

https://dspace.org/

Cohort Profile Africa WitsINDEPTH partnership for Genomic studies AWIGen in four subSaharan African countries

Furahini Tluway; Palwende Boua; Siyanda Madala; Ananyo Choudhury; Nigel Crowther ...

OXFORD UNIV PRESS
https://hdl.handle.net/10539/47391
WOS

Downloaded from a DSpace repository. https://dspace.org/

International Journal of Epidemiology, 2025, **54(1)**, dyae173
https://doi.org/10.1093/ije/dyae173
Cohort Profile

IEA International Epidemiological Association
OXFORD

Cohort Profile

Cohort Profile: Africa Wits-INDEPTH partnership for Genomic studies (AWI-Gen) in four sub-Saharan African countries

Furahini Tluway [1], Godfred Agongo [2,3], Vukosi Baloyi [4], Palwende Romuald Boua [1,5], Isaac Kisiangani [6], Moussa Lingani [5], Reneilwe Given Mashaba [7], Shukri F Mohamed [6], Engelbert A Nonterah [3], Cairo Bruce Ntimana [7], Toussaint Rouamba [5], Theophilous Mathema [1], Siyanda Madala [1,8], Dylan G Maghini [1,9], Ananyo Choudhury [1], Nigel J Crowther [10], Scott Hazelhurst [1,11], Dhriti Sengupta [1], Patrick Ansah [3], Solomon Simon Rampai Choma [12], Cornelius Debpuur [13], F Xavier Gómez-Olivé [14], Kathleen Kahn [14], Lisa K Micklesfield [4], Shane A Norris [4,15], Abraham R Oduro [13], Hermann Sorgho [5], Paulina Tindana [16], Halidou Tinto [5], Stephen Tollman [14], Alisha Wade [14], Michèle Ramsay [1,*] and as members of AWI-Gen and the H3Africa Consortium

[1]Sydney Brenner Institute for Molecular Bioscience, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa, [2]Department of Biochemistry and Forensic Sciences, School of Chemical and Biochemical Sciences, C. K. Tedam University of Technology and Applied Sciences, Ghana, [3]Navrongo Health Research Centre, Ghana Health Service, Navrongo, Ghana, [4]South African Medical Research Council/University of the Witwatersrand Developmental Pathways for Health Research Unit, Department of Paediatrics, School of Clinical Medicine, University of the Witwatersrand, Johannesburg, South Africa, [5]Clinical Research Unit of Nanoro, Institut de Recherche en Sciences de la Santè, Nanoro, Burkina Faso, [6]African Population and Health Research Center (APHRC), Nairobi, Kenya, [7]DIMAMO Population Research Centre, University of Limpopo, Polokwane, South Africa, [8]School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa, [9]Department of Medicine (Hematology), Stanford University, Stanford, CA, USA, [10]Department of Chemical Pathology, National Health Laboratory Service, University of the Witwatersrand, Faculty of Health Sciences, Johannesburg, South Africa, [11]School of Electrical & Information Engineering, University of the Witwatersrand, Johannesburg, South Africa, [12]Department of Pathology and Medical Sciences, University of Limpopo, Polokwane, South Africa, [13]Research and Development Division, Ghana Health Service, Accra, Ghana, [14]MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health, University of the Witwatersrand, Johannesburg, South Africa, [15]School of Human Development and Health, University of Southampton, Southampton, UK and [16]Department of Health Policy, Planning and Management, School of Public Health, University of Ghana, Accra Ghana

*Corresponding author. SBIMB, Faculty of Health Sciences, University of the Witwatersrand, Private Bag 3, Wits 2050, Gauteng, South Africa.
E-mail: Michele.Ramsay@wits.ac.za

Keywords: AWI-Gen, sub-Saharan Africa, cardiometabolic disease, adult population cohort, genetics, genomics, gut microbiome.

Key Features

- Africa Wits-INDEPTH partnership for Genomic studies (AWI-Gen) was established to examine genomic, environmental and behavioural factors influencing body composition and cardiometabolic diseases and traits in African populations.
- Population-based longitudinal cohort, involving four sub-Saharan African countries representing rural and urban settings: West Africa [Nanoro (Burkina Faso) and Navrongo (Ghana)]; East Africa [Nairobi (Kenya)]; and South Africa (Agincourt, Dikgale/DIMAMO and Soweto).
- Baseline data collected from 2013 to 2017, enrolling 12 032 adults [55.1% women; mean (SD) 51.9 (8.3) years, range 37–82], with follow-up ∼5 years later and a retention rate of almost 60%. In Wave 2 an additional 579 individuals were enrolled ($n = 7804$; 55.9% women; 57.0 (7.9), 39–98).
- Main categories of data collected at two time points include sociodemographic characteristics, history of chronic diseases and lifestyle behaviours, with spirometry, cognition and frailty added in the second data collection wave. Measurements at both time points include anthropometry, blood pressure, carotid intima–media thickness and body fat distribution. Blood and urine samples are collected to measure biomarkers for diabetes, HIV, dyslipidaemia and kidney disease, and stool samples are collected in a women's sub-sample for gut microbiome analyses. Genome-wide genotyping is available for all participants and whole genome sequences for a subset.
- We encourage collaboration, and data are accessible through the AWI-Gen Principal Investigator [Michele.Ramsay@wits.ac.za] in consultation with the steering committee or the H3Africa Data and Biospecimen Access Committee.

Received: 12 June 2024. **Editorial Decision:** 7 November 2024. **Accepted:** 16 December 2024
© The Author(s) 2025. Published by Oxford University Press on behalf of the International Epidemiological Association. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (https://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

Downloaded from https://academic.oup.com/ije/article/54/1/dyae173/7997654 by Cynthia Warren user on 14 April 2025

Why was the cohort set up?

Cardiometabolic diseases (CMDs) are on the rise worldwide and are leading causes of premature mortality,[1] a trend also observed in low- and middle-income countries (LMICs) including those in Africa.[2] There is a paucity of data from African countries on CMD prevalence, particularly on the genetic predisposition to common diseases, despite the recognition that African populations have high genetic diversity and unique population substructure.[3–5] To address this gap, we set up the Africa Wits-INDEPTH partnership for Genomic studies (AWI-Gen).

