Prevalence of Risk Factors and Established Cardiovascular Disease Among All of Us Participants: Benchmarking Against National Estimates

Mengjing Yang MPH[1], Tinsae Admassu MD MPH[1], Harlan M Krumholz MD SM[1,2,3], Chenxi Huang PhD[2,3], Karthik Murugiah MBBS MHS [2,3]

[1]Yale School of Public Health, New Haven, CT
[2]Section of Cardiovascular Medicine, Department of Internal Medicine, Yale School of Medicine, New Haven, CT
[3]Center for Outcomes Research and Evaluation, Yale-New Haven Hospital, New Haven, CT

Correspondence: Karthik Murugiah MD, 195 Church St, 6th Floor, New Haven, CT 06510; 203-487-1239 (tel). 203-764-5653 (fax).

Funding: Dr Murugiah received support from the National Heart, Lung, and Blood Institute of the National Institutes of Health (under award K08HL157727).

Disclosures: Dr Krumholz has received options for Element Science and Identifeye and payments from F-Prime for advisory roles; he is a co-founder of and holds equity in Hugo Health, Refactor Health, and ENSIGHT-AI; and he is associated with research contracts through Yale University from Janssen, Kenvue, Novartis, and Pfizer.

Abstract Word Count: 260

Manuscript Word Count: 2288

ABSTRACT

Introduction:
Knowing how demographics, cardiovascular risk factors and established cardiovascular disease (CVD) of All of Us (AoU) Research Program participants differ from the U.S population is important for cardiovascular researchers using these data, and to inform the generalizability of findings.

Methods:
We included AoU participants ≥18 years enrolled between May 2017-June 2022. 'AoU Survey' cohort included participants who completed the Personal and Family Health History survey. 'AoU EHR' included participants with linked EHR data. Cardiovascular risk factors and established CVD overall and by demographic sub-groups were compared to a weighted sample of adults ≥18 years from the National Health and Nutrition Examination Survey (NHANES) 2017-March 2020 cycle. Cardiovascular risk factors and established CVD were identified through survey questions in 'AoU Survey' and NHANES, and through SNOMED codes for 'AoU EHR'.

Results:
'AoU Survey' (N=185,232) was older than NHANES (N=9,683 representing 248 million) (56±17 vs 48±18 years), with more female (60.4% vs 51.8%) and Black (18.6% vs 11.5%, all P <0.001) participants. Prevalence of cardiovascular comorbidities in 'AoU Survey' was similar to NHANES overall, but lower when age stratified. However, this pattern was not observed among Black participants, who exhibited higher hypertension (42.2% vs 39%), diabetes (19.2% vs 12.5%), and congestive heart failure (3.8% vs 3.6%) rates than the U.S Black population. In contrast, female participants showed lower cardiovascular comorbidities. 'AoU EHR' with its code-based identification had higher prevalence of cardiovascular comorbidities.

Conclusion:
AoU includes more historically underrepresented groups, as intended. However, researchers should account for differences in disease prevalence and subgroup characteristics with the U.S. population when interpreting survey estimates.

INTRODUCTION

The All of Us (AoU) Research Program, launched by the National Institutes of Health (NIH), represents a transformative investment in precision medicine, assembling one of the largest, most diverse health cohorts ever created in the United States.(1) By integrating electronic health records (EHRs), detailed health surveys, biospecimens, physical measurements, and wearable-device data from over one million participants, particularly from those racial, ethnic, and socioeconomic groups historically underrepresented in biomedical research, AoU offers a unique lens into population health and disease mechanisms. In cardiovascular research, this unprecedented breadth and depth of data promises new insights into risk factors, disease progression, and treatment responses across diverse communities.

However, as AoU enrollment is based on convenience and extensive outreach rather than probability sampling, its participant pool can differ systematically from the general U.S. population. For investigators seeking to translate AoU findings into generalizable evidence, understanding how cardiovascular risk profiles and disease prevalence in AoU align with or diverge from national benchmarks is essential.

To address this need, we compare the demographic composition and the prevalence of major cardiovascular (CV) risk factors and established cardiovascular diseases (CVD) in AoU participants against contemporaneous estimates from the nationally representative National Health and Nutrition Examination Survey (NHANES).(2) By aligning AoU with a nationally representative reference and examining multiple cardiometabolic conditions across both survey-reported and EHR-derived data, our analysis will clarify the representativeness and supplies a critical context for interpreting future cardiovascular investigations that draw on this landmark resource.

METHODS

Data Sources and Study population

AoU Research Program:

We used the Controlled Tier Dataset v7 and included participants ≥18 years of age enrolled between May 31, 2017, and June 30, 2022. Comorbidities were identified either through self-reported diagnoses in the Personal Health and Family History survey or via the diagnosis codes in linked EHR data. To reflect these data sources, we constructed the following three cohorts based on data availability: 1) ‘AoU Overall’ cohort: all

participants regardless of survey or EHR availability. 2) ‘AoU Survey’ cohort: participants who had filled out the personal and family health history survey; 3) ‘AoU EHR’ cohort: participants with linked EHR data. Each study participant granted written informed consent, and the study is overseen by the AoU Institutional Review Board.

NHANES:

The NHANES is an ongoing program conducted by the Centers for Disease Control and Prevention’s National Center for Health Statistics, designed to assess the health and nutrition status of adults and children in the United States. Employing a complex, multistage, stratified and clustered sampling design, NHANES selects a representative sample of individuals annually from the U.S. population. Demographics, questionnaires, examination, and laboratory data are linked by unique participant identifiers (SEQN), and all analyses incorporate the provided sample weights and design variables to account for unequal selection probabilities and non-response. For this study, we analyzed data from the 2017-March 2020 cycle, including individuals ≥18 years of age who completed the following selected questionnaires: demographic questionnaire (DEMO), smoking questionnaire (SMQ), medical conditions (MCQ), blood pressure questionnaire (BPQ), kidney conditions questionnaire (KIQ), occupation questionnaire (OCQ), health insurance questionnaire (HIQ), diabetes questionnaire (DIQ), and health status questionnaire (HUQ). Use of NHANES data for this study received an exemption for review from the Institutional Review Board at Yale University because the NHANES data are publicly available and de-identified.

