates high school degree; LGBTQIA+, lesbian, gay, bisexual, transgender, queer, intersex, and asexual; and R/E, race/ethnicity.

with several hypotheses proposed as possible explanations, including behavioral (the acculturation theory and the healthy migrant hypothesis), nutritional, genetic, and psychosocial characteristics.[16] The prevalence of CVD was also lower in Asian individuals, as has been reported previously.[17]

Our results extend the existing knowledge in the field of CVD disparities by showing that individuals in underrepresented groups defined by factors other than race and ethnicity carry a disproportionate burden of CVD, including people who were older, disabled, or had a household income <$35 000. In addition, our findings highlight the prominent role of sex as an effect modifier within several of these underrepresented groups. Understanding the specific genetic, social, and biological mechanisms that

Table 2. Prevalence of Cardiovascular Disease in Underrepresented Groups Enrolled in All of Us

Group	Age in y, mean±SD	Female sex at birth, n (%)	Prevalence estimate (95% CI)	Univariable regression, OR (95% CI)	Multivariable regression, OR (95% CI)*
Race/ethnicity					
White	54.7±16.8	76 653 (61)	8.8 (8.7–9.0)	Reference	Reference
Asian	43.1±16.7	4138 (61)	4.0 (3.6–4.5)	0.43 (0.38–0.49)	0.85 (0.74–0.98)
Hispanic/Latino/a/x	44.6±15.8	28 949 (68)	4.7 (4.5–4.8)	0.51 (0.48–0.53)	0.84 (0.79–0.89)
Other‡	48.6±16.7	4532 (60)	8.0 (7.2–9.0)	0.90 (0.80–1.01)	1.20 (1.05–1.36)
>1	42.5±16.7	2362 (65.6)	4.9 (4.3–5.7)	0.54 (0.46–0.62)	1.01 (0.85–1.20)
Black	49.5±14.5	25 646 (59)	8.4 (8.2–8.7)	0.95 (0.91–0.98)	1.21 (1.16–1.27)
Age					
<75 y	49.2±15.5	134 079 (62)	6.8 (6.7–6.9)	Reference	Reference
≥75 y	79.0±3.0	7817 (52)	21.9 (21.3–22.6)	3.86 (3.70–4.02)	1.90 (1.81–1.99)
Disability					
No	51.0±17.0	125 532 (61)	7.1 (7.0–7.2)	Reference	Reference
Yes	53.1±13.9	14 689 (64)	13.5 (13.0–13.9)	2.03 (1.95–2.11)	1.60 (1.53–1.68)
Sex/gender					
LGBTQIA+, no	51.9±16.6	126 248 (63)	7.9 (7.8–8.1)	Reference	Reference
LGBTQIA+, yes	46.7±16.9	15 648 (53)	6.7 (6.4–7.0)	0.83 (0.79–0.87)	1.01 (0.95–1.07)
Education					
High school degree or more	51.3±17.0	126 799 (62)	7.8 (7.7–7.9)	Reference	Reference
Less than a high school degree	49.8±14.9	12 901 (60)	7.1 (6.8–7.5)	0.90 (0.86–0.95)	0.90 (0.85–0.96)
Income†					
>$35 000 US dollars	53.2±16.6	66 746 (61)	7.6 (7.5–7.8)	Reference	Reference
<$35 000 US dollars	48.4±16.3	47 470 (61)	8.0 (7.8–8.2)	1.06 (1.02–1.09)	1.22 (1.17–1.27)

LGBTQIA+ indicates lesbian, gay, bisexual, transgender, queer, intersex, and asexual; OR, odds ratio; and > 1, non-Hispanic >1 race selected.

*Following are the number of records excluded per model: race/ethnicity=28 370, age=24 997, disability=27 190, sex/gender=24 997, education=29 155, income=65 522.

†Income corresponds to annual household income.

‡Participants who did not self-report as "Hispanic, Latino or Spanish,",the "other" category comprises the following two categories from All of Us questionnaires: Another single population: participants self-reporting either Middle Eastern or North African or Native Hawaiian or other Pacific Islander (please note All of Us does not provide disaggregated data on these yet). None of these populations: participants self-reporting "None of these fully describe me" (options are White, Black, African American, or African, Asian, Middle Eastern or North African, Native Hawaiian or other Pacific Islander).

mediate the observed higher burden of CVD in these underrepresented groups is beyond the scope of this work, but the clear identification of these disparities sets the stage for follow-up research on this front. In accordance with prior reports, participants with a lower education level had an overall worse profile of cardiovascular risk factors.[18] However, the observed lower adjusted odds of CVD in this group indicates that the All of Us cohort may have differences with other cohorts that have been used to study health disparities.

