*International Journal of Epidemiology*, 2019, 366–367g
doi: 10.1093/ije/dyy261
Advance Access Publication Date: 6 December 2018
Data Resource Profile

Data Resource Profile

# Data Resource Profile: Cardiovascular H3Africa Innovation Resource (CHAIR)

**Mayowa O Owolabi,[1]* Onoja M Akpa,[1,2] Felix Made,[3,4] Sally N Adebamowo ,[5] Akinlolu Ojo,[6] Dwomoa Adu,[7] Ayesha A Motala,[8] Bongani M Mayosi,[9] Bruce Ovbiagele,[10] Clement Adebamowo,[5,11,12] Bamidele Tayo,[13] Charles Rotimi,[14] Rufus Akinyemi,[1] Mulugeta Gebregziabher,[15] Fred Sarfo,[16] Kolawole W Wahab,[17] Rulan S Parekh,[18] Mark E Engel,[19] Chisala Chisala,[19] Emmanuel Peprah,[20] George Mensah,[21] Ken Wiley,[22] Jennifer Troyer[23] and Michèle Ramsay[4]; as members of the CVD Working Group of the H3Africa Consortium**

[1]Center for Genomic and Precision Medicine, [2]Department of Epidemiology and Medical Statistics, College of Medicine, University of Ibadan, Ibadan, Nigeria, [3]Epidemiology and Surveillance Section, National Health Laboratory Services, Braamfontein, Gauteng Region, South Africa, [4]Sydney Brenner Institute for Molecular Bioscience and Division of Human Genetics, University of the Witwatersrand, Johannesburg, South Africa, [5]Department of Epidemiology and Public Health, and Greenebaum Comprehensive Cancer Center, University of Maryland School of Medicine, Baltimore, MD, USA, [6]Clinical Research and Global Health Initiatives, University of Arizona Health Sciences, Tucson, AZ, USA, [7]School of Medicine and Dentistry, University of Ghana, Accra, Ghana, [8]Department of Diabetes and Endocrinology, University of KwaZulu-Natal, Durban, South Africa, [9]Department of Medicine, Groote Schuur Hospital and University of Cape Town, Cape Town, South Africa, [10]Department of Neurology, University of California San Francisco, San Francisco, CA, USA, [11]Institute of Human Virology, University of Maryland School of Medicine, Baltimore, MD, and Institute of Human Virology, Abuja, Nigeria, [12]Center for Bioethics and Research, Department of Medicine, Ibadan, Nigeria, [13]Department of Preventive Medicine and Epidemiology, Loyola University Chicago Stritch School of Medicine, Maywood, IL, USA, [14]Center for Research on Genomics and Global Health, NHGRI, NIH, Bethesda, MD, USA, [15]Department of Public Health Sciences, Medical University of South Carolina, Charleston, SC, USA, [16]Department of Medicine, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana, [17]Department of Medicine, University of Ilorin, Ilorin, Nigeria, [18]Department of Pediatrics, Medicine and Epidemiology, Hospital for Sick Children, Research Institute, University Health Network and University of Toronto, Toronto, ON, Canada, [19]Division of Cardiology, Department of Medicine, University of Cape Town, Cape Town, South Africa, [20]Global Health Program, Department of Social and Behavioral Sciences, New York University, NY, USA, [21]National Heart, Lung, and Blood Institute, [22]Division of Genomic Medicine, and [23]Division of Genome Sciences, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA

*Corresponding author. Department of Medicine and Center for Genomic and Precision Medicine, University of Ibadan, Ibadan, Nigeria. E-mail: mayowaowolabi@yahoo.com

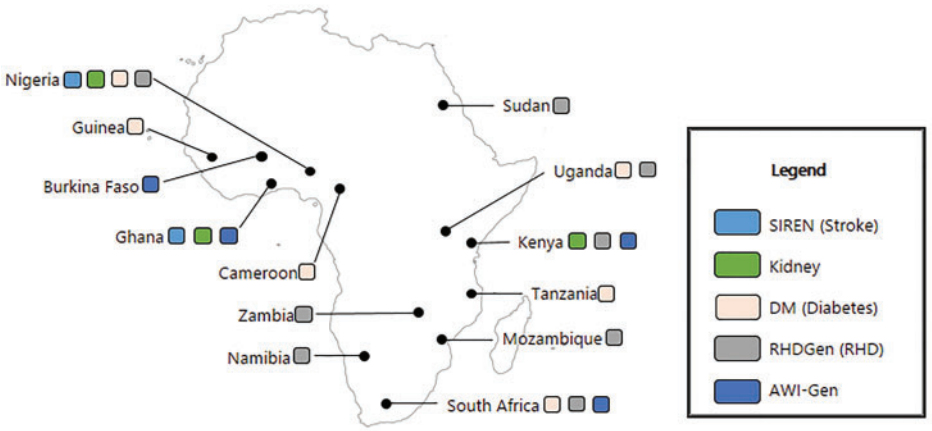
Editorial decision 18 October 2018; Accepted 30 October 2018

## Data resource basics

Low- and middle-income countries (LMIC) constitute the majority of the world's population and bear more than 80% of the global burden of cardiovascular disease (CVD).[1,2] The recent increases in CVD globally are also reflected in LMIC, where the prevalence of overall deaths from CVD was 28% in 2001[3] and premature CVD mortality was 37% in 2015.[4] The paucity of data regarding the drivers of the CVD epidemic and contextualized solutions is, in part, because less than 10% of the global research resources and facilities for implementation are found in LMIC.[5,6] Therefore LMIC are particularly disadvantaged in dealing with the CVD burden with targeted interventions. at the individual and community levels. due to the lack of data availability.

