Neuro
epidemiology

Neuroepidemiology 2015;45:73–82
DOI: 10.1159/000437372

Received: March 26, 2015
Accepted: June 28, 2015
Published online: August 19, 2015

Phenotyping Stroke in Sub-Saharan Africa: Stroke Investigative Research and Education Network (SIREN) Phenomics Protocol

Albert Akpalu[a, b] Fred Stephen Sarfo[c, d] Bruce Ovbiagele[e] Rufus Akinyemi[g] Mulugeta Gebregziabher[e] Reginald Obiako[h] Lukman Owolabi[i] Kwamena Sagoe[b] Carolyn Jenkins[e] Oyedunni Arulogun[j] Sheila Adamu[c] Lambert T. Appiah[c] Martin A. Adadey[a] Francis Agyekum[a] Joseph A. Quansah[c] Yaw B. Mensah[a] Abiodun M. Adeoye[j] Arti Singh[c] Aridegbe O. Tosin[g] Osimhiarherhuo Ohifemen[g] Abubabkar A. Sani[h] Eric Tabi-Ajayi[h] Ibinaiye O. Phillip[h] Suleiman Y. Isah[i] Nasir A. Tabari[i] Aliyu Mande[i] Atinuke M. Agunloye[j] Godwin I. Ogbole[j] Joshua O. Akinyemi[j] Onoja M. Akpa[j] Ruth Laryea[b] Sylvia Ezinne Melikam[j] Dorcas Adinku[c] Ezinne Uvere[j] Nina-Serena Burkett[e] Gregory F. Adekunle[j] Salaam I. Kehinde[j] Paschal C. Azuh[h] Abdul H. Dambatta[i] Naser A. Ishaq[i] Donna Arnett[f] Hemant K. Tiwari[f] Dan Lackland[e] Mayowa Owolabi[j, k]
on behalf of SIREN as part of the H3Africa Consortium

[a]Korle-Bu Teaching Hospital, Accra, [b]University of Ghana Medical School, Accra, Ghana; [c]Kwame Nkrumah University of Science and Technology, [d]Komfo Anokye Teaching Hospital, Kumasi, and [e]Medical University of South Carolina, Charleston, S.C., and [f]University of Alabama at Birmingham, Birmingham, Ala., USA; [g]Federal Medical Centre and Sacred Heart Hospitals, Abeokuta, [h]Ahmadu Bello University, Zaria, [i]Aminu Kano Teaching Hospital, Kano, [j]University of Ibadan, and University College Hospital, Ibadan, and [k]WFNR-Blossom Center for Neurorehabilitation, Ibadan, Nigeria

Key Words
Stroke · Phenomics · Sub-Saharan Africa · SIREN · Protocol

Abstract
Background: As the second leading cause of death and the leading cause of adult-onset disability, stroke is a major public health concern particularly pertinent in Sub-Saharan Africa (SSA), where nearly 80% of all global stroke mortalities occur, and stroke burden is projected to increase in the coming decades. However, traditional and emerging risk factors for stroke in SSA have not been well characterized, thus limiting efforts at curbing its devastating toll. The Stroke Investigative Research and Education Network (SIREN) project is aimed at comprehensively evaluating the key environmental and genomic risk factors for stroke (and its subtypes) in SSA while simultaneously building capacities in phenomics,

A.A. and F.S.S. contributed equally to this work.

0251–5350/15/0452–0073$39.50/0

E-Mail karger@karger.com
www.karger.com/ned

Dr. Mayowa Owolabi
College of Medicine
University of Ibadan, and Department of Medicine
University College Hospital, Ibadan 200001 (Nigeria)
E-Mail mayowaowolabi@yahoo.com

biobanking, genomics, biostatistics, and bioinformatics for brain research. ***Methods:*** SIREN is a transnational, multicentre, hospital and community-based study involving 3,000 cases and 3,000 controls recruited from 8 sites in Ghana and Nigeria. Cases will be hospital-based patients with first stroke within 10 days of onset in whom neurovascular imaging will be performed. Etiological and topographical stroke subtypes will be documented for all cases. Controls will be hospital- and community-based participants, matched to cases on the basis of gender, ethnicity, and age (±5 years). Information will be collected on known and proposed emerging risk factors for stroke. ***Study Significance:*** SIREN is the largest study of stroke in Africa to date. It is anticipated that it will shed light on the phenotypic characteristics and risk factors of stroke and ultimately provide evidence base for strategic interventions to curtail the burgeoning burden of stroke on the sub-continent.