Patient characteristics

For AoU, self-reported demographic information, including gender, race/ethnicity, employment status, income, insurance, etc., was obtained through the "Basic" survey. We examined eight cardiovascular conditions, including hypertension, hyperlipidemia, diabetes mellitus, coronary artery disease (CAD), acute myocardial infarction (AMI), congestive heart failure (CHF), cerebrovascular accident/transient ischemic attack (CVA/TIA), and chronic kidney disease (CKD),. In the AoU survey cohort, a condition was considered "Present" if participants selected ‘Self’ to questions about their personal history in the 'Personal/Family Health History’ survey, “Absent” if they selected “other family members”, “Did not answer” if they selected ‘Skip’, and “Missing” if no response was provided. In the AoU EHR cohort, a condition was classified as “present” if at least one corresponding Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) code appeared in the participant's medical records prior to the date of the survey. If no such code was found, the condition was classified as absent. The full list of

SNOMED CT codes used to identify each condition is provided in Supplementary Table 1.

For NHANES, sociodemographic data were collected through structured interviews, encompassing demographic variables such as age, gender, race, ethnicity, and socioeconomic status. CV comorbidities were defined by self-reported diagnoses, using targeted survey questions: CHF (MCQ.160b), CAD (MCQ.160c), angina (MCQ.160d), AMI (MCQ.160e), and CVA/TIA (MCQ.160f). CKD and dialysis status were assessed using KIQ.022 and KIQ.025, respectively, while diabetes, hypertension, and hyperlipidemia were identified using DIQ.010, BPQ.020, and BPQ.080. Responses were categorized as "Present" for "Yes", "Absent" for "No", “Did not answer” for responses such as “Refused” or “Don’t know”, and “Missing” when no response was recorded.

Statistical analysis

We conducted all analyses in R (version 4.4.0), using the survey package to account for NHANES’s complex design. For the AoU dataset, which employs a non-probabilistic, unweighted design, we estimated unweighted means, proportions, and 95% confidence intervals. For NHANES, we applied the provided sample weights and design variables to generate nationally representative estimates with corresponding variance estimates.

We first compared key sociodemographic characteristics, including age, sex, race and ethnicity, income, and insurance status, across NHANES and all three AoU cohorts (AoU Overall, AoU Survey, and AoU EHR) using two-sample t-tests for continuous variables and chi-square tests for categorical variables. For prevalence rates of CV comorbidities, we restricted comparisons to NHANES versus the AoU Survey and AoU EHR cohorts. Group differences were further examined within subgroups defined by age (18-40, 40-65 and >65 years), sex, and race/ethnicity using the same test framework. We considered two-sided p-values <0.05 statistically significant, noting that, given large sample sizes and multiple comparisons, these tests are intended to highlight patterns rather than serve as definitive hypothesis tests.

RESULTS

The ‘Overall AoU’ cohort included 413,457 participants (Figure 1). Of these, 185,232 (44.8%) participants completed the Person and Family Health Survey data (‘AoU Survey’ cohort) and 287,012 (69.4%) had linked their EHR records (‘AoU EHR’ cohort). The NHAHES cohort from the 2017-March 2020 cycle with data on the selected

questionnaires included 9,683 participants representing an estimated 248 million adult US population after weighting.

Compared to NHANES, AoU participants were older (mean±SD: 56±17 years vs 48±18 years), more often female (60.4% vs 51.8%) and included a higher proportion of Black participants (18.6% vs 11.5%, all P <0.001, Table 1). Within AoU Survey, these age differences persisted across demographic subgroups. The mean age of male participants was 62.2 years versus 46.7 years for the male U.S population in NHANES, while the mean age of AoU Survey female participants was 56.5 years versus 48.3 years in NHANES. The mean age of White participants in AoU Survey was 61 years versus 49.8 years for the White U.S population in NHANES, while the mean age of AoU Survey Black participants was 56.5 years versus 45.3 years in NHANES; all P <0.001.

AoU Survey participants differed from AoU EHR participants. They were marginally older and more often female, but had much higher proportions of White (69.7% vs 52.4%), better educated (83.5% vs 67.3% with >12 years education), and more frequently employed (53.3% vs 45.8%, Table 1) participants.

The overall prevalence of CV risk factors and established CVD is shown in Figure 2 and Supplementary Table 2. Overall, self-reported prevalence in the AoU Survey cohort were similar to or marginally lower than NHANES estimates. In contrast, the AoU EHR cohort had a significantly higher prevalence of all comorbidities except for CVA/TIA, which was lower than in NHANES.

The prevalence of CV risk factors and established CVD stratified by age, sex and race/ethnicity is shown in Figures 3 and 4 and Supplementary Table 3. Across age groups, the prevalence of CV risk factors and established CVD was lower in the AoU Survey cohort compared to NHANES, and this pattern was consistent across most comorbidities. A similar pattern was also observed among women. Stratified by race, the prevalence of CV risk factors and established CVD was consistently lower among White participants in the AoU Survey cohort when compared to the White population in NHANES, despite a higher mean age of White participants in AoU Survey. However, Black participants deviated from this pattern, with higher rates of hypertension, diabetes, CHF, but lower prevalence of CKD.

DISCUSSION

In this study comparing AoU participants to the US adult population we noted that the AoU participants were older and had a higher proportion of women and Black adults. Overall prevalence of cardiovascular risk factors and established disease in the AoU

Survey cohort closely matched national estimates; however, stratified analyses revealed consistently lower comorbidity rates across all age groups. Notably, women in AoU reported fewer cardiovascular conditions than their U.S. counterparts, whereas Black participants exhibited higher rates. Within AoU, the AoU EHR cohort, which had a lower educational attainment and employment, and a higher proportion of Black participants, demonstrated a markedly higher comorbidity prevalence than the survey cohort.