A key development poised to improve the understanding and control of CVD in LMICs in Africa is the Human Heredity and Health in Africa (H3Africa) Consortium, of which the H3Africa Cardiovascular Disease Working Group (CVDWG) is a core component.[6,7] The CVDWG includes six of the H3Africa projects spread across 13 African countries (Figure 1) . The six projects include: the African Collaborative Center for Microbiome and Genomics Research (ACCME), which also collects data on CVD; the Genomic and Environmental Risk Factors for Cardiometabolic Disease in Africans (AWI-Gen); the burden, spectrum and aetiology of type 2 diabetes mellitus in sub-Saharan Africa (DM group); the H3Africa Kidney Disease Research Network (Kidney group); the Genetics of Rheumatic Heart Disease Network (RHDGen); and the Stroke Investigative Research and Educational Network (SIREN).

The CVDWG is focused on developing genomic research and training infrastructure that will enable Africa to benefit from the genomic revolution. Specifically, the CVDWG is building the Cardiovascular H3Africa Innovation Resource (CHAIR), which aims to include >55 000 participants, with 30 044 participants in the first phase, as the largest cohort of continental Africans for studies on genomic and environmental contributions to CVD in Africa. CHAIR comprises phenotype data relevant to the cardiovascular disease spectrum, including sociodemographic, anthropometric, clinical, biochemical (blood and urine biomarkers) and genomic data as well as stored biological samples for future research.




| | | SIREN (Stroke) | | Kidney (CKD) | | DM (Diabetes | | RHDGen (RHD) | | AWI-Gen | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Country | Total | Cases | Controls | Cases | Controls | Cases | Controls | Cases | Controls | Cases | Controls |
| Burkina Faso | 2082 | - | - | - | - | - | - | - | - | - | 2082 |
| Cameroon | 117 | - | - | - | - | 117 | - | - | - | - | - |
| Ghana | 6666 | 733 | 736 | 530 | 2833 | - | - | - | - | - | 1834 |
| Guinea | 74 | - | - | - | - | 74 | - | - | - | - | - |
| Kenya | 2555 | - | - | 3 | 17 | - | - | 265 | 269 | - | 2001 |
| Mozambique | 358 | - | - | - | - | - | - | 182 | 176 | - | - |
| Namibia | 481 | - | - | - | - | - | - | 149 | 332 | - | - |
| Nigeria | 7826 | 1155 | 1187 | 985 | 3596 | 195 | 579 | 30 | 99 | - | - |
| South Africa | 7880 | - | - | - | - | 1022 | 600 | 508 | 610 | - | 5140 |
| Sudan | 340 | - | - | - | - | - | - | 173 | 167 | - | - |
| Tanzania | 572 | - | - | - | - | 572 | - | - | - | - | - |
| Uganda | 780 | - | - | - | - | 498 | - | 69 | 213 | - | - |
| Zambia | 313 | - | - | - | - | - | - | 93 | 220 | - | - |
| **Total** | **30044** | **1888** | **1923** | **1518** | **6446** | **2478** | **1179** | **1469** | **2086** | **-** | **11057** |

**Figure 1.** Map of Africa showing location or study sites of the participants from the six H3Africa studies, based on recruitment as at 30 September 2016.

**Table 1.** Participating H3Africa Consortium Projects that constitute CHAIR

| Project (PI reference) | Descriptive title | Study design | Countries involved in recruitment | Current sample size (target sample size)/sources |
|---|---|---|---|---|
| ACCME (Adebomowo)[8] | African Collaborative Center for Microbiome and Genomics Research | Prospective cohort study | Nigeria | Population-based sample of 11 700 female participants (targeted) |
| AWI-Gen (Ramsay and Sankoh)[9] | Genomic and Environmental Risk Factors for Cardiometabolic Disease in Africans | Cross-sectional: population-based cohort of adults | Burkina Faso<br>Ghana<br>Kenya<br>South Africa | Community-based (randomized by household and unrelated individuals) 11 057 (targeted: ~12 000) participants |
| DM Group (Motala)[10] | Burden, spectrum and aetiology of type 2 diabetes in sub-Saharan Africa | Cros-sectional: clinical and population-based cohort | Cameroon<br>Guinea<br>Malawi<br>Nigeria<br>South Africa<br>Tanzania<br>Uganda | Health facility-based recruitment of cases and community(population)-based recruitment of controls 3657 (targeted: ~6000 DM cases and ~6000 DM-free controls) |
| Kidney Group (Adu and Ojo)[11] | H3Africa Kidney Disease Research Network | Prospective and population-based cohort | Ghana<br>Ethiopia<br>Kenya<br>Nigeria | Cases: outpatient clinics and patient admissions<br>Controls: outpatient clinics<br>7964 (targeted: ~4000 kidney disease cases and ~4000 kidney disease-free controls) |
| RHDGen (Mayosi)[12] | Genetics of rheumatic heart disease (RHDGen) Network | Case-control cohort and family trios | Kenya<br>Mozambique<br>Namibia<br>Nigeria<br>South Africa<br>Sudan<br>Uganda<br>Zambia | Hospital-based recruitment of cases and community-based controls with no valvular heart disease 3555 (targeted: ~2700 RHD cases, ~2700 RHD-free controls and ~200 family trios) |
| SIREN (Owolabi and Ovbiagele)[13,14] | Stroke Investigative Research and Educational Network | Prospective and population-based cohort | Ghana<br>Nigeria | Hospital-based cases, community and hospital-based controls 3811 (~3000 stroke cases and ~3000 stroke-free controls) |

PI, Principal investigator (www.h3a.org).