Our work has a number of limitations. First, the cross-sectional design precludes the possibility of deriving any causal conclusions from these analyses. Second, observational studies can be subject to “volunteer bias,” as healthy people are more likely to enroll in these studies. Although recall bias could also be present, the combined use of electronic health record and self-reported data to ascertain outcomes and risk factors limits the impact of this type of bias. Third, information on disaggregated Hispanic/Latino/a/x ethnicity and Asian descent is not yet available in All of Us, and data from the overall Hispanic/Latino/a/x and Asian categories are not necessarily representative of any individual subgroup within these 2 broad race and ethnic categories. Finally, although unmeasured confounders could play a role, these results are descriptive in nature and are not intended to draw causal conclusions.

In summary, we report the findings of the first study based on All of Us focused on evaluating the burden of CVD in underrepresented groups. We offer important evidence of the resource’s potential and report a higher burden of CVD in underrepresented groups defined by factors other than race and ethnicity. Our results underscore the urgency of addressing cardiovascular health disparities across broadly defined underrepresented groups and point to All of Us, which will soon release genomic and other layers of data, as a promising resource to advance research in this area.

ARTICLE INFORMATION

Received March 19, 2021; accepted July 15, 2021.

Affiliations

Department of Neurology, Yale School of Medicine, New Haven, CT (J.N.A., A.C.L., C.P.B., N.S., V.T.-L, S.A., J.S., N.P., L.S., K.N.S., G.J.F.); John A. Burns School of Medicine, University of Hawaii, Honolulu, HI (S.B.); and Department of Internal Medicine, Yale School of Medicine, New Haven, CT (T.M.G.).

Acknowledgments

The All of Us Research Program would not be possible without the partnership of its participants.

Sources of Funding

Dr Acosta is supported by the American Heart Association Bugher Fellowship in Hemorrhagic Stroke Research (817874). Ms Leasure is supported by the American Heart Association Medical Student Research Fellowship. Dr Gill is supported by the Yale Claude D. Pepper Older Americans Independence Center (P30AG021342). Dr Sheth is supported by the National Institutes of Health (R03NS112859, R01NS110721, R01NS075209, U01NS113445, U01NS106513, R01NR01833, U24NS107215, and U24NS107136) and the American Heart Association (17CSA33550004, 817874). Dr Falcone is supported by the National Institutes of Health (K76AG059992, R03NS112859 and P30AG021342), the American Heart Association (18IDDG34280056, 817874), the Yale Pepper Scholar Award (P30AG021342), and the Neurocritical Care Society Research Fellowship. The All of Us Research Program is supported by the National Institutes of Health, Office of the Director, Regional Medical Centers (1 OT2 OD026549, 1 OT2 OD026554, 1 OT2 OD026557, 1 OT2 OD026556, 1 OT2 OD026550, 1 OT2 OD 026552, 1 OT2 OD026553, 1 OT2 OD026548, 1 OT2 OD026551, 1 OT2 OD026555, IAA: AOD 16037) and federally qualified health centers (HHSN, 263201600085U; Data and Research Center, 5 U2C OD023196; Biobank, 1 U24 OD023121; The Participant Center, U24 OD023176; Participant Technology Systems Center, 1 U24 OD023163; Communications and Engagement, 3 OT2 OD023205; 3 OT2 OD023206; and Community Partners, 1 OT2 OD025277; 3 OT2 OD025315; 1 OT2 OD025337; 1 OT2 OD025276), https://aousupporthelp.zendesk.com/hc/en-us/articles/360040452471-How-do-I-cite-the-Researcher-Workbench-in-my-grants-or-publications-.

Disclosures

Dr Sheth reports grants from Hyperfine, grants from Bard, grants from Biogen, grants from Novartis, consultant pay from Ceribell, personal fees from Zoll, and equity from Alva unrelated to the submitted work. The remaining authors have no disclosures to report.