The AoU program has an important focus on recruiting participants from populations historically underrepresented in research, and the efforts of the program are reflected by the higher proportion of female and Black participants compared with national benchmarks. However, other key groups, such as Asians or those with Hispanic ethnicity remain under-represented, highlighting the need for more targeted outreach to ensure inclusive participation.

There were notable differences between AoU participants who had filled out the health survey compared with participants with EHR linkage. The AoU Survey cohort, compared with the AoU EHR cohort, had higher educational attainment and employment levels and a lower proportion of Black participants. In fact, the AoU survey cohort had a lower proportion of Black participants than the US population. This gap underscores the persistent socioeconomic and logistic barriers to completing surveys which limit our ability to obtain self-reported health behaviors and health experiences of underrepresented populations.(3) The AoU program used instruments and questions tested in diverse populations and targeted a fifth grade reading level.(1) However, clearly efforts beyond improving the ability to interpret survey questions are needed to enhance diverse representation of survey data.

The AoU Survey cohort's overall prevalence of cardiovascular risk factors and established CVD closely mirrored the NHANES' survey-based estimates. However, when stratified by age, AoU participants consistently reported slightly lower rates of comorbidities across every age group, suggesting a modest "healthy volunteer" effect compared with the broader U.S. population, which has been noted in other similar large cohorts such as the UK Biobank.(4) Our findings though contrast with results from two recent studies. One study compared the disease prevalence in AoU to Global Burden of Diseases (GBD) data for the US showed a higher prevalence of comorbidities in AoU. However, that analysis included multisystem conditions and was not focused on CVD, further, conditions were defined via ICD codes, which limits direct comparability.(5) Similarly, another study used NHANES survey identified comorbidities and compared it with EHR identified comorbidities in AoU data and reported a higher comorbidity burden in AoU.(6) In contrast in our study we made comparisons using survey-based comorbidity estimates in both AoU and NHANES. Further, the NHANES time frame

used in this prior study was not optimally aligned to the time frame of enrollment in AoU which further limits the comparison. Using survey based estimates we demonstrate that the absolute CVD burden within AoU is slightly lower than the US population. Importantly though, despite this slightly lower CVD burden it remains substantial enough to support in depth cardiovascular research.

The finding of lower CV comorbidity burden, however, was not consistent across demographic subgroups, and some patterns were noted. Black participants in AoU had generally a higher prevalence of comorbidities than the Black US population, while White participants had a lower prevalence despite being older. Similarly, female participants in AoU had a lower prevalence of comorbidities than the female US population while male participants had higher prevalence. These findings underscore the need for caution when using AoU survey data to study racial or sex specific cardiovascular risks.

Prevalence of CV comorbidities in the EHR cohort as identified by SNOMED codes was substantially higher than the survey cohort. It likely reflects a combination of true differences in the populations captured as well as measurement disparities. EHR linked participants were demographically distinct (e.g., lower education, higher Black representation) and may have more frequent healthcare encounters, leading to greater diagnostic capture. Conversely, discrepancies between self-report and clinical records are well documented.(7) EHR and self-report both can be incomplete, the former due to incorrect coding/documentation or issues with EHR interoperability, and the latter due to limitations in health literacy or recall bias. Recognizing these complementary strengths and limitations is essential when selecting data sources for cardiovascular research with AoU.

Our study has limitations. It is cross-sectional and relies on a single survey timepoint and EHR snapshot, precluding assessment of how cardiovascular status evolves over time. Misclassification may arise from self-report inaccuracies and variability in diagnostic coding across EHR systems. The non-probabilistic, convenience-sampling design of AoU limits direct comparability with NHANES and may introduce selection bias despite our subgroup analyses. Finally, unmeasured factors such as access to care and health behaviors are not captured in our datasets and could confound observed differences in disease prevalence. Nevertheless, our findings highlight critical considerations for future research using AoU data, underscoring the need to account for sampling design, data source heterogeneity, and subgroup-specific patterns when investigating cardiovascular outcomes.

CONCLUSION

AoU's enriched enrollment of underrepresented groups makes it a powerful resource for cardiovascular research, but our comparison with NHANES reveals systematic differences in disease prevalence—both overall and within key demographic subgroups. Researchers should therefore consider these cohort and measurement differences when designing studies and interpreting findings to strengthen their generalizability.

ACKNOWLEDGEMENTS

We gratefully acknowledge All of Us participants for their contributions, without whom this research would not have been possible. We also thank the National Institutes of Health's All of Us Research Program for making available the participant data examined in this study.

TABLES AND FIGURES

Table 1. Comparison of Socio-Demographic Factors between NHANES and AoU cohorts

Socio-demographic status	**NHANES Weighted**		**AoU General (N = 413,457)**		**AoU Person and Family Survey (N = 185,232)**		**AoU with EHR (N = 287,012)**	
	N	**%**	**N**	**%**	**N**	**%**	**N**	**%**
Age (years±SD)	47.49 ± 17.87		56.06±17.11		58.14±17.14		57.63± 16.96	
Age Group								
18-40	94125616	37.96	91731	22.19	37045	20.00	57725	20.11
40-65	103189321	41.62	174886	42.30	71991	38.87	120291	41.91
>=65	50632123	20.42	146840	35.52	76196	41.14	108996	37.98
Sex								
Male	119592918	48.23	155169	37.53	62560	33.77	108739	37.89
Female	128354142	51.77	249565	60.36	118491	63.97	172401	60.07
Did not answer	0	0.00	4439	1.07	1237	0.67	3169	1.10
Missing	0	0.00	4284	1.04	2944	1.59	2703	0.67
Race/ Ethnicity								
White	154162965	62.18	222646	53.85	129090	69.69	150402	52.40
Black	28420787	11.46	77069	18.64	16918	9.13	57177	19.92
Asian	14781102	5.96	13838	3.35	6086	3.29	8038	2.80
Hispanic/Latino/a/x	40485465	16.33	64672	15.64	17236	9.31	47699	16.62
Other/ >1	10096742	4.07	23443	5.67	10266	5.54	15690	5.47
Did not answer	0	0.00	11692	2.83	5636	3.04	8005	2.79
Missing	0	0.00	97	0.02	0	0.00	1	0.00
Marital Status								
Married	148072184	59.72	204011	49.34	107790	58.19	139534	48.62
Widowed, Separated or divorced	45281859	18.26	92150	22.29	35758	19.30	67086	23.37
Never married	47000622	18.96	102480	24.79	36484	19.70	70297	24.49
Did not answer	94828	0.04	14719	3.56	5200	2.81	10094	3.52
Missing	7497567	3.02	97	0.02	0	0.00	1	0.00