## Data collected

### Datasets and population coverage

CHAIR is the combined resource of six independent research projects (Table 1; Supplementary Table 1, available as Supplementary data at *IJE* online). In this first phase, combined data from 30 044 participants of both genders and across age groups, from 15 African countries, have been combined using a careful harmonization process to ensure data equivalency. It is anticipated that the CHAIR resource will have data on close to 100 CVD-related phenotypes and genomic data for >55 000 participants. At the current wave of data collection, data completeness is >90% for most important CVD risk factors and almost 100% for stored samples (Table 2).

Participants include cases (selected by disease status) and controls (specific disease-free controls and cross-sectional population-based controls) and are resident in Burkina Faso, Cameroon, Ghana, Guinea, Kenya, Mozambique, Namibia, Nigeria, South Africa, Sudan, Tanzania, Uganda or Zambia (Figure 1). Recruitment strategies and CVD phenotypes that were measured are similar but not identical across the projects. Most projects have also published their research protocols.[8,15–19] Each study had its own inclusion and exclusion criteria, and

**Table 2.** Key variables being collected across different projects in CHAIR

| | ACCME | | AWI-Gen | | DM Group | | Kidney group | | RHDGen | | SIREN | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $n = 11\,700$ | % Complete | $n = 11\,057$ | % Complete | $n = 3657$ | % Complete | $n = 7964$ | % Complete | $n = 3555$ | % Complete | $n = 3811$ | % Complete |
| Age distribution | >18 years | 100.0 | 40–60 years | 100.0 | ≥25 / ≥18 years | 100.0 | 0–74 years | 100.0 | Paediatrics and adult | 100.0 | >18–100 years | 100.0 |
| Sex | Female only | 100.0 | M/F | 100.0 | M/F | 100.0 | M/F | 100.0 | M/F | 100.0 | M/F | 100.0 |
| Anthropometrics: | | | | | | | | | | | | |
| Weight (kg) | X | 99.9 | X | 100.0 | X | 100.0 | X | 100.0 | X | 98.5 | X | 87.9 |
| Height (m) | X | 99.8 | X | 100.0 | X | 100.0 | X | 99.9 | X | 95.4 | X | 91.0 |
| Waist circumference (cm) | X | 99.8 | X | 100.0 | X | 100.0 | X | 99.7 | | | X | 94.6 |
| Hip circumference (cm) | X | 99.9 | X | 100.0 | X | 100.0 | | | | | X | 95.5 |
| General health: | | | | | | | | | | | | |
| Smoking/tobacco | X | 99.8 | X | 100.0 | X | 100.0 | X | 100.0 | | | X | 98.3 |
| Alcohol | X | 99.8 | X | 99.9 | X | 100.0 | X | 99.9 | | | X | 98.9 |
| Cancer history | X | 99.8 | X | 99.9 | X | 100.0 | X | 99.8 | | | X | 98.5 |
| Diet | X | 99.7 | X | 99.8 | X | 100.0 | | | | | X | 95.4 |
| Exercise | X | 99.8 | X | 100.0 | X | 100.0 | | | | | X | 98.3 |
| Cardiovascular health: | | | | | | | | | | | | |
| Blood pressure | X | 99.9 | X | 100.0 | X | 100.0 | X | 100.0 | X | 86.8 | X | 96.6 |
| Atrial fibrillation (ECG) | | | | | | | X | 100.0 | X | 12.1 | X | 98.3 |
| Stroke and stroke-free status | X | 99.8 | X | 99.9 | X | 100.0 | X | 100.0 | | | X | 100.0 |
| Myocardial infarction | X | 99.8 | | 77.0 | X | 100.0 | | | | | X | 98.6 |
| Blood collection for biomarkers | | | | | | | | | | | | |
| Lipid profile | | | X | 100.0 | X | 100.0 | | | | | X | 99.0 |
| Fasting plasma Glucose | | | X | 100.0 | X | 100.0 | | | | | X | 60.6 |
| HbA1c | | | X | | X | 100.0 | | | | | X | 41.7 |
| Insulin | | | X | 100.0 | X | 100.0 | | | | | | |
| Infection history | | | | | | | | | | | | |
| TB infection | | | X | 99.9 | X | 100.0 | X | 99.9 | | | X | 98.6 |
| HIV status | X | 100.0 | X | 99.9 | X | 100.0 | X | 99.9 | | | X | 98.6 |
| Malaria | | | X | 99.9 | X | 100.0 | X | 100.0 | | | X | 98.4 |
| Urine collection | | | X | | X | 100.0 | | | | | | |
| Albumin (microalbumin) | | | X | 99.8 | X | 100.0 | X | 100.0 | | | | |
| Total protein | | | X | 99.9 | | | X | 99.9 | | | | |
| Creatinine | | | X | 99.9 | X | 100.0 | X | 100.0 | | | X | 80.0 |
| Samples to be stored | | | | | | | | | | | | |
| DNA | X | 99.9 | X | 99.9 | X | 100.0 | | | X | 99.0 | X | 99.9 |
| Buffy coat | X | 99.9 | X | 99.9 | | | | | X | 70.0 | X | 99.9 |
| Plasma | X | 99.9 | X | 99.9 | X | 100.0 | | | X | 82.0 | X | 99.9 |
| Serum | X | 99.9 | X | 99.9 | X | 100.0 | | | X | 79.9 | X | 99.9 |
| Urine | X | 99.7 | X | 99.8 | X | 100.0 | | | | | X | |