Introduction

There is a rising epidemic of stroke in Sub-Saharan Africa (SSA) directly resulting from the epidemiologic transition [1–4]. This epidemiologic transition is driven by an exploding but neglected burden of non-communicable diseases including vascular risk factors such as systemic arterial hypertension, diabetes mellitus, and dyslipidemia culminating in stroke [3, 4]. The WHO estimates indicate that death from stroke and disability adjusted life years lost to stroke are at least 7 times higher in low- and middle-income countries – mostly in Africa – than in high-income countries [5]. The age-standardized stroke incidence and prevalence rates in Africa with a current population of over 1 billion are up to 316 per 100,000 and up to 981 per 100,000 population respectively with a 3-year mortality rate of 84% [1, 4]. Furthermore, stroke affects a relatively younger age group and the productive workforce in developing than in developed economies [1, 6] and has been shown to exert its severest toll in the physical, psychoemotional, cognitive, and social domains [7–13].

To tackle the burden of stroke and unravel the reasons for the escalating epidemic, there is an urgent need to determine the role played by environmental and genetic factors to the risk and outcomes of stroke and its subtypes in SSA. The traditional risk factors for stroke including hypertension, diabetes mellitus, cigarette smoking, sedentary lifestyle, unhealthy diet, excessive alcohol intake, dyslipidemia, and truncal obesity are predicted to be on the increase among African populations but their relative contributions to stroke predisposition have not been explicitly quantified. This is because considerable evidence from several observational studies shows a remarkable heterogeneity in the profile of risk factors, prevalence, incidence, types, subtypes, natural history, and outcome of stroke across different geographical locations and different ethnicities. For instance, significantly higher age-standardized stroke mortality rates have been reported among males than females in Ghana and Nigeria [14]. Furthermore, notable differences in prevalence and mortality from stroke were noted among rural Tanzanians compared with rural South Africans, plausibly due to differential exposures of lifestyle risk factors and health transitions in these 2 populations [15]. Again, data from the INTERSTROKE study [16] showed that while ischemic and hemorrhagic strokes accounted for 66 and 34% of strokes in Africa, these figures were 91 and 9% respectively in high-income countries. Thus, several questions remain unanswered regarding the true proportions, etiologic sub-types, risk factor profiles, and genetic underpinnings of ischemic and hemorrhagic stroke in SSA. Although stroke is a discrete phenotype, it is the clinical culmination of several complex and interacting biological processes; hence, any step towards lessening its burden and formulating successful tailored prevention programs in SSA will require a deep understanding of the relative contributions of its predisposing risk factors.

Thus far, the underlying risk factors for stroke in SSA have not been sufficiently characterized. In addition, there is a paucity of expertise, experience, and infrastructure for sustainable cutting-edge stroke research in SSA. To this end, the Systematic Investigation of Blacks with Stroke-Phenomics (SIBS-Phenomics), an arm of the Stroke Investigative Research and Education Network (SIREN) project, will involve several geographically dispersed and ethnically diverse sites in SSA countries of Ghana and Nigeria with the aim of properly evaluating the key risk factors for stroke (and its subtypes) in these countries. It is hoped that the outcome of this study will provide accurate and reliable stroke phenotypes for a subsequent genomic study (SIBS-Genomics) which aims to explore the genetic risk factors for stroke (and its subtypes).

The aim of this manuscript is to describe the protocol for the evaluation of risk factors and the clinical, etiological, and pathological types of stroke in the current largest study on stroke in SSA.