Supplementary Material

Tables S1–S3
Figure S1

REFERENCES

1. Virani SS, Alonso A, Aparicio HJ, Benjamin EJ, Bittencourt MS, Callaway CW, Carson AP, Chamberlain AM, Cheng S, Delling FN, et al. Heart disease and stroke statistics-2021 update: a report from the American Heart Association. *Circulation*. 2021;143:e254–e743. DOI: 10.1161/CIR.0000000000000950.
2. Youmans QR, Hastings-Spaine L, Princewill O, Shobayo T, Okwuosa IS. Disparities in cardiovascular care: past, present, and solutions. *Cleve Clin J Med*. 2019;86:621–632. DOI: 10.3949/ccjm.86a.18088.
3. Alvidrez J, Castille D, Laude-Sharp M, Rosario A, Tabor D. The National Institute on Minority Health and Health Disparities research framework. *Am J Public Health*. 2019;109:S16–S20. DOI: 10.2105/AJPH.2018.304883.
4. Abdalla SM, Yu S, Galea S. Trends in cardiovascular disease prevalence by income level in the United States. *JAMA Netw Open*. 2020;3:e2018150. DOI: 10.1001/jamanetworkopen.2020.18150.
5. Alzahrani T, Nguyen T, Ryan A, Dwairy A, McCaffrey J, Yunus R, Forgione J, Krepp J, Nagy C, Mazhari R, et al. Cardiovascular disease risk factors and myocardial infarction in the transgender population. *Circ Cardiovasc Qual Outcomes*. 2019;12:e005597. DOI: 10.1161/CIRCOUTCOMES.119.005597.
6. All of Us Research Program Investigators, Denny JC, Rutter JL, Goldstein DB, Philippakis A, Smoller JW, Jenkins G, Dishman E. The "All of Us" research program. *N Engl J Med*. 2019;381:668–676. DOI: 10.1056/NEJMsr1809937.
7. The All of Us Research Program Initial Protocol. National Institutes of Health (NIH)—All of Us. Published August 11, 2020. Available at: https://allofus.nih.gov/news-events-and-media/announcements/all-us-research-program-initial-protocol Accessed November 18, 2020.
8. Hripcsak G, Ryan PB, Duke JD, Shah NH, Park RW, Huser V, Suchard MA, Schuemie MJ, DeFalco FJ, Perotte A, et al. Characterizing treatment pathways at scale using the OHDSI network. *Proc Natl Acad Sci USA*. 2016;113:7329–7336. DOI: 10.1073/pnas.1510502113.
9. Roumie CL, Shirey-Rice J & Kripalani S MidSouth CDRN—Coronary Heart Disease Algorithm. PheKB. Available at: https://phekb.org/phenotype/midsouth-cdrn-coronary-heart-disease-algorithm Accessed October 26, 2020.
10. Woodfield R, Grant I; UK Biobank Stroke Outcomes Group, UK Biobank Follow-Up and Outcomes Working Group, Sudlow CLM. Accuracy of electronic health record data for identifying stroke cases in large-scale epidemiological studies: a systematic review from the UK Biobank Stroke Outcomes Group. *PLoS One*. 2015;10:e0140533. DOI: 10.1371/journal.pone.0140533.
11. Rannikmäe K, Ngoh K, Bush K, Al-Shahi Salman R, Doubal F, Flaig R, Henshall DE, Hutchison A, Nolan J, Osborne S, et al. Accuracy of identifying incident stroke cases from linked health care data in UK Biobank. *Neurology*. 2020;95:e697–e707. DOI: 10.1212/WNL.0000000000009924.
12. Graham G. Disparities in cardiovascular disease risk in the United States. *Curr Cardiol Rev*. 2015;11:238–245.
13. Bell CN, Thorpe RJ, Bowie JV, LaVeist TA. Race disparities in cardiovascular disease risk factors within socioeconomic status strata. *Ann Epidemiol*. 2018;28:147–152. DOI: 10.1016/j.annepidem.2017.12.007.
14. Yang Q, Zhong Y, Ritchey M, Cobain M, Gillespie C, Merrit R, Hong Y, George MG, Bowman BA. Vital signs: predicted heart age and racial disparities in heart age among U.S. adults at the state level. *MMWR Morb Mortal Wkly Rep*. 2015;64:950–958. DOI: 10.15585/mmwr.mm6434a6.
15. Cunningham TJ, Croft JB, Liu Y, Lu H, Eke PI, Giles WH. Vital signs: racial disparities in age-specific mortality among blacks or African Americans—United States, 1999–2015. *MMWR Morb Mortal Wkly Rep*. 2017;66:444–456. DOI: 10.15585/mmwr.mm6617e1.
16. Medina-Inojosa J, Jean N, Cortes-Bergoderi M, Lopez-Jimenez F. The Hispanic paradox in cardiovascular disease and total mortality. *Prog Cardiovasc Dis*. 2014;57:286–292. DOI: 10.1016/j.pcad.2014.09.001.
17. Villaroel MA, Blackwell DL & Jen A. Tables of summary health statistics for U.S. adults: 2018 National Health Interview Survey. Published online 2019. Available at: http://www.cdc.gov/nchs/nhis/SHS/tables.htm. Accessed February 1, 2021.
18. Caleyachetty R, Echouffo-Tcheugui JB, Muennig P, Zhu W, Muntner P, Shimbo D. Association between cumulative social risk and ideal cardiovascular health in US adults: NHANES 1999–2006. *Int J Cardiol*. 2015;191:296–300. DOI: 10.1016/j.ijcard.2015.05.007.