M/F, male/female.

most studies started recruitment in 2012 and completed recruitment in 2017. Most studies are planning longitudinal follow-up of participants. Study instruments used by the participant studies were administered by well-trained fieldworkers and research professionals, using validated approaches to harmonize data.

## Survey frequency

One of the two cross-sectional studies is a population-based study (both have plans for longitudinal follow-up in the next wave of data collection), and four studies are population-based longitudinal cohorts or case-control disease-based studies. Studies with longitudinal data collections within the first phase of H3Africa (mid-2012 to mid-2017) are shown in Supplementary Table 2, available as Supplementary data at *IJE* online; and one is planning follow-up in a second phase. For instance, the Kidney group used a case-control study design and plan annual follow-up visits of the cases over a period of 5 years, to assess kidney disease progression and other clinical events.[18] The SIREN study followed up patients with stroke at 1, 3, 6, 9 and 12 months after discharge from hospital.[8] The ACCME group conducted a prospective cohort study where demographic and clinical data were collected at baseline and then during follow-up visits every 6 months for 2 years.[9,15] The AWI-Gen study is a cross-sectional population-based study of older adults which will be developed into a longitudinal cohort with follow-up after approximately 5 years from the first recruitment, to assess progression of, or transition to, cardiometabolic disease states.[16] The DM group is a cross-sectional study comprising two arms: a 'clinic/case-series' arm which includes participants with known diabetes attending a health facility; and in parallel, from the same geographical area, are drawn participants for the 'population survey' arm.[17] RHDGen is a case-control study with cases with rheumatic disease recruited in a hospital-based setting and community-based controls, which can be contacted for further data collection.[19] The CHAIR studies have obtained broad consent from most participants for the reuse of data and samples, and participants either been consented to be approached for future studies or will be reconsented when approached again for longitudinal data collection.

## Measures

Each participating study in CHAIR has independent but related robust case report forms (CRFs) for phenotype and laboratory data collection. Study CRFs consist of validated measures for capturing data on several phenotypes. For example, the Questionnaire for Verifying Stroke-Free Status (QVSFS)[8] is used for determining stroke-free status with baseline blood pressure measured according to the American Heart Association Council on High Blood Pressure Research.[9] Standing height, body weight, waist circumferences and obesity are measured in accordance with the recommendations of the World Health Organization (WHO) Multinational Monitoring of Trends in Cardiovascular Disease (MONICA) project.[10]

A total of 67 phenotypes including: sociodemographic details (age, sex, country of residence, ethnicity, marital status, educational attainment and occupational status); socioeconomic status; medical history (infection with HIV, TB or malaria; history of hypertension, diabetes, stroke, myocardial infarction, cancer and other infectious and non-infectious diseases and traits); lifestyle data (tobacco and alcohol exposure); family health history; and biomarkers of CVD relevance in blood and urine; are included in the CHAIR dataset (Table 2; and Supplementary Table 3, available as Supplementary data at *IJE* online). The questionnaires used for collecting data on these phenotypes are based on existing validated questionnaires and differ, in some cases, between participating studies. Where possible, when comparable measures were used, data have been harmonized across datasets. For instance, hypertension data were harmonized and defined as sustained systolic BP $\geq$ 140 mmHg or diastolic BP $\geq$ 90 mmHg, history of hypertension, or taking antihypertensive medications. Diabetes mellitus was harmonized and defined as history of diabetes mellitus, use of medications for diabetes mellitus, fasting glucose levels $\geq$126 mg/dL (or $\geq$200 mg/dL if failing to fast) and/or HBA1c >6.5%. Obesity was based on body mass index (kg/m$^2$) calculated and classified as underweight (<18.5), normal (18.5–24.9), overweight (25–29.9) or obese (30+) or based on waist-to-hip ratio tertiles.

All studies collect blood and/or other tissues or body fluids (e.g. urine) for DNA extraction and biomarker analysis. Several aliquots of each DNA sample from all participants are stored in an H3Africa Biorepository in Nigeria, Uganda or South Africa. Genome-wide variation data will be available for all studies. Four studies have used the H3Africa SNP genotyping array (~2.3 M SNP Illumina Array enriched for common variation across African populations). RHDGen has used the Illumina Omni 2.5 M array, and the DM group will use a different array. African-based imputation panels are available to harmonize the disparate genomic data types.