The AWI-Gen population cohort, initiated in 2012, was funded by the National Institutes of Health (NIH) as part of the Human Heredity and Health in Africa (H3Africa) Consortium.[6] AWI-Gen is a collaboration between the University of the Witwatersrand (Wits) in Johannesburg and the International Network for the Demographic Evaluation of Populations and Their Health (INDEPTH). Five of the six study sites reside within existing health and demographic surveillance systems (HDSS). The remaining site is a Wits research unit in urban Soweto, Johannesburg.

AWI-Gen's main goal is to provide a research resource to investigate the prevalence of CMD and associated risk factors, and to explore gene–gene and gene–environment interactions.[7,8] Additional aims are to strengthen the capacity for genomic research in Africa and to understand the key ethical, legal and sociocultural issues that would affect implementation.

In 2012, a meeting was held with ethics committee members to consider appropriate models for informed consent, data sharing and community benefit.[9] Community sensitization was followed by context-relevant stepwise permissions and individualized consent before entry into the study. The approach of broad consent for the future use of data and biological samples was agreed on by the research team and was also the preferred model adopted by H3Africa. The AWI-Gen process was documented at all sites, highlighting community engagement experiences and lessons learned.[10,11] General health awareness and education were provided and point-of-care tests performed (glucose, lipids and HIV, the latter with supportive trained counsellors). Participants with health concerns (e.g. high blood pressure and high fasting glucose levels) were referred to the local health facilities with explanatory notes. Appropriate feedback was provided in reasonable time frames, and ongoing engagement and awareness of research progress were shared with communities.

The longitudinal nature of the cohort permits assessment of changes in measurements and behaviour over about a 5-year period, and their impact on cardiometabolic endpoints. Knowledge gained from the AWI-Gen study has the potential for translational impact, including policy briefs and development of strategies for precision medicine applications in participating communities.

Who is in the cohort?

AWI-Gen study participants are from five INDEPTH member HDSS centres in Africa, with diverse representation of populations from rural and urban environments in West, East and South Africa.[7,8] These centres are Nanoro (Burkina Faso),[12] Navrongo (Ghana),[13] Nairobi (Kenya),[14] Agincourt (South Africa)[15] and Dikgale (later renamed DIMAMO) (South Africa)[16] and the SAMRC/Wits Developmental Pathways for Health Research Unit (DPHRU) (Soweto, South Africa).[17] The geographical location, number and biological sex distribution of participants from each study centre and a timeline for training, community engagement and recruitment, are shown in Figure 1, with study participant characteristics in Table 1.

At baseline (Wave 1), AWI-Gen was a population-based cross-sectional survey that aimed to enrol 2000 participants, 40–60 years of age, approximately equal numbers of men and women, from each of the six study centres. Random sampling was used to select participants from existing populations based at the HDSS. Each study centre used mechanisms to minimize selection bias. Criteria excluded closely related individuals, pregnant women and individuals resident in the community for <10 years. Closely related individuals were considered first-degree relatives, but despite this, genetic analysis revealed significant relatedness among participants in some centres. Ethics approval was obtained from the Wits Human Research Ethics Committee (Medical) and participating institutions and/or national ethics committees. Participants gave informed consent to participate in the study before taking part.

We enrolled 12 032 participants, with 88.1% between 40 and 60 years of age. Agincourt had 447 individuals >60 years of age due to harmonization with the Health and Aging in Africa: A Longitudinal Study in South Africa (HAALSI) cohort.[8] Participants were selected using purposeful sampling based on existing population sample frames from the respective study sites; details are provided in Ali *et al.*[8] and in the Supplementary Material (available as Supplementary data at *IJE* online).

How often have they been followed up?

Data and sample collection was conducted at two time points (Figure 1a and Table 1). During Wave 1 (2013–17), 12 032 participants were enrolled. Wave 2 of data collection took place between 2018–22, as a closed cohort with one exception: the Dikgale/DIMAMO study centre was augmented with an additional 579 new participants in Wave 2. A 60% retention of the original cohort was achieved (range 47–72% across centres) and data and samples collected on 7804 participants (7225 recalled and 579 new participants) in Wave 2 (Table 1). Those remaining in the cohort were on average younger [50.0 (34–82) years vs $52^{29–81}$] and female (57.1% vs 52.0% women).

What has been measured?

The AWI-Gen cohort variables, collection instruments and protocols for Wave 1 are described in detail in Ali *et al.*[8] and include sociodemographic characteristics, lifestyle behaviour, medical history of chronic and infectious diseases, anthropometry, body fat distribution, blood pressure and blood and urine biomarkers. In Wave 2, additional variables and measurements were included (Table 2).

Genomic data: about 11 000 AWI-Gen participants were genotyped on the H3Africa Consortium single nucleotide variant (SNV) genotyping array, and additional genotypes imputed using the African Genome Reference panel at the Sanger Imputation Service,[18] with a final dataset of 10 603 participants and 13.9 million SNVs. About 100 South African participants and 60 from Burkina Faso and Ghana have whole genome sequence (WGS) data (30x coverage), and eight have PacBio WGS data (10–20x). A further ~600 South African participants have WGS (4x) data and matching transcriptome data from RNA extracted from whole blood

Downloaded from https://academic.oup.com/ije/article/54/1/dyae173/7997654 by Cynthia Warren user on 14 April 2025