Education (years)								
<12 y	26538169	10.70	36432	8.81	5976	3.23	27867	9.71
12 y	64776211	26.13	77014	18.63	19654	10.61	56778	19.78
>12 y	148884269	60.05	286481	69.29	154659	83.49	193184	67.31
Did not answer	2508442	0.10	13433	3.25	4943	2.67	9182	3.20
Missing	7497567	3.02	97	0.00	0	0.00	1	0.00
Self-rated health								
Good	201938577	81.44	293368	70.95	144907	78.23	205987	71.77
Fair	39302718	15.85	81000	19.59	31524	17.02	60044	20.92
Poor	6419188	2.59	20555	4.97	7352	3.97	15358	5.35
Did not answer	286576	0.12	17288	4.18	1449	0.78	5467	1.90
Missing	0	0.00	1246	0.30	0	0.00	156	0.05
Health Insurance								
Covered	214599200	86.55	371633	89.88	173855	93.86	260359	90.71
Not covered	32892008	13.27	26779	6.48	6058	3.27	16733	5.83
Did not answer	455852	0.18	13158	3.18	3690	1.99	8685	3.03
Missing	0	0.00	1887	0.46	1629	0.88	1235	0.43
Smoking								
Yes	102779026	41.45	155901	37.71	65107	22.68	114595	39.93
No	145104923	58.52	232630	56.26	116753	40.68	164021	57.15
Did not answer	63111	0.03	23680	5.73	3372	1.17	8240	2.87
Missing	0	0.00	1246	0.3	0	0.00	156	0.05
Employment status								
Employed	158614770	63.97	199392	48.23	98660	53.26	131320	45.75
Unemployed	89276381	36.01	197741	47.83	81166	43.82	144650	50.40
Did not answer	55909	0.02	16227	3.92	5406	2.92	11041	3.85
Missing	0	0.00	97	0.02	0	0.00	1	0.00

Figure 1: Venn diagram representing the overlap of participants with Person and Family Survey data and EHR data