Table 1. Study design

Cases	Control
Sampling frame All patients admitted to medical ward, neurology or neurosurgical ward, ICU, emergency room, or reviewed in an outpatient clinic in the participating centres fulfilling the case definition for stroke will be logged and invited to participate in the study. Those who meet the eligibility criteria and provide informed consent will be included in the study.	Controls will be primarily community-based, but hospital-based controls will also be accepted. Hospital-based controls may be attendants or relatives of another (non-stroke) patient, or patients admitted to the hospital or visiting the hospital for conditions or procedures not related to stroke or TIA. Each control will be matched for sex and age (±5 years). There will be at least a control (with or without CV risks) for every case recruited. Where there is a significant representation from multiple ethnic groups, controls will also be matched with cases based on ethnicity.
Inclusion criteria Patients who meet the following criteria will be included in the study (1) Patients attending/admitted to study centres with first stroke (see definition below); (2) Within 8 days of current symptom onset or 'last seen without deficit'; (3) CT or MRI is planned within 10 days of symptom onset (note: beyond 10 days, ischemic and hemorrhagic strokes may be difficult to distinguish with great certainty); (4) Adult aged 18 years or older (younger age group: difficulties with obtaining consent).	Persons, with no clinical evidence of stroke, with or without cardiovascular risk factors.
Exclusion criteria Patients who meet the following criteria will be excluded from the study (1) Patients unable to communicate because of severe stroke, aphasia, or dementia and without a valid surrogate respondent (spouse or first degree relative who has lived with patient within the last year); (2) Extra-axial hemorrhage, tumor or brain abscess; (3) Subarachnoid hemorrhage; (4) Current hospitalization for coronary heart disease; (5) Unable to provide consent and no surrogate available; (6) A known previous history of stroke.	(1) A known previous history of stroke; (2) Current hospitalization for acute coronary syndrome/myocardial infarction; (3) Unable to provide consent and no surrogate Available.

Methods

Specific Aims and Hypotheses

SIREN has 4 broad objectives and underlying hypotheses: our first objective is to evaluate the qualitative and quantitative contributions of traditional and novel socio-demographic, clinical, and radiologic risk factors to stroke occurrence, type, and outcome among individuals of Black race in SSA. This will be done via a cross-national, multicentre case-control study. The underlying hypothesis for this sub-objective is that the distribution of pre-specified risk factors will contribute to differences in the occurrence, type, and outcome of stroke among individuals of Black race in SSA. Second, we propose to compare socio-demographic, clinical, and radiographic risk factors for stroke among individuals of Black race in SSA to cohorts of individuals of Black race in the United States. We hypothesize that differences in the pattern of socio-demographic, clinical, and radiographic risk factors may account for differences in stroke occurrence, type, and outcome between Blacks in SSA and Blacks in the United States. Third, we will survey and elucidate the knowledge, attitude, beliefs, and practices of individuals of Black race in SSA about stroke and its risk factors, performance of genetic testing, and participation in research studies. Our hypothesis for this sub-objective is that systematic community engagement using established qualitative research methodologies will result in information that can be used to educate and mobilize the community regarding the nature of stroke, role of genetic testing, and benefits of involvement in research studies. And fourth, our objective is to generate comprehensive and accurate phenomic data on stroke patients and controls that will assure reliable phenotyping for a concurrent genomics project (SIBS-Genomics) focused on evaluating genetic risk factors for stroke (and its subtypes). This is underscored by the hypothesis that coordinated expertise and experience leveraged from seasoned investigators and resources within 2 ongoing NIH-funded stroke studies in SSA will facilitate reliable stroke phenotyping.

Study Design

SIBS-Phenomics will be based on a case-control study design where every case will be matched to a control based on demographics (age, sex, ethnicity) using inclusion and exclusion criteria shown in table 1. Matching on these factors will minimize the potential confounding effect of these variables on the relationship between stroke and the main environmental and genetic risk factors. This matched case-control design is based on a risk-set sam-

Fig. 1. Map showing the location of participating centres.

pling framework, where cases and controls are sampled from a cohort of Africans who meet the inclusion criteria. Subjects with known outcomes (cases and controls) are measured for past exposures to estimate the relative risk of 'disease' since direct estimation of absolute risk is difficult [17]. By exploiting the setup of risk sets, case-control designs have been suggested as efficient ways of studying the association between outcomes and exposure [18–20].