## Data resource use

Although each project has its own research objectives, the shared resource will be able to tackle cross-cutting research

questions and test hypotheses which the individual projects are underpowered to explore. For instance, we have collected and harmonized data on potential risk factors, including age, setting (urban or rural), weight, height, waist and hip circumference across the entire CHAIR cohort. In addition, we have stored biological samples including urine, serum, plasma and DNA from the majority of the participants. We have also developed an inclusive model for data harmonization which makes it possible to investigate multiple hypotheses, to address a wide spectrum of CVD risk factors in disease causation or influence. Therefore, CHAIR provides the opportunity to examine genomic associations with cardiovascular phenotypes across various African regions and urban/rural settings. This enhances the potential for discovery, validation and generalizability. CHAIR offers a pan-African perspective with unique opportunities for collaboration, including sharing of data on controls, provision of validation cohorts, meta-analyses, capacity-building for genomic research and development of infrastructure for CVD epidemiological, genomic, biobanking and bioinformatics research.

The findings from the CHAIR research promise to inform the development of cost-effective diagnostic, preventive and therapeutic interventions at individual, health care system and population levels.[12,13,20] Such interventions could be implemented through robust advocacy efforts involving all stakeholders working within and across African countries and regions. This would enhance the emergence of a healthy and economically vibrant continent.

Baseline characteristics for 30 044 participants of CHAIR with available data as of 30 September 2016 are presented in Table 3. Overall, 57.0% were women. The proportion with hypertension (defined as systolic blood pressure ≥140 mm Hg or diastolic blood pressure ≥90 mm Hg or confirmed diagnosis of hypertension by a health care worker or current use of antihypertensive medication) was 48.8% among men and 47.4% among women. Based on the 2017 guidelines for hypertension diagnosis (i.e. systolic blood pressure ≥120 mm Hg or diastolic blood pressure ≥80 mm Hg in addition to confirmed diagnosis of hypertension by a health care worker or current use of antihypertensive medication), the proportion with hypertension was 74.8% among men and 69.6% among women, and the proportion with obesity [body mass index (BMI) ≥30 kg/m$^2$] was 10.8% among men and 32.8% among women (Table 3). The proportion with hypertension and obesity varied among projects, influenced by individual participant selection. For instance, the proportion with hypertension was highest in the SIREN study on stroke, and

**Table 3.** Preliminary baseline characteristics for 30 044 CHAIR participants from five studies as at 30 September 2016

| Studies (number of participants) | Sex (%) | *n* | Age in years (SD) | Proportion with obesity[a] | Proportion with hypertension (definition 1: ≥140/90 mmHg)[a] | Proportion with hypertension (definition 2: ≥120/80 mmHg)[a] |
|---|---|---|---|---|---|---|
| AWI-Gen (11 057) | Male (45.2) | 4999 | 50.61±6.96 | 7.8 | 34.4 | 64.3 |
| | Female (54.8) | 6058 | 50.83±6.95 | 31.2 | 37.9 | 60.7 |
| Subtotal | | 11 057 | 50.73±6.95 | 20.6 | 36.3 | 62.3 |
| DM (3657) | Male (32.8) | 1199 | 49.01±16.69 | 18.5 | 55.2 | 77.1 |
| | Female (67.2) | 2458 | 50.71±14.46 | 51.8 | 64.8 | 80.5 |
| Subtotal | | 3657 | 50.16±15.24 | 40.9 | 61.7 | 79.4 |
| Kidney (7964) | Male (42.7) | 3398 | 47.40±14.69 | 11.7 | 60.8 | 82.7 |
| | Female (57.3) | 4566 | 46.40±13.70 | 26.9 | 52.2 | 73.5 |
| Subtotal | | 7964 | 46.83±14.14 | 20.4 | 55.9 | 77.5 |
| RHDGen (3555) | Male (34.2) | 1217 | 34.00±12.81 | 8.0 | 21.6 | 67.1 |
| | Female (65.8) | 2338 | 38.78±13.78 | 27.9 | 21.4 | 58.1 |
| Subtotal | | 3555 | 37.14±13.64 | 21.1 | 21.5 | 61.1 |
| SIREN (3811) | Male (55.3) | 2108 | 57.43±12.89 | 14.4 | 75.7 | 90.1 |
| | Female (44.7) | 1703 | 58.62±14.10 | 32.5 | 78.6 | 90.7 |
| Subtotal | | 3811 | 57.96±13.45 | 22.4 | 77.9 | 90.4 |
| Combined CHAIR (30 044) | Male(43.0) | 12 921 | 49.17±13.35 | 10.8 | 48.8 | 74.8 |
| | Female(57.0) | 17 123 | 48.76±13.02 | 32.8 | 47.4 | 69.6 |
| Grand total | | 30 044 | 48.94±13.17 | 23.4 | 48.0 | 71.8 |

Percentages of completed data for the baseline characteristics were: 100.0% for sex; 100.0% for age; 97.0% for height; 97.1% for weight; 98.8% for systolic blood pressure; 98.8% for diastolic blood pressure.

SD, standard deviation.

[a]These rates may have been affected by the design and disease focus of the individual study.

similarly the proportion with obesity was highest in the type 2 diabetes study (Table 3).