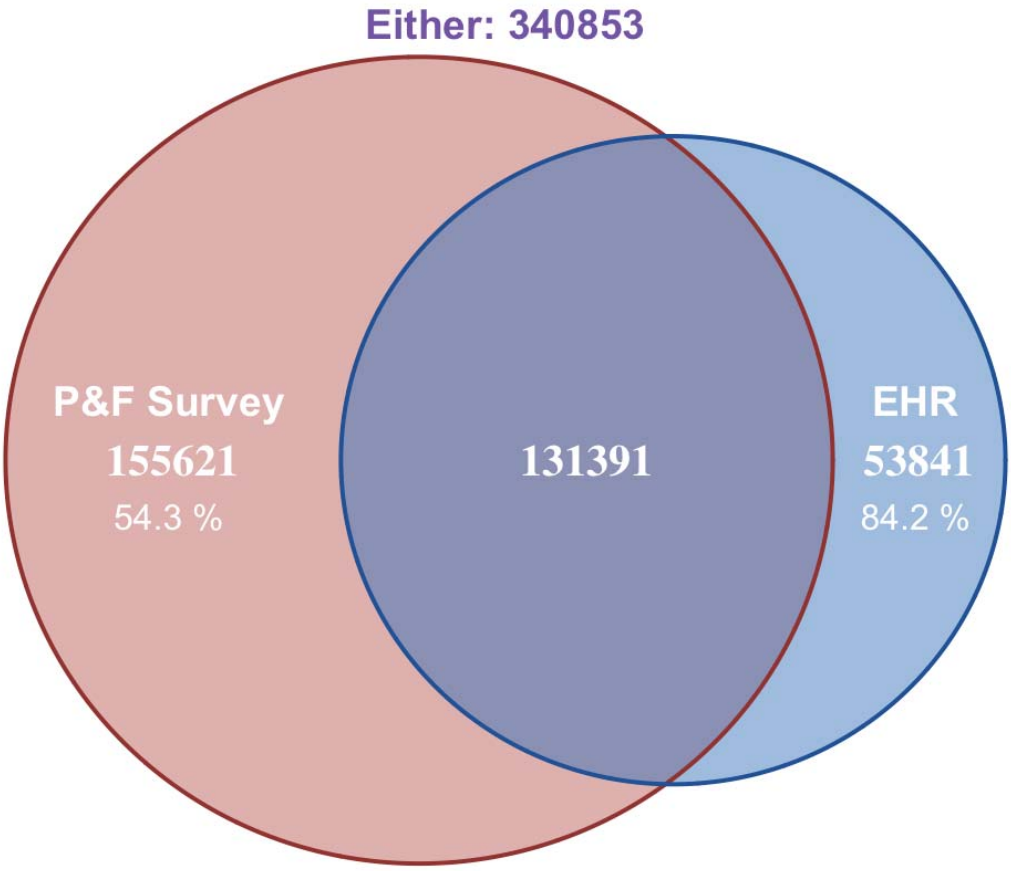
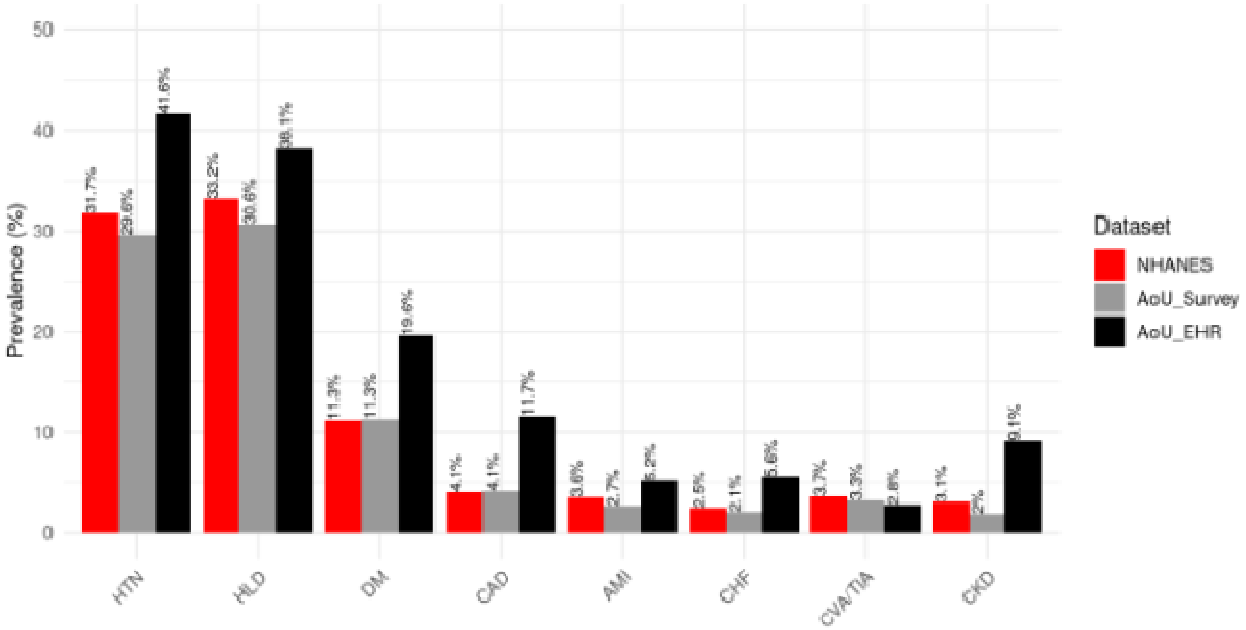
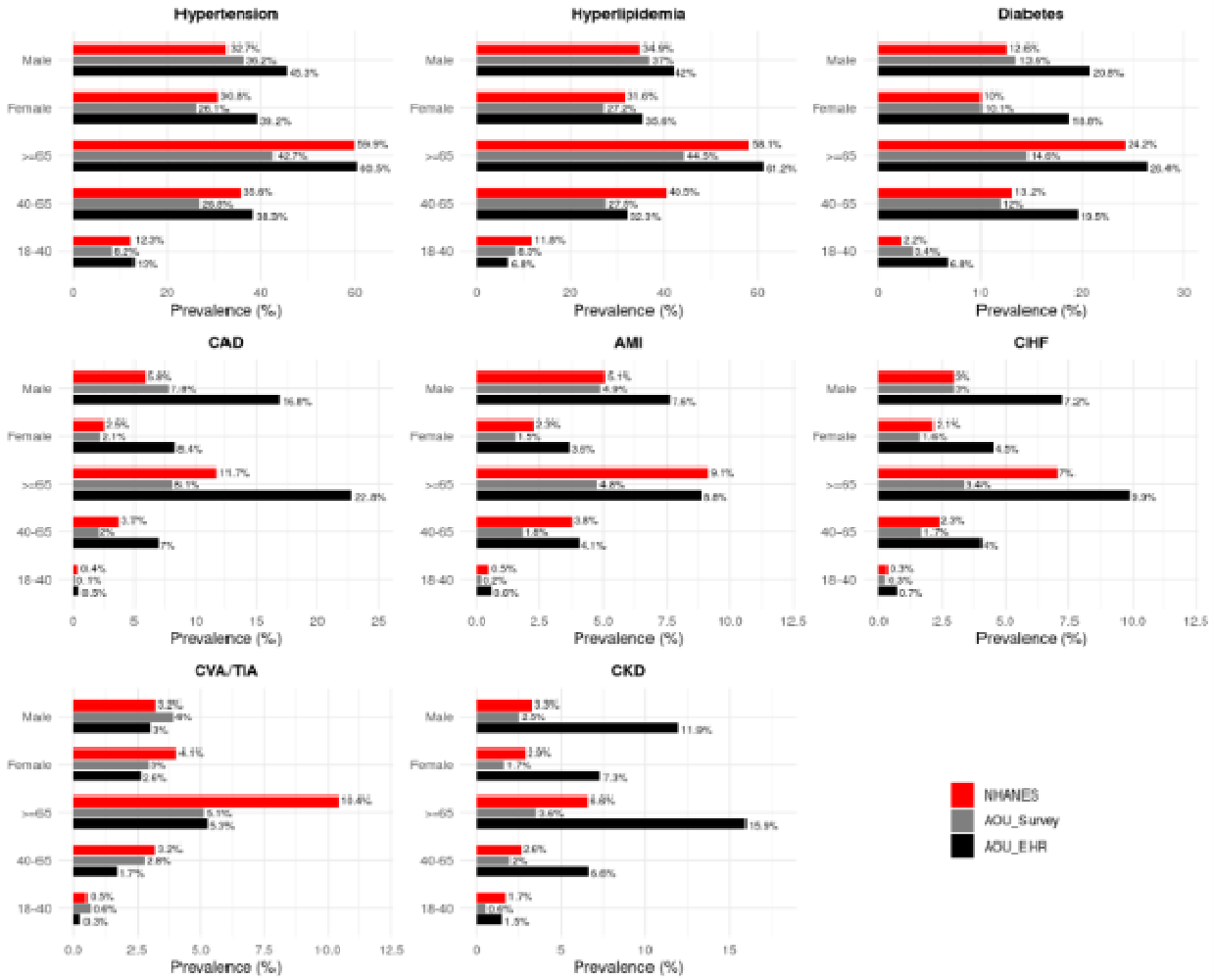