Study Sites

Eight collaborating centres in Nigeria and Ghana will use a standardized and comprehensive protocol based on the design of the WHO-STEPS Stroke Surveillance [21], INTERSTROKE [16], and REGARDS [22] protocols with relevant adaptations/modifications (to reflect the uniqueness of the African people, culture, lifestyles and the environment). This will be used to accrue detailed demographic, clinical, psychosocial, and neuroimaging data from 3,000 stroke patients in comparison with 3,000 demographically-matched controls. To ensure accurate phenotyping, the stroke patients will be evaluated in hospitals well equipped with neuroimaging (CT and/or MRI), vascular, and cardiac (ultrasonography, electrocardiography (ECG), ECHO) facilities, while control recruitment will be enhanced through community stroke awareness programs. At each site, a community engagement team will be constituted. This team will collaborate with the Community Advisory Board made up of community leaders, faith-based leaders and key opinion leaders, stroke survivors, public health leaders, to engage communities at churches, mosques, market places, community meetings – to give talks on stroke awareness. At these meetings, the objectives of SIREN will also be highlighted and an invitation given to interested community members to serve as 'community-based controls' after screening them to rule out stroke using an 8-item questionnaire for verifying stroke free status (QVSFS). Potential 'hospital-based controls' will be approached at out-patient clinics/in-patient settings by research assistants and stroke as a diagnosis ruled out using the 8-item QVSFS those without stroke will be recruited as controls as shown in table 1.

There are 6 participating sites in Nigeria including the University College Hospital and the WFNR-Blossom Centre for Neurorehabilitation, Ibadan; Federal Medical Centre and Sacred Heart Hospital, Abeokuta; Aminu Kano Teaching Hospital, Kano; Ahmadu Bello University, Zaria, and 2 sites in Ghana namely Korle-Bu Teaching Hospital, Accra and Komfo Anokye Teaching Hospital, Kumasi (fig. 1 and 2). The demographic characteristics and the participating hospitals/study sites are shown in table 2. The participating study sites are predominantly tertiary teaching hospitals (where stroke cases can be adequately investigated) situated such that stroke referrals are received from urban, semi-urban, and rural communities. Our community engagement core will also ensure that stroke cases in rural areas get referred to tertiary centers involved in the study by increasing the awareness on stroke.

Evaluation of Cases and Controls

The SIBS questionnaires (cases and controls) (available on request) will be used to accrue basic demographic and lifestyle data including ethnicity and native language of the subjects and their parents, socioeconomic status, dietary patterns, routine physical activity, stress (validated INTERSTROKE instrument), depression, cigarette smoking, and alcohol use (table 3). Others are infections, exposure to internal and external pollution, and personal medical and drug history. Also, cardiovascular and anthropometric (e.g. waist-hip ratio and weight) measurements will be obtained. Hypertension will be based on blood pressure levels and will be classified in the traditional categories of awareness, treatment, and control. In addition, resistant hypertension will be determined based on blood pressure values and the number of hypertension medications. Diabetes mellitus and dyslipidemia will