The CVDWG has published position papers[6,7] describing the current knowledge of CVD in Africa and also articulating the opportunities and challenges for future CVD research in Africa.[6,7,21–30] To ensure accurate exclusion of stroke mimics, a pictographic version of the Questionnaire for Verifying Stroke-Free Status with superior diagnostic properties including 98% certainty for determining stroke-free status, which was validated in three languages commonly spoken in West Africa (including Ghana), was developed.[31] Such free but robust instruments for stroke diagnosis, a leading cause of CVD in Africa, could be deployed in the follow-up of study participants across the CHAIR consortium. Some cardiovascular disease-specific findings from some of the collaborating studies[32] suggest the need for a unique cardiovascular risk calculator for Africans. In a recent report, one of the studies (SIREN)[14,33] identified leading risk factors for stroke in the continent, that is: hypertension, dyslipidaemia, regular meat consumption, elevated waist-to-hip ratio, diabetes, low green leafy vegetable consumption, stress, salt added at the table, cardiac disease, physical inactivity and tobacco use. These factors will be investigated further using the CHAIR resource.[14,33]

The AWI-Gen study found significant sex differences in the prevalence of hypertension in four of their six study centres and regional differences in prevalence, awareness and control of hypertension.[34] This suggests the need for the implementation of regionally appropriate and context-driven targeted interventions for hypertension in Africa. In a cohort of sexually active HIV-negative adult women with no previous history of cervical abnormalities, cervical cancer or total abdominal hysterectomy, the ACCME study found a 21% prevalence of persistent human papilloma virus (HPV) infection.[15] There was also a reported history of physician-diagnosed hypertension (15%), diabetes (2%), hypercholesterolaemia (4%) or heart disease (0.3%).[15] These results provide important data for studies of associations between these characteristics and HPV-associated cancers including cervical cancer.[15] The RHDGen project, in a 5-year follow-up study of latent Rheumatic heart diseases (RHD) among school pupils and using the World Heart Federation (WHF) echocardiographic criteria, found that latent RHD resolves to normal in nearly half of school pupils.[19] It is therefore a necessity to repeat echocardiography in cases of latent RHD before considering an intervention.

A comprehensive list of publications from the CVDWG is available on the H3Africa website: [http://h3africa.org/links/publications].

### Strengths and weaknesses

The main strength of the CHAIR cohort is its large anticipated final sample size of over 55 000 African participants, which is powered to validate and examine many genetic risk factors for quantitative traits including BMI and lipid levels, where genetic markers have low or moderate effects.

With a sample size of at least 50 000, the study will have >95% power to detect associations with quantitative traits with an effect size (beta value) of 0.2 for single nucleotide polymorphisms (SNPs) with a minor allele frequency (MAF) of 0.05, and >90% power to detect a beta value of 0.1 for SNPs with MAF 0.20. In case-control studies with 6000 cases and 6000 controls, there will be >95% power to detect an odds ratio (OR) of 1.2 (MAF 0.05) and >90% to detect an OR of 1.1 (MAF 0.20). Other strengths include the harmonized high-quality phenotype data and the availability of genome-wide genotyping data on all participants and access to DNA samples.

All studies also have community engagement, training and capacity building components and have produced freely available protocols, guidelines, training materials and analysis tools that can be used in similar studies and are adapted to the African context. Also, although electronic medical records are scarce and diseases and vital registries are incomplete in Africa in general, some of the hospital-based studies could do passive follow-up from hospital records and the AWI-Gen study, which is nested in HDSS, can refer to verbal autopsies in the cases of deaths.

A limitation of CHAIR is the lack of standardization of data collection tools across all six studies. The harmonization process, however, has ensured that data are compared in a meaningful way for the cardinal CVD risk factors. Furthermore, most studies will also have genomic data and, for those not using the H3Africa Illumina SNP array, genetic analyses will use imputation strategies to develop a common dataset to analyse pooled phenotypic data using advanced meta-analytical techniques.[35,36]

## Data resource access

This CHAIR resource is available to researchers (members and non-members of the H3Africa Consortium) worldwide according to the H3Africa guidelines and policies [https://www.h3africa.org/consortium/documents]. Data will be submitted to the European Genome-phenome Archive (EGA) and will have a publication embargo of 12 months, or following the first publication, whichever is the soonest. The harmonized phenotype data of CHAIR will also be stored in a dedicated REDCap database[37] and will be linked through unique identifiers to genomic data. Genotype and phenotype data will be available at

H3ABioNet and the European Genome-phenome Archive (EGA).[38] DNA samples will be accessed from the H3Africa biorepositories.[39] A managed access model has been implemented for data and samples, and they will be available after approval by the H3Africa Data and Biospecimen Access Committee (DBAC). Imputation panels will be available to ensure that a rich joint genomic dataset can be developed for meta-analyses. Please contact the H3Africa Coordinating Centre for more information [h3africa@mail.nih.gov].