SUPPLEMENTAL MATERIAL

Table S1. Codes used to ascertain outcomes in the electronic health records (EHR).

Coding standard	Coronary heart disease*	Stroke
ICD-9 CM	410, 411, 412, 413, 414, 429.7, V45.81, V46.82	430, 431, 434.0, 434.1, 434.9, 436.X
ICD-9 procedure codes	36.01, 36.02, 36.03, 36.05, 36.09, 36.10, 36.12, 36.13, 36.14, 36.15, 36.16, 36.17, 36.18, 36.19	
ICD-10 CM	I20.0, I20.1, I20.8, I20.9, I21.09, I21.11, I21.19, I21.29, I21.3, I21.4, I23.0, I24.0, I24.1, I24.8, I25.10, I25.2, I25.3, I25.41, I25.42, I25.5, I25.810, I25.811, I25.812, I25.82, I25.83, I25.84, I25.89, I25.9, I51.0. I51.1, I51.2, I51.3, I51.4, I51.5, I51.7, I51.81, I51.89, I51.9, I97.0, I97.110, I97.130, I97.190, Z95.1, Z98.61	I60, I60.0 , I60.1, I60.2, I60.3, I60.4, I60.5, I60.6, I60.7, I60.8, I60.9, I61, I61.0 , I61.1, I61.2, I61.3, I61.4, I61.5, I61.6, I61.8, I61.9, I63, I63.0 , I63.1, I63.2, I63.3, I63.4, I63.5, I63.6, I63.8, I63.9, I63.X, I64, I64.X
ICD-10 procedure codes	0210093, 0210098, 0210099, 021009C, 021009F, 021009W, 02100A3, 02100A8, 02100A9, 02100AC, 02100AF, 02100AW, 02100J3, 02100J8, 02100J9, 02100JC, 02100JF, 02100JW, 02100K3, 02100K8, 02100K9, 02100KC, 02100KF, 02100KW, 02100Z3, 02100Z8, 02100Z9, 02100ZC, 02100ZF, 0210493, 0210498, 0210499, 021049C, 021049F, 021049W, 02104A3, 02104A8, 02104A9, 02104AC, 02104AF, 02104AW, 02104J3, 02104J8, 02104J9, 02104JC, 02104JF, 02104JW, 02104K3, 02104K8, 02104K9, 02104KC, 02104KF, 02104KW, 02104Z3, 02104Z8, 02104Z9, 02104ZC, 02104ZF, 0211098, 0211099, 021109C, 021109W, 02110A8, 02110A9, 02110AC, 02110AW, 02110J8, 02110J9, 02110JC, 02110JW, 02110K8, 02110K9, 02110KC, 02110KW, 02110Z8, 02110Z9, 02110ZC, 0211498, 0211499, 021149C, 021149W, 02114A8, 02114A9, 02114AC, 02114AW, 02114J8, 02114J9, 02114JC, 02114JW, 02114K8, 02114K9, 02114KC, 02114KW, 02114Z8, 02114Z9, 02114ZC, 021209C, 021209W, 02120AC, 02120AW, 02120JC, 02120JW, 02120KC, 02120KW, 02120ZC, 021249C, 021249W, 02124AC, 02124AW, 02124JC, 02124JW, 02124KC, 02124KW, 02124ZC, 021309C, 021309W, 02130AC, 02130AW, 02130JC, 02130JW, 02130KC, 02130KW, 02130ZC, 021349C, 021349W, 02134AC, 02134AW, 02134JC, 02134JW, 02134KC, 02134KW, 02134ZC, 02700ZZ, 02710ZZ, 02720ZZ, 02730ZZ, 02C00ZZ, 02C03ZZ, 02C04ZZ, 02C10ZZ, 02C13ZZ, 02C14ZZ, 02C20ZZ, 02C23ZZ, 02C24ZZ, 02C30ZZ, 02C33ZZ,02C34ZZ	

*Following the MidSouth Clinical Data Research Network Coronary Heart Disease Algorithm. Abbreviations: ICD = International Classification of Diseases.