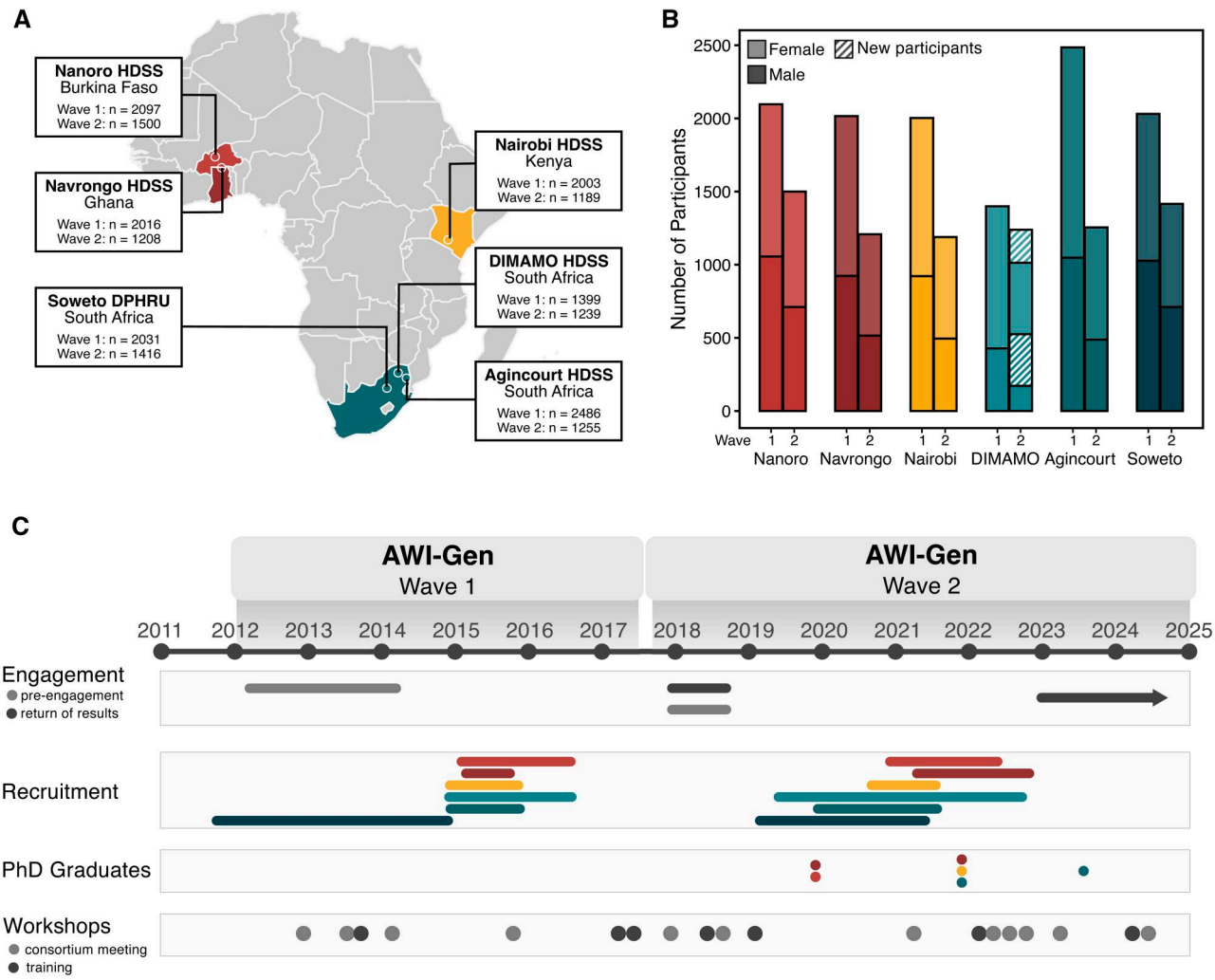


Figure 1 Description of the AWI-Gen study across two waves of data collection. (A) Location of the six study locations in health and demographic surveillance systems (HDSS) and the Developmental Pathways to Health Research Unit (DPHRU), showing the number of participants recruited in Wave 1 and the number recalled in Wave 2. Note that the DIMAMO cohort was enhanced in Wave 2 due to the lower recruitment numbers in Wave 1. (B) Number of men and women per study centre with the augmentation in DIMAMO shown with the cross-hatching. (C) Timeline showing community engagement that is still ongoing, and recruitment during Wave 1 and Wave 2. Recruitment started earlier in Soweto since we onboarded women from a study on menopause which was already on the files from 2012. Six PhD students graduated through the Africa Wits-INDEPTH partnership for Genomic studies (AWI-Gen) study and a further three are in progress. AWI-Gen meetings coincided with some of the Human Heredity and Health in Africa (H3Africa) consortium meetings, and training workshops were held to ensure that data were collected according to the same protocols and standard operating procedures. The colour coding per centre is consistent across all figures

(study performed in collaboration with Variant Bio). There are 16S and shotgun sequence data from gut microbiome samples of 169 participants at two study sites from Wave 1, and shotgun sequence data from 1824 participants at six study sites (overlap of 76) from Wave 2.

What has it found?

Demographic characterization and key variables at baseline (Wave 1) and follow-up, roughly 5 years later, (Wave 2) are shown in Figure 2 and Table 3. The age distributions are shown across the two waves of data collection (Figure 2A) and the genetic-informed population sub-structure is demonstrated in a plot based on principal component analysis (Figure 1B). The percentage of individuals with specific risk factors [body mass index (BMI); Figure 2C] and CMD outcomes (obesity and overweight, diabetes, chronic kidney disease and hypertension; Figure 2D) increased over the 5-year period, as would be expected for participants in this age group.