**Profile in a nutshell**

- CHAIR is the largest harmonized resource of cardiovascular diseases in Africans with genomic, environmental and phenotypic data from participants across the entire cardiovascular cascade. CHAIR is the first and only such initiative in Africa with the potential for addressing cross-cutting research questions with substantial power and validity.
- With an initial target of >30 000 (male and female) participants, the resource comprises data on adults from 15 African countries, recruited by six H3Africa-member studies including the ACCME, AWI-Gen, DM group, Kidney group, RHDGen and SIREN at baseline enrolment, and with a subset having longitudinal follow-up data.
- Follow-up within CHAIR has been study-specific as follows: ACCME (every 6 months for 2 years), Kidney group (annually), SIREN (every 3 months for 1 year). Studies like AWI-Gen and DM group have planned for follow-up in the next wave of data collection. RHDGen is a cross-sectional study with no planned follow-up.
- Associated biological samples are available for future studies on genetic and epigenetic determinants of cardiovascular health, diseases and morbidity among individuals of African ancestry.
- CHAIR can be accessed through the H3Africa Coordinating Centre.

## Supplementary Data

Supplementary data are available at *IJE* online.

## Funding

The harmonization process for the CHAIR consortium is supported by the administrative supplement 3U54HG007479-03S1. SIREN is funded by the NIH (Grant no. U54HG007479). The AWI-Gen Collaborative Centre is funded by the NIH (Grant no. U54HG006938) as part of the H3Africa Consortium. The ACCME cohort is funded by the National Institutes of Health (NIH/NHGRI grant 1U54HG006947). The DM study (burden, spectrum and aetiology of type 2 diabetes in sub-Saharan Africa) is funded by the Wellcome Trust (Grant no. WT 099316AIA). The RHDGen Network is funded by the Wellcome Trust (099313/Z/12/Z and 099313/B/12/Z). The H3Africa Kidney Disease Research Consortium is funded by the NIH (Grant no. 5U54HG006939-05). C.R. is supported by the intramural programme of the NHGRI/NIH at the Center for Research on Genomics and Global Health (CRGGH); the CRGGH is also supported by NIDDK, CIT and the NIH Office of the Director. The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the Wellcome Trust.

## Acknowledgements

We thank Romuald Boua for performing the power calculations using Quanto. The authors kindly thank Harry Wedel for illustrating Figure 1.

## Author Contributions

All authors contributed to study concept, design and data acquisition. O.M.A. and F.M. are CHAIR statisticians who contributed to data harmonization and analysis. M.O.O., M.R., O.M.A. and F.M. contributed to interpretation of results and drafting of the manuscript. All authors contributed to critical revision of the manuscript and approval of the final draft.

**Conflict of interest:** None declared.

## References

1. Roth GA, Johnson C, Abajobir A. Global, regional, and national burden of cardiovascular diseases for 10 causes, 1990 to 2015. *J Am Coll Cardiol* 2017;**70**:1–25.
2. Yusuf S, Rangarajan S, Teo K *et al*. Cardiovascular risk and events in 17 low-, middle-, and high-income countries. *N Engl J Med* 2014;**371**:818–27.
3. Gaziano TA, Bitton A, Anand S, Abrahams-Gessel S, Murphy A. Growing epidemic of coronary heart disease in low- and middle-income countries. *Curr Probl Cardiol* 2010;**35**:72–115.
4. World Health Organization. *Cardiovascular Diseases (CVDs)*. Updated May 2017. http://www.who.int/mediacentre/factsheets/fs317/en/ (12 November 2017, date last accessed).
5. Owolabi MO, Bower JH, Ogunniyi A. Mapping Africa's way into prominence in the field of neurology. *Arch Neurol* 2007;**64**:1696–700.
6. Owolabi MO, Mensah GA, Kimmel PL *et al*. Understanding the rise in cardiovascular diseases in Africa: harmonising H3Africa genomic epidemiological teams and tools. *Cardiovasc J Afr* 2014;**25**:134–36.
7. Peprah E, Wiley K, Troyer J *et al*. Building a platform to enable NCD research to address population health in Africa: CVD Working Group Discussion at the Sixth H3Africa Consortium Meeting in Zambia. *Glob Heart* 2016;**11**:165–70.
8. Adebamowo SN, Dareng EO, Famooto AO *et al*.; ACCME Research Group as part of the H3Africa Consortium. Cohort