Table S2. Cardiovascular risk factors stratified by sex at birth.

Group	Female sex at birth (n = 141,896)					Male sex at birth (n = 85,625)				
	Hypertension	Hyperlipidemia	Type 2 DM	Ever smoked	BMI	Hypertension	Hyperlipidemia	Type 2 DM	Ever smoked	BMI
All	57,252 (40.3)	52,851 (37.2)	21,727 (15.3)	48,525 (35.0)	30.37 (8.21)	40,300 (47.1)	37,858 (44.2)	15,664 (18.3)	41,256 (49.6)	29.07 (6.54)
Race/ethnicity										
White	29,327 (38.3)	32,387 (42.3)	8,673 (11.3)	29,556 (39.3)	29.20 (7.78)	24,446 (50.3)	25,695 (52.9)	8,189 (16.9)	22,540 (47.4)	29.15 (6.13)
Black	14,544 (56.7)	9,081 (35.4)	6,371 (24.8)	9,858 (39.9)	33.55 (9.14)	8,053 (47.5)	5,028 (29.6)	3,344 (19.7)	10,112 (62.2)	28.79 (7.43)
Hispanic/Latino/a/x	10,018 (34.6)	8,210 (28.4)	5,301 (18.3)	6,274 (22.2)	31.09 (7.52)	5,161 (39.4)	4,430 (33.8)	3,004 (22.9)	5,827 (45.9)	29.91 (6.79)
Asian	882 (21.3)	1,010 (24.4)	385 (9.3)	433 (10.7)	24.77 (5.40)	788 (30.4)	980 (37.8)	370 (14.3)	646 (25.4)	26.16 (4.78)
Other	796 (36.7)	729 (33.6)	356 (16.4)	783 (37.2)	30.29 (8.37)	709 (41.9)	700 (41.3)	303 (17.9)	820 (50.3)	28.54 (6.25)
More than one	727 (30.8)	626 (26.5)	252 (10.7)	825 (35.9)	29.97 (8.66)	378 (31.8)	374 (31.5)	142 (11.9)	511 (44.4)	28.27 (6.40)
Did not answer	958 (48.4)	808 (40.8)	389 (19.7)	796 (42.9)	30.67 (8.17)	765 (50.7)	651 (43.1)	312 (20.7)	800 (57.4)	28.65 (6.25)
Age										
<75 years	51,652 (38.5)	47,217 (35.2)	20,112 (15.0)	45,020 (34.4)	30.51 (8.30)	35,076 (44.6)	32,581 (41.5)	13,935 (17.7)	37,464 (49.1)	29.14 (6.66)
>75 years	5,600 (71.6)	5,634 (72.1)	1,615 (20.7)	3,505 (45.8)	28.03 (6.10)	5,224 (74.4)	5,277 (75.1)	1,729 (24.6)	3,792 (55.1)	28.33 (4.96)
Disability										
Without disability	47,304 (37.7)	45,071 (35.9)	16,543 (13.2)	41,366 (33.6)	29.89 (7.87)	34,720 (45.4)	33,556 (43.9)	12,801 (16.7)	35,920 (48.1)	28.91 (6.30)
With disability	9,186 (62.5)	7,159 (48.7)	4,834 (32.9)	6,705 (46.9)	34.30 (9.80)	4,909 (62.7)	3,741 (47.7)	2,563 (32.7)	4,747 (62.5)	30.76 (8.31)
Did not answer	762 (45.5)	621 (37.1)	350 (20.9)	454 (37.7)	30.71 (7.92)	671 (49.2)	561 (41.1)	300 (22.0)	589 (60.8)	28.54 (6.41)
Sex/Gender										
LGBTQIA+: No	51,901 (41.1)	48,364 (38.3)	19,469 (15.4)	42,259 (34.2)	30.28 (8.13)	35,471 (47.6)	33,520 (45.0)	13,772 (18.5)	35,826 (49.3)	29.14 (6.51)
LGBTQIA+: Yes	5,351 (34.2)	4,487 (28.7)	2,258 (14.4)	6,266 (41.6)	31.07 (8.81)	4,829 (43.6)	4,338 (39.2)	1,892 (17.1)	5,430 (51.1)	28.61 (6.70)
Education										
High school completed	49,818 (39.3)	47,279 (37.3)	17,737 (14.0)	42,616 (34.4)	30.15 (8.19)	35,497 (47.2)	34,407 (45.7)	13,220 (17.6)	34,614 (47.2)	29.15 (6.48)
Less than high school	6,305 (48.9)	4,727 (36.6)	3,492 (27.1)	5,128 (41.0)	32.15 (8.21)	3,854 (46.2)	2,763 (33.1)	2,043 (24.5)	5,532 (68.7)	28.57 (6.89)
Did not answer	1,129 (51.4)	845 (38.5)	498 (22.7)	781 (38.4)	31.34 (8.14)	949 (45.9)	688 (33.3)	401 (19.4)	1,110 (58.8)	28.34 (6.81)
Income										
Household income > 35 K	22,921 (34.3)	25,590 (38.3)	6,481 (9.7)	20,343 (31.0)	28.95 (7.43)	19,829 (47.9)	22,091 (53.4)	6,549 (15.8)	15,828 (39.0)	29.22 (5.85)
Household income <= 35 K	22,024 (46.4)	17,138 (36.1)	10,035 (21.1)	20,150 (43.6)	32.23 (8.96)	13,060 (45.3)	9,678 (33.6)	5,775 (20.0)	17,787 (63.7)	29.01 (7.24)
Did not answer	12,307 (44.5)	10,123 (36.6)	5,211 (18.8)	8,032 (30.0)	30.34 (7.87)	7,411 (48.0)	6,089 (39.4)	3,340 (21.6)	7,641 (51.7)	28.82 (6.75)