At baseline, we observed the highest prevalence of obesity (BMI ≥30) at the three South African study sites (42.3–66.6% in women and 2.81–17.5% in men) and the lowest in West Africa (1.2–4.2% in women and 1.19–2.20% in men).[19] The South African study sites had a mean prevalence of hypertension of over 40.0%, followed by Nairobi (25.6%) and Navrongo (24.5%) [hypertension was defined as self-report and/or systolic blood pressure (SBP) ≥140 mmHg and/or diastolic blood pressure (DBP) ≥90].[20] Indications of kidney damage [estimated glomerular filtration rate (eGFR) <60 mL/min per 1.73 m^2 and/or albuminuria (urine creatinine ratio >3 mg/mmol)] were present in 10.7%[21]; and diabetes prevalence was 5.5% (95% CI 4.4% to 6.5%) (defined as self-report and/or fasting plasma glucose ≥7 mmol/L and/

Downloaded from https://academic.oup.com/ije/article/54/1/dyae173/7997654 by Cynthia Warren user on 14 April 2025

Table 1 Africa Wits-INDEPTH partnership for Genomic studies (AWI-Gen) study site and study characteristics

	Study site					
	Nanoro	Navrongo	Nairobi	Dikgale/DIMAMO	Agincourt	Soweto
Country	Burkina Faso	Ghana	Kenya	South Africa		
Study site characteristics						
HDSS Catchment Area, km^2	594	1675	6	545	420	200
HDSS Catchment Area population, thousands	63	156	75	36	115	1200
Population density per km^2	105	91	14 833	113	274	6357
Map coordinates	12.57N, 1.92W 12.72N, 2.31W	10.89N, 1.09W	1.25S, 36.89E 1.31S 36.87E	23.65S, 29.65E 23.90S 29.85E	24.50S, 31.08E 24.56S 31.25E	26.24S, 27.84E
Altitude, m	313	196	1790	1250	400-600	1632
AWI-Gen study characteristics						
Wave 1 ($n = 12\,032$)						
Participants (n)	2097	2016	2003	1399	2486	2031
Female (%)	49.6	54.2	53.9	69.4	57.8	49.3
Mean (SD) age (years)	49.8 (5.9)	51.1 (5.8)	49.1 (6.1)	52.2 (8.3)	58.7 (11.0)	49.3 (5.9)
Age range (years)	40–62	41–65	38–67	37–82	41–81	40–69
Wave 2 ($n = 7804$)						
Participants (n)	1498	1208	1189	1239[b]	1255	1415
Female (%)	52.6	57.4	58.4	57.6	61.1	49.8
Mean (SD) age (years)	55.0 (5.8)	57.0 (5.7)	54.3 (6.1)	58.4 (9.4)	62.6 (10.5)	55.3 (6.0)
Age range (years)	45–67	47–71	43–73	39–98	46–86	44–73
Retention in Wave 2: n (%)	599 (28.6)	808 (40.1)	814 (40.6)	739 (52.8)	1231 (49.5)	616 (30.0)[a]

HDSS, health and demographic surveillance survey; SD, standard deviation.
[a] For the Soweto cohort, recall was finalized when 70% of participants had been recalled into Wave 2. Not all participants were recontacted.
[b] For the Dikgale/DIMAMO cohort an additional 579 participants were recruited in Wave 2.

or random plasma glucose ≥11.1 mmol/L).[22] A cross-sectional comparison of CMD risk factors between pre- and postmenopausal women in Wave 1 highlighted that differences were more prominent in the West African cohorts despite, or perhaps because, the CMD risk factors were more prevalent at all ages in the South and East African cohorts.[23] Multimorbidity was common across the AWI-Gen study sites, with overall prevalence of 47.2% in women and 35.0% in men.[24] Commonly used 10-year CVD risk calculators were applied to all AWI-Gen participants, with risk varying from 2.6% to 6.5%, and with differences between study sites with highest risk in South Africa.[25] These data highlighted the inappropriateness of current CVD risk calculators for use in an African setting. Each of the AWI-Gen publications highlights differences between the study sites, with the South African sites having a higher prevalence of CMD indicators as they are further along the health epidemiological transition, and the West African sites having the lowest levels.

A high prevalence of HIV infection was observed at the South (14–34%) and East African (12%) study sites, with low levels in West Africa (<1%). Interactions with HIV and the impact of HIV on CMD was examined in the South and East African study sites.[19]

A pilot gut microbiome study in Agincourt ($n = 118$) and Soweto ($n = 51$) in Wave 1 showed important differences between the two sites, and revealed previously undescribed gut microbial diversity.[26] The larger metagenomic study of 1820 female participants in Wave 2, across all six study sites, revealed unique taxonomic patterns and distributions and many novel species.[27]

Patterns of tobacco and alcohol use showed important regional and sex-specific differences, with 68.4% of men and 33.3% of women being current alcohol users, and 34.5% of men but only 2.1% of women being current smokers.[28] We examined gene–smoking interactions and their effect on the carotid intima–media thickness (CIMT) and detected sex differences.[29] The *RCBTB1* region was involved in gene–smoking interactions and identified novel associations. Genetic associations identified eight independent variants associated with either smoking initiation or cessation.[30]

A key objective of AWI-Gen was to perform genome-wide association studies (GWAS) with different indicators of CMD disease and specific outcomes. With representation from West, East and Southern Africa, and demonstrated clear population structure (Figure 2B), we used different approaches to adjust for population structure during the analyses; 'mega-analysis', which includes the full dataset and uses principal component-based correction for population structure, and independent GWASs for each geographical region followed by a meta-analysis. Our lipid trait GWAS compared results from the two approaches and found the outcomes to be largely similar. Our studies replicated many well-known genetic associations suggesting global transferability of loci associated with several cardiometabolic traits, but the identification of novel associations highlighted the value of working with data from African populations. Harnessing the increased genetic diversity and the generally lower linkage disequilibrium in African populations allowed us to fine-map some well-known associated loci. The underperformance of European-based polygenic scores (PGSs) in our dataset emphasizes the need for more PGS-based research on ancestrally diverse populations. GWAS on: carotid intima–media thickness (CIMT)[31]; lipid traits[32]; hypertension and blood pressure traits[33]; and urinary albumin creatinine ratios as a marker of kidney function[34]; are published, and other GWASs are in the publication pipeline.