Figure 2: Prevalence of CV risk factors and established CVD



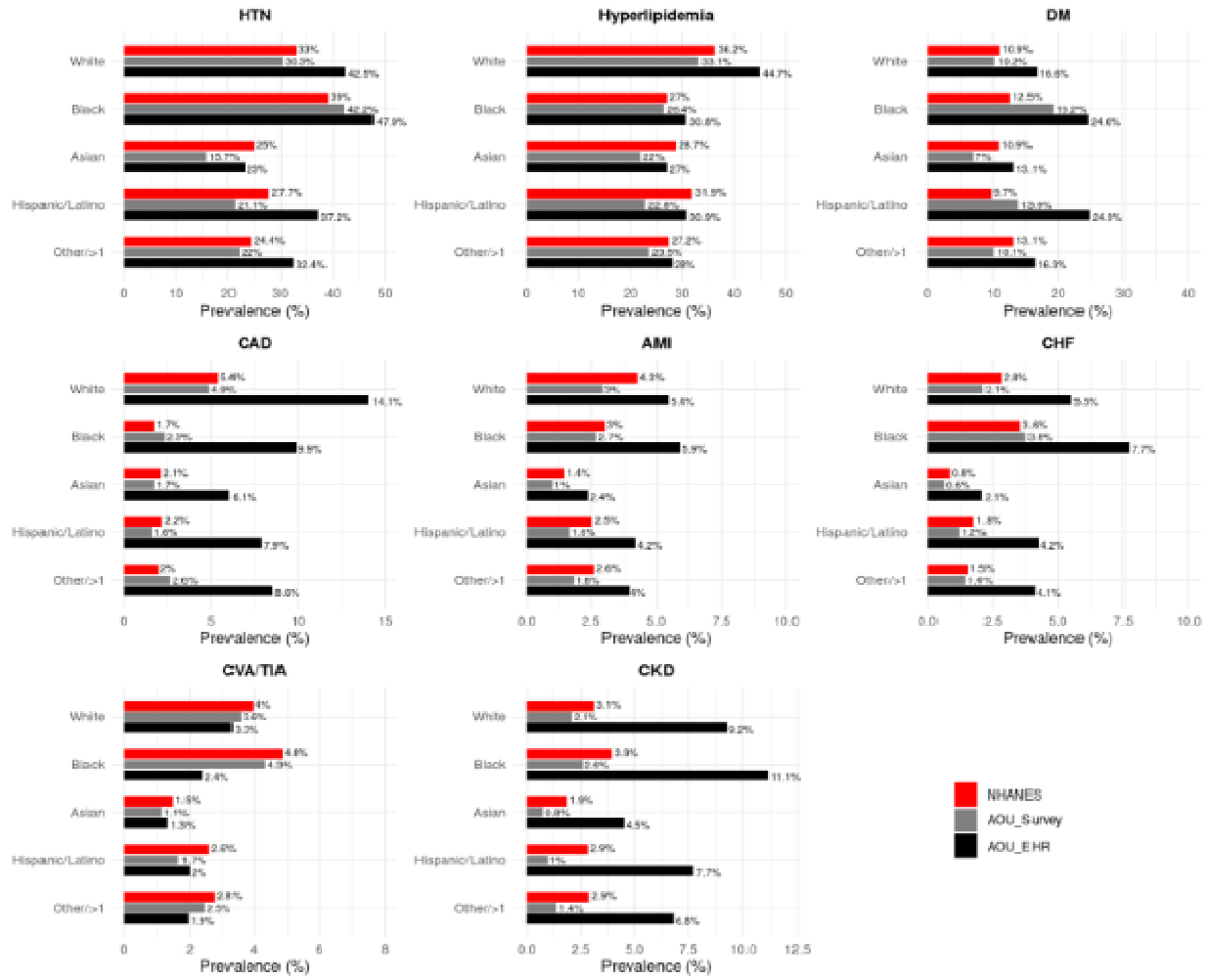
This figure compares the estimated prevalence of major cardiometabolic conditions including acute myocardial infarction (AMI), coronary artery disease (CAD), congestive heart failure (CHF), chronic kidney disease (CKD), cerebrovascular accident/transient ischemic attack (CVA/TIA), diabetes mellitus (DM), hyperlipidemia (HLD), and hypertension (HTN) across three data sources: NHANES (self-report and exam-based), All of Us (AoU) EHR, and AoU survey responses. EHR-based prevalence estimates were consistently higher than survey-based or NHANES estimates for most conditions, particularly for AMI, CAD, and CKD.

Figure 3: Prevalence of CV risk factors and established CVD by age and sex



This figure presents the prevalence of eight cardiometabolic conditions by demographic subgroups (age groups: 18–40, 40–65, ≥65; and sex: Male, Female) across three datasets. EHR-based data from All of Us generally show higher prevalence rates than self-reported survey responses and NHANES, with the largest age-associated increases observed for hypertension, diabetes, and CAD. The sex-specific trends vary by condition and data source.

Figure 4: Prevalence of CV risk factors and established CVD by race/ethnicity



This figure shows race/ethnicity-stratified prevalence for key cardiovascular and metabolic conditions, comparing NHANES and both EHR and survey data from All of Us. Disparities in disease burden are evident across racial/ethnic groups, with Black participants generally exhibiting higher prevalence of hypertension and diabetes across all datasets, while Asian participants show lower rates for most conditions.