Columns for risk factors present the prevalence of each risk factor in each subgroup. Abbreviations: DM = Diabetes mellitus. BMI = body mass index. LGBTQIA+ = lesbian, gay, bisexual, transgender, queer, intersex and asexual.

Table S3. Adjusted odds of cardiovascular disease in underrepresented groups by sex at birth.

Group	Female sex at birth		Male sex at birth		Interaction p value
	Univariable regression OR (95%CI)	Multivariable regression OR (95%CI)	Univariable regression OR (95%CI)	Multivariable regression OR (95%CI)	
Race/Ethnicity					
White	Reference	Reference	Reference	Reference	
Asian	0.39 (0.31-0.47)	0.73 (0.58-0.91)	0.47 (0.40-0.54)	0.95 (0.79-1.13)	
Hispanic/Latino/a/x	0.64 (0.60-0.68)	0.92 (0.85-0.99)	0.44 (0.41-0.47)	0.78 (0.71-0.85)	<0.001
Other	0.94 (0.78-1.12)	1.15 (0.93-1.40)	0.82 (0.70-0.96)	1.26 (1.05-1.50)	
More than one	0.62 (0.50-0.77)	0.99 (0.77-1.24)	0.48 (0.38-0.60)	1.06 (0.81-1.37)	
Black	1.42 (1.35-1.50)	1.40 (1.32-1.49)	0.59 (0.56-0.63)	1.02 (0.95-1.10)	
Age					
< 75 years	Reference	Reference	Reference	Reference	<0.001
>= 75 years	3.46 (3.25-3.69)	1.69 (1.57-1.81)	3.90 (3.68-4.13)	2.11 (1.98-2.25)	
Disability					
No	Reference	Reference	Reference	Reference	<0.001
Yes	2.70 (2.55-2.85)	1.82 (1.71-1.94)	1.51 (1.41-1.61)	1.37 (1.27-1.48)	
Sex/Gender					
LGBTQIA+: No	Reference	Reference	Reference	Reference	0.07
LGBTQIA+: Yes	0.75 (0.69-0.81)	1.03 (0.94-1.12)	0.78 (0.73-0.84)	0.99 (0.91-1.07)	
Education					
>= High school degree	Reference	Reference	Reference	Reference	0.001
< High school degree	1.17 (1.08-1.25)	0.98 (0.90-1.06)	0.67 (0.61-0.72)	0.84 (0.76-0.92)	
Income*					
> USD 35k	Reference	Reference	Reference	Reference	<0.001
< USD 35k	1.53 (1.45-1.60)	1.34 (1.26-1.41)	0.74 (0.70-0.78)	1.16 (1.09-1.23)	

SD = standard deviation. 95%CI = 95% confidence interval. OR = odds ratio. USD 35k = 35 thousand American dollars. LGBTQIA+ = lesbian, gay, bisexual, transgender, queer, intersex and asexual. *Income corresponds to annual household income.

Figure S1. Definitions of underrepresented groups used in the All of Us Research.