We explored runs of homozygosity as a marker of inbreeding depression, to assess their effect on CMD-associated traits and to examine the effect of different levels of urbanization (using night-time luminosity as a proxy marker).[35] Sex effects were evident, with inbreeding having a decreasing effect in men and an increasing effect in women for BMI,

Downloaded from https://academic.oup.com/ije/article/54/1/dyae173/7997654 by Cynthia Warren user on 14 April 2025

Table 2 Data and variables collected during Africa Wits-INDEPTH partnership for Genomic studies (AWI-Gen) Wave 1 (baseline) and Wave 2

Data domain	Variables	Wave 1	Wave 2
Phenotype data collected			
Sociodemographic characteristics	Age at enrolment, sex, ethno-linguistics, country, level of education, employment status, marital status	Yes	Yes
Family composition and household attributes	Number of siblings and children, household size and other socioeconomic characteristics	Yes	Yes
Reproductive history and menopause status (women)	Number of pregnancies and live births, menstrual and family planning history, date of final menstrual period	Yes	Yes
Lifestyle behaviours and exposure to pesticides	Diet, physical activity, sleep, tobacco and alcohol use, exposure to pesticides	Yes	Yes
Infectious diseases	Tuberculosis, HIV and malaria	Yes	Yes
History of chronic diseases and risk factors	History of diabetes, hypertension, cardiac diseases, chronic kidney disease, dyslipidaemia, stroke, cancers	Yes	Yes
Cognition and frailty	Immediate and delayed recall, orientation, grip strength, chair rise sit to stand, gait speed	No	Yes
Trauma	History of assault, injuries, serious illness, domestic violence, bereavement, loss of employment,	No	Yes
Respiratory health[a]	History of upper and lower respiratory infections, asthma	No	Yes
Microbiome[a]	History of antibiotic use, diarrhoea, intestinal worms, probiotic use	No	Yes
Measurements			
Anthropometry	Weight, height, waist and hip circumferences	Yes	Yes
Vital signs	Blood pressure, pulse	Yes	Yes
Ultrasound measurements	Visceral and subcutaneous adipose tissue, left, and right intima–media thickness	Yes	Yes
Spirometry[a]	Spirometry test	No	Yes
Biomarkers			
Blood	Serum HDL-C and LDL-C and triglycerides, serum creatinine, insulin and fasting blood glucose, sex hormones measurements[a]	Yes	Yes
Urine	Albumin, creatinine and total protein	Yes	Yes
Stool[a]	Gut microbiome characterization	Yes	Yes

HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol.
[a] These variables are available for a subset of the AWI-Gen cohort.

subcutaneous adipose tissue, low-density lipoprotein and total cholesterol levels, with more intense effects in men.

In the South African subset of AWI-Gen, the genetic data revealed population structure that aligns with language subfamilies and geography, which is largely driven by differential gene flow from the Khoe-San hunter gatherers.[5] We call for a shift from the practice of combining all South-Eastern Bantu speakers from South Africa into a single group in health-related studies to recognize the genetic differences between the groups and their potential impact on the phenotype. AWI-Gen played a key role in leading the data analysis and interpretation for the pan-African H3Africa publication in *Nature* in 2020 and contributed data from the West African study sites.[4] Over 60 AWI-Gen papers have been published over the past 12 years.

What are the main strengths and weaknesses?

A major strength of AWI-Gen is that it capitalized on the unique characteristics of existing HDSS centres that offer established infrastructure for surveillance, research and longitudinal follow-up, and sustained contact with their communities. This provided access to participants, ensuring geographical spread with highly diverse demographic and phenotype data in addition to genome-wide genetic data. Another strength is that these populations are largely treatment naïve and, therefore, measurements reflect base physiological characteristics. Importantly, data were collected according to harmonized standard operating procedures, measurements were done using the same equipment sourced from a single supplier, laboratory assays were performed in a single laboratory, and training was conducted to ensure uniform approaches by fieldworkers and laboratory staff across the sites. A broad consent model was followed, and participants were asked for specific consent for the AWI-Gen study and broad consent for future research studies, provided these studies had ethics approval from the Wits HREC (Medical).

The two waves of data collection allow genomic association studies with disease incidence and with rates of change of key parameters that may affect onset and progression of disease. AWI-Gen and its members supported and were part of the Pan-African Bioinformatics Network for H3Africa (H3ABioNet),[36,37] which contributed significantly to bioinformatics capacity and infrastructure developed to support genetics and genomic research in Africa.

Despite the significant strengths, AWI-Gen is limited to a specific age group (middle-aged adults), specific geographical locations and a ~60% cohort retention, and therefore findings may not be generalizable to all African populations.

Can I get hold of the data? Where can I find out more?

AWI-Gen datasets from data collection Wave 1 have been submitted to the European Genome-phenome Archive (EGA) [https://www.ebi.ac.uk/ega/studies/EGAS00001002482O] and

Downloaded from https://academic.oup.com/ije/article/54/1/dyae173/7997654 by Cynthia Warren user on 14 April 2025