Supplementary Table 1: SNOMED CT codes used to identify each disease condition

Condition	**SNOMED ID**
Hypertension	194791005, "10725009", "10757401000119104", "118781000119108", "1201005",
	"123799005", "132721000119104", "140121000119100", "16229371000119106", "194783001",
	"194785008", "194788005", "198941007", "198942000", "198944004", "198945003",
	"198946002", "198947006", "198949009", "198965005", "198966006", "198967002",
	"198983002", "198984008", "198985009", "198986005", "198997005", "198999008",
	"199000005", "199002002", "199005000", "199007008", "199008003", "206596003",
	"23130000", "237279007", "28119000", "288250001", "307632004", "31992008",
	"367390009", "371125006", "37618003", "38341003", "397748008", "398254007",
	"40521000119100", "41114007", "429457004", "443482000", "461301000124109",
	"46764007", "48146000", "48194001", "52698002", "541000119105", "5501000119106",
	"56218007", "57684003", "59621000", "63287004", "65518004", "67359005",
	"697929007", "697930002", "70272006", "706882009", "712832005", "71421000119105",
	"72022006", "73410007", "765182005", "78808002", "78975002", "8218002",
	"82771000119102", "86041002", "8762007", "89242004", "95605009"
Hyperlipidemia	55822004, "129589009", "299465007", "13644009","238076009",
	"398036000","238079002","238080004","267432004","302870006" ,
	"238083002","34528009","267435002","34349009","238085009",
	"402473001","267433009","129591001","267434003","238088006" ,
	"238040008","238089003","1571000119104","701000119103","402727002",
	"426161002","137941000119106","137931000119102"
Diabetes	“5368009”, "5969009", "8801005", "11687002", "23045005",
	"28032008", "44054006", "46635009", "46894009", "49817004",

	"51002006", "59079001", "73211009", "75022004", "76751001",
	"81531005", "91352004", "111552007", "190368000", "190389009",
	"190447002", "199223000", "199225007", "199226008", "199227004",
	"199228009", "199229001", "199230006", "237599002", "237601000",
	"237627000", "313435000", "313436004", "359642000", "408540003",
	"413183008", "426705001", "426875007", "427089005", "445260006",
	"609561005", "609562003", "609563008", "609564002", "609566000",
	"609567009", "609568004", "609570008", "703136005", "703138006",
	"737212004", "1481000119100", "31321000119102", "40791000119105",
	"40801000119106", "105401000119101", "106281000119103",
	"10753491000119101", "10754881000119104"
CAD	53741008, "371804009", "429673002", "11018701000119109", "139011000119104",
	"15960061000119102", "15960141000119102", "16891151000119103", "194842008",
	"194843003", "233817007", "285141000119106", "285151000119108", "371803003",
	"371805005", "429245005", "442224005", "442240008", "442421004", "443502000",
	"444855007", "451041000124103", "451361000124102", "719678003", "724431008",
	"791000119109", "92517006"
AMI	22298006, "57054005", "401303003", "703164000", "15713081000119108",
	"15713121000119105", "16837681000119104", "17531000119105", "1755008",
	"194802003", "194856005", "194857001", "194858006", "23311000119105",
	"233835003", "233837006", "233840006", "285981000119103", "307140009",
	"314207007", "380001000004106", "401314000", "418044006", "54329005",
	"58612006", "59063002", "65547006", "70211005", "703164000",
	"703211006", "703360004", "70422006", "73795002", "76593002",
	"79009004", "836294006"
CHF	42343007, "10633002", "194767001", "92506005", "88805009",

	"426263006", "426611007", "67441000119101", "23341000119109",
	"82523003", "96311000119109", "194779001", "194781004",
	"5148006", "698594003", "698296002", "66989003",
	"15629541000119106", "15781000119107", "83105008",
	"7401000175100", "16838951000119100"
CVA/TIA	10349009, "14070001", "25772007", "56267009", "57981008",
	"70936005", "195213000", "195217004", "230713003", "230715005",
	"371040005", "371121002", "16371781000119100", "16000351000119109",
	"16000471000119107", "16000511000119103", "16002031000119102",
	"16002071000119104", "16002111000119106", "16002151000119100",
	"16002231000119106", "16002271000119109", "16002351000119105",
	"16002391000119100", "16002431000119105", "16002471000119108",
	"16002511000119104", "16023911000119108", "16024111000119109",
	"16026951000119102", "195212005", "195216008", "230690007",
	"230691006", "281240008", "292671000119104", "292681000119101",
	"292691000119103", "329401000119103", "329621000119105",
	"329641000119104", "329651000119102", "329671000119106",
	"330791000119108", "371041009", "373606000", "39925003",
	"413758000", "422504002", "434141000124103", "457551000124104",
	"5571000124103", "716051003", "78569004", "9901000119100",
	"99451000119105", "12237911000119109", "12237951000119105",
	"195200006", "195209007", "195210002", "195211003",
	"230716006", "230717002", "266257000", "64009001",
	"710575003", "88032003"
CKD	709044004, "46177005", "49708008", "90688005", "236433006",
	"236436003", "431855005", "431856006", "431857002", "433144002",
	"433146000", "700378005", "700379002", "713313000", "714152005",
	"714153000", "723190009", "711000119100", "721000119107",
	"731000119105", "741000119101", "771000119108", "90741000119107",

	"90761000119106", "90791000119104", "96441000119101",
	"129151000119102", "129161000119100", "129181000119109",
	"691421000119108", "104931000119100"

Supplementary Table 2. Prevalence of CV risk factors and established CVD

	NHANES Weighted		**AoU Person and Family Survey (N=185232)**		**AoU with EHR (N=287012)**	
	N	**%**	**N**	**%**	**N**	**%**
Hypertension						
Yes	78613893	31.71	54917	29.65	119554	41.65
No	169132190	68.21	65520	35.37	167458	58.35
Do not answer	200977	0.08	49763	26.87		
Missing	0	0.00	15032	8.12		
Hyperlipidemia						
Yes	82319763	33.20	56749	30.64	109435	38.13
No	164499677	66.34	46095	24.89	177577	61.87
Do not answer	1127620	0.45	62836	33.92		
Missing	0	0.00	19552	10.56		
Diabetes						
Yes	27931576	11.27	20988	11.33	56223	19.59
No	219839616	88.66	56709	30.62	230789	80.41
Do not answer	175869	0.07	84751	45.75		
Missing	0	0.00	22784	12.30		
CAD						
Yes	10166914	4.10	7661	4.14	33521	11.68
No	229863340	92.71	36314	19.60	253491	88.32
Do not answer	419240	0.17	107980	58.29		
Missing	7497567	3.02	33277	17.97		
AMI						
Yes	8977335	3.62	4997	2.70	14875	5.18
No	231146169	93.22	65960	35.61	272137	94.82
Do not answer	325989	0.13	86627	46.77		
Missing	7497567	3.02	27648	14.93		
CHF						