Profile: African Collaborative Center for Microbiome and Genomics Research's (ACCME) Human Papillomavirus (HPV) and Cervical Cancer Study. *Int J Epidemiol* 2017;**46**:1745.
9. 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.
10. Ekoru K, Young EH, Adebamowo C *et al*. H3Africa multi-centre study of the prevalence and environmental and genetic determinants of type 2 diabetes in sub-Saharan Africa: study protocol. *Glob Health Epidemiol Genom* 2016;**1**:e5.
11. Osafo C, Raji YR, Burke D *et al*. Human Heredity and Health (H3) in Africa kidney disease research network: a focus on methods in sub-saharan Africa. *Clin J Am Soc Nephrol* 2015;**10**: 2279–87.
12. Zühlke L, Engel ME, Lemmer CE *et al*. The natural history of latent rheumatic heart disease in a 5 year follow-up study: a prospective observational study. *BMC Cardiovasc Disord* 2016;**16**:46.
13. Akpalu A, Sarfo FS, Ovbiagele B *et al*. Phenotyping stroke in sub-saharan Africa: stroke investigative research and education network (SIREN) phenomics protocol. *Neuroepidemiology* 2015;**45**:73–82.
14. Oga EA, Brown JP, Brown C *et al*. Recurrence of cervical intraepithelial lesions after thermo-coagulation in HIV-positive and HIV-negative Nigerian women. *BMC Women's Health* 2016;**16**:25.
15. Pickering TG, Hall JE, Appel LJ *et al*. Subcommittee of Professional and Public Education of the American Heart Association Council on High Blood Pressure Research. Recommendations for blood pressure measurement in humans and experimental animals. Part 1: blood pressure measurement in humans. *Hypertension* 2005;**45**:142–61.
16. WHO. *MONICA Manual Part III: Population Survey*. Section 1: Population survey data component. 1997. http://www.thl.fi/publications/monica/manual/part3/iii-1.htm (12 November 2017, date last accessed).
17. Owolabi M, Olowoyo P, Miranda JJ *et al*. Gaps in hypertension guidelines in low- and middle-income versus high-income countries: a systematic review. *Hypertension* 2016;**68**:1328–37.
18. Owolabi M, Miranda JJ, Yaria J, Ovbiagele B. Controlling cardiovascular diseases in low and middle income countries by placing proof in pragmatism. *BMJ Glob Health* 2016;**1**:e000105.
19. Johnson W, Onuma O, Owolabi M, Sachdev S. Stroke: a global response is needed. *Bull World Health Organ* 2016;**94**:634–34A.
20. Akinyemi RO, Ovbiagele B, Akpalu A *et al*. Stroke genomics in people of African ancestry: charting new paths. *Cardiovasc J Afr* 2015;**26**:S39–49.
21. Gibbons GH, Sampson UK, Cook NL, Mensah GA. NHLBI perspectives on the growth of heart, lung, blood and sleep conditions in Africa: global and domestic insights, challenges and opportunities. *Cardiovasc J Afr* 2015;**26**:S18–20.
22. Hall MD, Dufton AM, Katso RM, Gatsi SA, Williams PM, Strange ME. Strategic investments in non-communicable diseases (NCD) research in Africa: the GSK Africa NCD Open Lab. *Cardiovasc J Afr* 2015;**26**:S15–17.
23. Mensah GA, Peprah EK, Sampson UK, Cooper RS. H3Africa comes of age. *Cardiovasc J Afr* 2015;**26**:S3–5.
24. Mensah GA, Roth GA, Sampson UK *et al*. Mortality from cardiovascular diseases in sub-Saharan Africa, 1990-2013: a systematic analysis of data from the Global Burden of Disease Study 2013. *CVJA* 2015;**26**:S6–10.
25. Mocumbi AO. Rheumatic heart disease in Africa: is there a role for genetic studies? *Cardiovasc J Afr* 2015;**26**:S21–26.
26. Owolabi MO, Karolo-Anthony S, Akinyemi R *et al*. The burden of stroke in Africa: a glance at the present and a glimpse into the future. *Cardiovasc J Afr* 2015;**26**:S27–38.
27. Raji Y, Mabayoje O, Bello T. Familial clustering of risk factors for cardiovascular disease among first-degree relatives of patients with chronic kidney disease in a sub-Saharan African population. *Cardiovasc J Afr* 2015;**26**:S11–14.
28. Sampson UK, Engelgau MM, Peprah EK, Mensah GA. Endothelial dysfunction: a unifying hypothesis for the burden of cardiovascular diseases in sub-Saharan Africa. *Cardiovasc J Afr* 2015;**26**:S56–60.
29. Wonkam A, Makani J, Ofori-Aquah S *et al*. Sickle cell disease and H3Africa: enhancing genomic research on cardiovascular diseases in African patients. *Cardiovasc J Afr* 2015;**26**:S50–55.
30. Sarfo F, Gebregziabher M, Ovbiagele B *et al*. Stroke investigative research educational networks. Multilingual validation of the questionnaire for verifying stroke-free status in West Africa. *Stroke* 2016;**47**:167–72.
31. Owolabi MO, Akpa OM, Agunloye AM. Carotid IMT is more associated with stroke than risk calculators. *Acta Neurol Scand* 2016;**133**:442–50.
32. Owolabi MO, Sarfo F, Akinyemi F *et al*. Dominant modifiable risk factors for stroke in Ghana and Nigeria (SIREN): a case-control study. *Lancet Glob Health* 2018;**6**:e436–46.
33. Kengne AP, Mayosi BM. Modifiable stroke risk factors in Africa: lessons from SIREN. *Lancet Glob Health* 2018;**6**:e363–64.
34. Gómez-Olivé FX, Ali SA, Made F *et al*. Stark 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.
35. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;**42**:377–81.
36. Roetzheim RG, Freund KM, Corle DK. Analysis of combined data from heterogeneous study designs: a methodological proposal from the patient navigation research program. *Clin Trials* 2012;**9**:176–87.
37. Blettner M, Sauerbrei W, Schlehofer B, Scheuchenpflug T, Friedenreich C. Traditional reviews, meta-analyses and pooled analyses in epidemiology. *Int J Epidemiol* 1999;**28**:1–9.
38. Lappalainen I, Almeida-King J, Kumanduri V *et al*. The European genome-phenome archive of human data consented for biomedical research. *Nat Genet* 2015;**47**:692–95.
39. Beiswanger CM, Abimiku A, Carstens N *et al*. Accessing biospecimens from the H3Africa consortium. *Biopreserv Biobank* 2017;**15**:95–98.