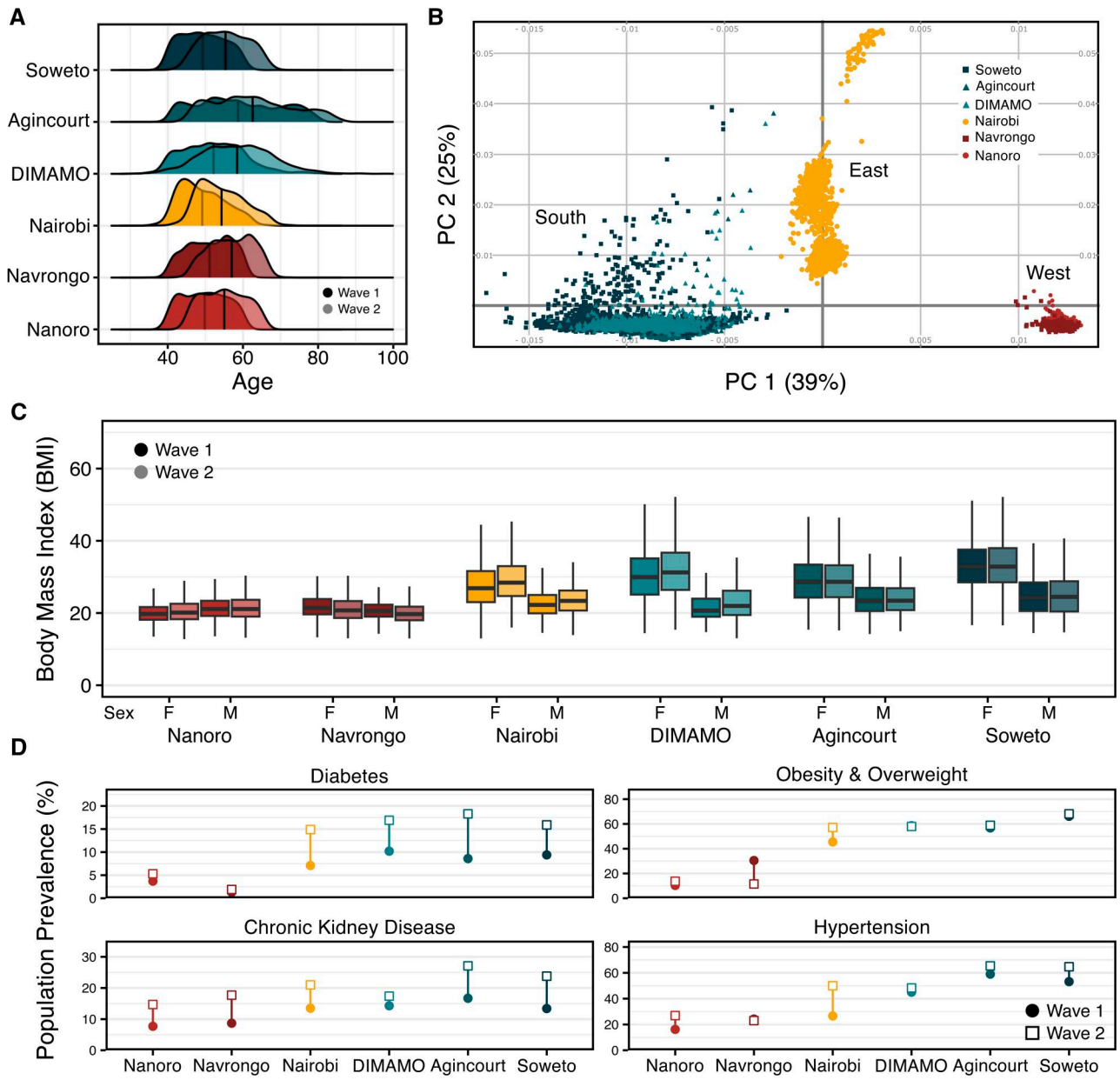


Figure 2 Outcomes from the longitudinal AWI-Gen study highlighting genetic diversity and key cardiometabolic diseases risk factors and endpoints. (A) Distribution of ages of participants in Wave 1 and Wave 2. (B) Principal component (PC) analysis using genome-wide genotyping data from the H3Africa single nucleotide polymorphism (SNP) array, showing the clustering of participants from the South, East and West Africa study centres. It highlights the population sub-structure and emphasizes the need to adjust for ancestry in genome-wide association studies. (C) Distribution of body mass index (BMI) in women (F) and men (M) in Wave 1 and Wave 2, noting the higher BMI in women and overall, in the East and South African cohorts. Note that outliers and individuals with BMI >70 ($n = 2$) were excluded from this visualization. Lower and upper hinges correspond to first and third quartiles, and whiskers indicate largest and smallest values within 1.5 * interquartile range. (D) The prevalence % of the combined cohort of four key cardiometabolic outcomes, diabetes, obesity and overweight, chronic kidney disease and hypertension, at two different time points (Wave 1 and Wave 2). Almost all conditions have a higher prevalence at Wave 2 when participants are about 5 years older. Note the different scales for the percentages

can be accessed through the H3Africa data and Biospecimen Access Committee (DBAC). Non-human microbiome data have been submitted to the European Nucleotide Archive (ENA) data repository [https://www.ebi.ac.uk/ena/browser/home] and can be accessed without restriction. Wave 2 data will also be available through the EGA. The DNA samples have been submitted to the H3Africa Biorepository in South Africa. For new collaborations and data and sample access, contact the AWI-Gen Principal Investigator Michele Ramsay at [Michele.Ramsay@wits.ac.za] or the H3Africa DBAC [https://catalog.h3africa.org].

Ethics approval

Human Research Ethics Committee (Medical) of the University of the Witwatersrand (M121029, M16021, M170880,

Downloaded from https://academic.oup.com/ije/article/54/1/dyae173/7997654 by Cynthia Warren user on 14 April 2025

Table 3 Demographic characteristics and key outcome variables for AWI-Gen at baseline (Wave 1) and follow-up (Wave 2), roughly 5 years later, comparing men and women

	Wave 1			Wave 2		
Key variable	Total *n* (%)	Male *n* (%)	Female *n* (%)	Total *n* (%)	Male *n* (%)	Female *n* (%)
Age groups in years	12 032			7807		
<50	5164 (42.9)	2365 (43.8)	2799 (42.2)	1441 (18.5)	702 (20.4)	739 (16.9)
50–60	5535 (46.0)	2431 (45.0)	3104 (46.8)	3884 (49.8)	1625 (47.2)	2259 (51.8)
>60	1333 (11.1)	609 (11.3)	724 (10.9)	2482 (31.8)	1116 (32.4)	1366 (31.3)
Age in years	12 032			7807		
Mean (SD)	51.9 (8.3)	51.8 (8.4)	52.0 (8.3)	57.0 (7.9)	57.0 (8.2)	57.1 (7.7)
Body mass index	11 945			7715		
Underweight	1317 (11.0)	744 (13.9)	573 (8.7)	984 (12.7)	555 (16.2)	429 (10.0)
Normal	5706 (47.8)	3246 (60.5)	2460 (37.4)	3338 (43.3)	1863 (54.4)	1475 (34.4)
Overweight	2371 (19.9)	971 (18.1)	1400 (21.3)	1574 (20.4)	690 (20.2)	884 (20.6)
Obese	2551 (21.4)	409 (7.6)	2142 (32.6)	1819 (23.6)	314 (9.2)	1505 (35.1)
Hypertension	12 032			7734		
No	7494 (62.3)	3513 (65.0)	3981 (60.1)	4158 (53.8)	879 (54.7)	2279 (53.0)
Yes	4538 (37.7)	1892 (35.0)	2646 (39.9)	3576 (46.2)	1553 (45.3)	2023 (47.0)
Diabetes	12 032			7748		
No	11240 (93.4)	5100 (94.4)	6140 (92.7)	6814 (87.9)	3068 (89.2)	3746 (86.9)
Yes	792 (6.6)	305 (5.6)	487 (7.4)	934 (12.1)	372 (10.8)	562 (13.1)
Chronic kidney disease	9836			6862		
No	8648 (87.9)	4295 (88.9)	4353 (87.0)	5454 (79.5)	2482 (80.6)	2972 (78.6)
Yes	1188 (12.1)	539 (11.1)	649 (13.0)	1408 (20.5)	599 (19.4)	809 (21.4)