Yes	6287028	2.54	3890	2.10	15994	5.57
No	233775925	94.28	40928	22.10	271018	94.43
Do not answer	386549	0.16	107241	57.90		
Missing	7497567	3.02	33173	17.91		
CVA/TIA						
Yes	9087945	3.67	6156	3.32	7969	2.78
No	231020064	93.17	46124	24.90	279043	97.22
Do not answer	341484	0.14	102600	55.39		
Missing	7497567	3.02	30352	16.39		
CKD						
Yes	7687625	3.10	3629	1.96	26221	9.14
No	232461842	93.75	1219	0.66	260791	90.86
Do not answer	300026	0.12	139414	75.26		
Missing	7497567	3.02	40970	22.12		

Supplementary Table 3: Prevalence of CV risk factors and established CVD by demographic subgroups

	Category	**NHANES Weighted**	**AOU Person and Family Survey (N=185232)**	**AOU with EHR (N=287012)**
Hypertension	18-40	12.29	8.16	12.96
	40-65	35.61	26.84	38.31
	>=65	59.85	42.74	60.55
Hyperlipidemia	18-40	11.79	8.27	6.85
	40-65	40.50	27.50	32.26
	>=65	58.14	44.47	61.17
Diabetes	18-40	2.21	3.38	6.82
	40-65	13.18	11.97	19.53
	>=65	24.20	14.59	26.42
CAD	18-40	0.42	0.11	0.48
	40-65	3.71	1.98	6.95
	>=65	11.73	8.14	22.83
AMI	18-40	0.51	0.17	0.61

	40-65	3.76	1.81	4.06
	>=65	9.12	4.77	8.84
CHF	18-40	0.32	0.27	0.70
	40-65	2.34	1.69	4.03
	>=65	7.05	3.37	9.86
CVA/TIA	18-40	0.53	0.64	0.29
	40-65	3.21	2.78	1.70
	>=65	10.41	5.14	5.28
CKD	18-40	1.74	0.59	1.53
	40-65	2.62	1.99	6.64
	>=65	6.60	3.57	15.91
Hypertension	Male	32.70	36.18	45.29
	Female	30.78	26.08	39.22
Hyperlipidemia	Male	34.92	36.97	42.01
	Female	31.60	27.18	35.59
Diabetes	Male	12.58	13.56	20.79
	Female	10.04	10.11	18.79
CAD	Male	5.84	7.84	16.81
	Female	2.48	2.14	8.39
AMI	Male	5.06	4.90	7.61
	Female	2.28	1.51	3.64
CHF	Male	2.95	2.95	7.25
	Female	2.15	1.64	4.50
CVA/TIA	Male	3.23	3.97	3.04
	Female	4.07	2.97	2.61
CKD	Male	3.33	2.51	11.94
	Female	2.88	1.66	7.33
Hypertension	White	32.98	30.26	42.45
	Black	39.04	42.23	47.89
	Asian	24.98	15.72	23.05
	Hispanic/Latino/a/x	27.73	21.12	37.22
	Other/ >1	24.41	22.04	32.38
Hyperlipidemia	White	36.16	33.08	44.73
	Black	26.96	26.35	30.76
	Asian	28.66	22.00	26.96
	Hispanic/Latino/a/x	31.90	22.83	30.87
	Other/ >1	27.23	23.52	28.02
Diabetes	White	10.91	10.20	16.64
	Black	12.49	19.20	24.60

	Asian	10.86	7.03	13.13
	Hispanic/Latino/a/x	9.68	13.92	24.86
	Other/ >1	13.08	10.09	16.32
CAD	White	5.40	4.88	14.09
	Black	1.72	2.31	9.87
	Asian	2.12	1.74	6.07
	Hispanic/Latino/a/x	2.17	1.55	7.90
	Other/ >1	1.98	2.64	8.56
AMI	White	4.28	2.96	5.45
	Black	3.04	2.70	5.90
	Asian	1.45	1.02	2.38
	Hispanic/Latino/a/x	2.54	1.64	4.21
	Other/ >1	2.61	1.81	3.95
CHF	White	2.81	2.10	5.51
	Black	3.56	3.75	7.69
	Asian	0.83	0.59	2.08
	Hispanic/Latino/a/x	1.78	1.20	4.21
	Other/ >1	1.54	1.44	4.07
CVA/TIA	White	3.97	3.56	3.31
	Black	4.83	4.33	2.40
	Asian	1.50	1.15	1.34
	Hispanic/Latino/a/x	2.61	1.68	2.00
	Other/ >1	2.79	2.46	1.94
CKD	White	3.13	2.08	9.25
	Black	3.93	2.59	11.12
	Asian	1.85	0.77	4.50
	Hispanic/Latino/a/x	2.87	1.01	7.72
	Other/ >1	2.91	1.40	6.77

REFERENCES

1. Investigators TAoURP. The “All of Us” Research Program. New England Journal of Medicine 2019;381:668-676.
2. Prevention UCfDCa. US National Health and Nutrition Examination Survey.
3. Jang M, Vorderstrasse A. Socioeconomic Status and Racial or Ethnic Differences in Participation: Web-Based Survey. JMIR Res Protoc 2019;8:e11865.
4. Stamatakis E, Owen KB, Shepherd L, Drayton B, Hamer M, Bauman AE. Is Cohort Representativeness Passé? Poststratified Associations of Lifestyle Risk Factors with Mortality in the UK Biobank. Epidemiology 2021;32:179-188.
5. Zeng C, Schlueter DJ, Tran TC et al. Comparison of phenomic profiles in the All of Us Research Program against the US general population and the UK Biobank. J Am Med Inform Assoc 2024;31:846-854.
6. Wang J, Ferguson EL, Buto P et al. Sociodemographic, health-related, and clinical characteristics and their associations with mortality among all of us participants compared with the United States general population. Am J Epidemiol 2025.
7. Sulieman L, Cronin RM, Carroll RJ et al. Comparing medical history data derived from electronic health records and survey answers in the All of Us Research Program. J Am Med Inform Assoc 2022;29:1131-1141.