M2210108), with additional ethics approvals from the participating institutions and/or national ethics committees.

Data availability

See 'Can I get hold of the data?' above.

Supplementary data

Supplementary data are available at *IJE* online.

Author contributions

Implementation of field activities and supervision of data collection and project administration: G.A., P.R.B., V.B., I.K., M.L., R.G.M, S.F.M, E.A.N. and C.B.N. Data management: quality control of preliminary analyses, tables and figure generation: D.G.M., T.M., S.M. and T.R.. Site principal investigators (PIs) and co-PI, project leads, senior scientists: P.A., A. C., S.S.R.C., N.J.C., S.H., F.X.G., F, K.K., L.K.M, S.A.N., A. R.O., C.P., D.S., H.S., H.T., P.T., S.T., A.W.). Senior project manager (F.T.) and the principal investigator (M.R.) were responsible for drafting the initial version of this manuscript. All authors had the opportunity to revise and critically review the manuscript, and each accepts accountability for the accuracy and integrity of the manuscript.

Use of artificial intelligence (AI) tools

No AI tools were used collecting and/or analysing data, producing images or graphical elements nor in writing the paper.

Funding

The AWI-Gen Collaborative Centre is funded by the National Human Genome Research Institute (NHGRI), the National Institute of Environmental Health Sciences (NIEHS), of the National Institutes of Health (NIH) under award number U54HG006938 and its supplements, as part of the H3Africa Consortium, and by the Department of Science and Innovation, South Africa, award number DST/CON 0056/2014. D.G.M was supported by the NIH Fogarty Global Health Equity Scholars Program (NIH FIC D43TW010540).

Acknowledgements

We acknowledge the funders, the AWI-Gen participants, the H3Africa Consortium and the extraordinary study coordinators, field teams and investigators at all six study centres across the duration of the study.

Conflict of interest

None declared.

References

1. Forouzanfar MH, Afshin A, Alexander LT *et al.* Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* 2016;**388**:1659–724.
2. Gouda HN. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2018;**392**:1923–94. Erratum in: *Lancet* 2019; 393:132. Erratum in: *Lancet* 2019;393:e44.
3. Gomez F, Hirbo J, Tishkoff SA. Genetic variation and adaptation in africa: implications for human evolution and disease. *Cold Spring Harb Perspect Biol* 2014;**6**:a008524.
4. Choudhury A, H3Africa Consortium, Aron S, Botigué LR *et al.* High-depth African genomes inform human migration and health. *Nature* 2020;**586**:741–48. Erratum in: *Nature* 2021;592:E26.
5. Sengupta D, H3Africa Consortium, Choudhury A, Fortes-Lima C *et al.* Genetic substructure and complex demographic history of South African Bantu speakers. *Nat Commun* 2021;**12**:2080.
6. The H3Africa Consortium; C Rotimi, A Abayomi, A Abimiku et al Enabling the genomic revolution in Africa. *Science* 2014; **344**:1346–48.

Downloaded from https://academic.oup.com/ije/article/54/1/dyae173/7997654 by Cynthia Warren user on 14 April 2025

7. Ramsay M, Crowther N, Tambo E *et al.* H3Africa AWI-Gen Collaborative Centre: a resource to study the interplay between genomic and environmental risk factors for cardiometabolic diseases in four sub-Saharan African countries. *Glob Health Epidemiol Genom* 2016;**1**:e20.
8. Ali SA, Soo C, Agongo G *et al.* Genomic and environmental risk factors for cardiometabolic diseases in Africa: methods used for Phase 1 of the AWI-Gen population cross-sectional study. *Glob Health Action* 2018;**11**:1507133.
9. Ramsay M, de Vries J, Soodyall H, Norris SA, Sankoh O; aH3Africa Consortium. Ethical issues in genomic research on the African continent: experiences and challenges to ethics review committees. *Hum Genomics* 2014;**8**:15.
10. Agongo G, Debpuur C, Amenga-Etego L *et al.* Community engagement and feedback of results in the H3Africa AWI-Gen project: experiences from the Navrongo Demographic and Health Surveillance site in Northern Ghana. *AAS Open Res* 2021;**4**:15.
11. Kisiangani I, Mohamed SF, Kyobutungi C *et al.* Perspectives on returning individual and aggregate genomic research results to study participants and communities in Kenya: a qualitative study. *BMC Med Ethics* 2022;**23**:27.
12. Derra K, Rouamba E, Kazienga A *et al.* Profile: Nanoro health and demographic surveillance system. *Int J Epidemiol* 2012;**41**:1293–301.
13. Oduro AR, Wak G, Azongo D *et al.* Profile of the Navrongo health and demographic surveillance system. *Int J Epidemiol* 2012;**41**:968–76.
14. Beguy D, Elung'ata P, Mberu B *et al.* Health & demographic surveillance system profile: the Nairobi Urban Health and Demographic Surveillance System (NUHDSS). *Int J Epidemiol* 2015;**44**:462–71.
15. Kahn K, Collinson MA, Gómez-Olivé FX *et al.* Profile: Agincourt health and socio-demographic surveillance system. *Int J Epidemiol* 2012;**41**:988–1001.
16. Alberts M, Dikotope SA, Choma SR *et al.* Health & demographic surveillance system profile: the Dikgale health and demographic surveillance system. *Int J Epidemiol* 2015;**44**:1565–71.
17. Richter L, Norris S, Pettifor J, Yach D, Cameron N. Cohort Profile: Mandela's children: The 1990 birth to twenty study in South Africa. *Int J Epidemiol* 2007;**36**:504–11.
18. Sengupta D, Botha G, Meintjes A *et al.*; H3Africa Consortium. Performance and accuracy evaluation of reference panels for genotype imputation in sub-Saharan African populations. *Cell Genom* 2023;**3**:100332.
19. Ramsay M, Crowther NJ, Agongo G *et al.*; as members of AWI-Gen and the H3Africa Consortium. Regional and sex-specific variation in BMI distribution in four sub-Saharan African countries: the H3Africa AWI-Gen study. *Glob Health Action* 2018;**11**:1556561.
20. Gómez-Olivé FX, Ali SA, Made F *et al.*; AWI-Gen and the H3Africa Consortium. Regional and sex differences in the prevalence and awareness of hypertension: an H3Africa AWI-Gen Study across 6 sites in Sub-Saharan Africa. *Glob Heart* 2017;**12**:81–90.
21. George JA, Brandenburg J-T, Fabian J *et al.*; AWI-Gen and the H3Africa Consortium. Kidney damage and associated risk factors in rural and urban sub-Saharan Africa (AWI-Gen): a cross-sectional population study. *Lancet Glob Health* 2019;**7**:e1632–43.
22. Wade AN, Maposa I, Agongo G *et al.*; AWIGen and the H3Africa Consortium. Diabetes care cascade and associated factors in 10 700 middle-aged adults in four sub-Saharan African countries: a cross-sectional study. *BMJ Open* 2023;**13**:e069193.
23. Chikwati RP, Mahyoodeen NG, Jaff NG *et al.* Cardiometabolic disease risk factors in pre- and postmenopausal women from four sub-Saharan African countries: a cross-sectional study. *Maturitas* 2023;**172**:60–68.
24. Micklesfield LK, Munthali R, Agongo G *et al.*; as members of AWI-Gen and the H3Africa Consortium. Identifying the prevalence and correlates of multimorbidity in middle-aged men and women: a cross-sectional population-based study in four African countries. *BMJ Open* 2023;**13**:e067788.
25. Wagner RG, Crowther NJ, Micklesfield LK *et al.* Estimating the burden of cardiovascular risk in community dwellers over 40 years old in South Africa, Kenya, Burkina Faso and Ghana. *BMJ Glob Health* 2021;**6**:e003499.
26. Tamburini FB, Maghini D, Oduaran OH *et al.*; H3Africa AWI-Gen Collaborative Centre. Short- and long-read metagenomics of urban and rural South African gut microbiomes reveal a transitional composition and undescribed taxa. *Nat Commun* 2022;**13**:926.
27. Maghini DG, Oduaran OH, Wirbel J *et al.* Expanding the human gut microbiome atlas of Africa. *bioRxiv*; doi:10.1101/2024.03.13.584859, 14 March 2024, preprint: not peer reviewed.
28. Boua PR, Soo CC, Debpuur C *et al.*; as members of AWI-Gen and the H3Africa Consortium. Prevalence and socio-demographic correlates of tobacco and alcohol use in four sub-Saharan African countries: a cross-sectional study of middle-aged adults. *BMC Public Health* 2021;**21**:1126.
29. Boua PR, Brandenburg J-T, Choudhury A *et al.* Novel and known gene-smoking interactions with cIMT identified as potential drivers for atherosclerosis risk in West-African populations of the AWI-Gen study. *Front Genet* 2019;**10**:1354.
30. Piga N-N, Boua PR, Soremekun C *et al.* Genetic insights into smoking behaviours in 10,558 men of African ancestry from continental Africa and the UK. *Sci Rep* 2022;**12**:18828.
31. Boua PR, Brandenburg J-T, Choudhury A *et al.*; H3Africa Consortium. Genetic associations with carotid intima-media thickness link to atherosclerosis with sex-specific effects in sub-Saharan Africans. *Nat Commun* 2022;**13**:855.
32. Choudhury A, Brandenburg J-T, Chikowore T *et al.*; H3Africa Consortium. Meta-analysis of sub-Saharan African studies provides insights into genetic architecture of lipid traits. *Nat Commun* 2022;**13**:2578.
33. Singh S, Choudhury A, Hazelhurst S *et al.* Genome-wide association study meta-analysis of blood pressure traits and hypertension in sub-Saharan African populations: an AWI-Gen study. *Nat Commun* 2023;**14**:8376.
34. Brandenburg J-T, Chen WC, Boua PR *et al.* Genetic association and transferability for urinary albumin-creatinine ratio as a marker of kidney disease in four Sub-Saharan African populations and non-continental individuals of African ancestry. *Front Genet* 2024;**15**:1372042.
35. Ceballos FC, Hazelhurst S, Clark DW *et al.* Autozygosity influences cardiometabolic disease-associated traits in the AWI-Gen sub-Saharan African study. *Nat Commun* 2020;**11**:5754.
36. Mulder NJ, Adebiyi E, Alami R *et al.*; H3ABioNet Consortium. H3ABioNet, a sustainable pan-African bioinformatics network for human heredity and health in Africa. *Genome Res* 2016;**26**:271–77.
37. Mulder NJ, Adebiyi E, Adebiyi M *et al.*; H3ABioNet Consortium, as members of the H3Africa Consortium. Development of Bioinformatics Infrastructure for Genomics Research. *Glob Heart* 2017;**12**:91–98.

© The Author(s) 2025. Published by Oxford University Press on behalf of the International Epidemiological Association.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (https://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.
International Journal of Epidemiology, 2025, 54, 1–8
https://doi.org/10.1093/ije/dyae173
Cohort Profile

Downloaded from https://academic.oup.com/ije/article/54/1/dyae173/7997654 by Cynthia Warren user on 14 April 2025