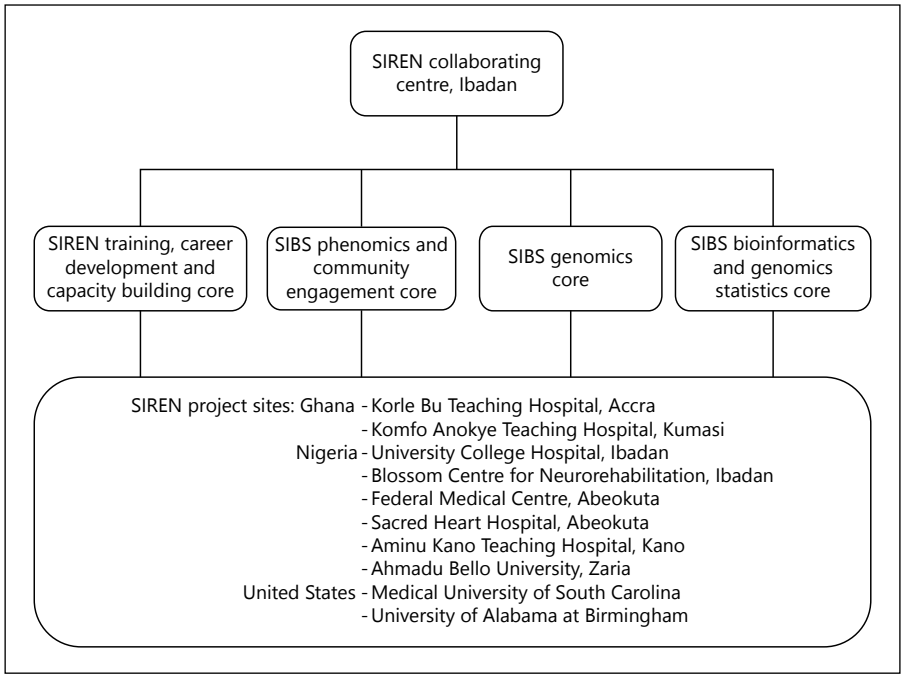


Fig. 2. Organogram of the SIREN.

be based on fasting glucose levels and/or HBA_{1c} and fasting lipid profile. A detailed neurologic evaluation will be conducted to assess neurologic deficits and determine stroke severity using the National Institute of Health Stroke Severity Score, Stroke Levity Scale [23], modified-Rankin Scale, and modified Barthel's Index scores on admission and serially monitored until discharge. Data on the therapy the patient received will also be documented as well as complications arising during hospital admission and vital status at discharge of patient.

In addition, the following evaluations will be performed: brain neuroimaging (CT or MRI of brain), ECG and/or Holter monitoring, transthoracic and/or trans-oesophageal echocardiography (for all cases of ischemic stroke, mandatory for all suspected cardioembolic stroke), and carotid Doppler USS (for all cases of ischemic stroke, anterior circulation). Ischemic stroke will be typed clinically using the Oxford Community Stroke Project criteria [24] and etiological sub-types will be defined using the Trial of Org 10172 in Acute Stroke Treatment [25] and Atherosclerosis-small vessel disease-cardiac source-other [26] criteria (fig. 3). Intracerebral hemorrhage will be classified etiologically into structural, medication-related, amyloid angiopathy, systemic/other disease, hypertension and undetermined causes as proposed by Meretoja et al. [27] The annotation and image markup on ClearCanvas software which we developed for phenotyping stroke patients based on radiological, clinical, and laboratory features, will also capture the above information including volume and location of bleeds, which is important for this classification.

Radiologic typing of stroke will be carried out at 3 levels of adjudication: first by the site PI-neurologist/neuroradiologist; second by site PIs-neurologists/neuroradiologists from other sites for verification; third by a stroke phenotyping committee with all site PIs and neuroradiologists (for ambiguous cases) and by carrying out random checks on all classified cases.

The 8-item QVSFS [28] has been validated in a sub-study (manuscript in preparation) and will be used to ascertain stroke free status in control subjects. Socio-demographic, lifestyle, medical, and drug history as well as anthropometric and vascular risk factors will be evaluated in controls.

Follow-Up Strategy

Stroke survivors will be followed up at dedicated neurology out-patient clinics upon discharge from medical wards in the participating institutions. The schedule for follow-up is at months 1, 3, 6, 9, and 12-post discharge. At each of these visits, patients will be assessed for control of risk factors, presence or absence of functional neurologic deficits using scales outlined earlier and taken through a neuro-psychological battery including the Montreal Cognitive Assessment [29] and an assessment of the impact of stroke on the quality of life and cognitive domains using the Health-Related Quality of Life in Stroke Patients [12]. The overall strategy for recruitment and follow-up of SIREN protocol is depicted in figure 4.

Blood Sampling Strategy

Thirty five millilitres of venous sample will be obtained from cases and controls for the following subsequent analyses: sera and plasma separation for novel and emerging biomarkers of stroke and also for buffy coat preparation for DNA extraction to be used in SIBS genomics. Refer to the online supplementary information (for all online suppl. material, see www.karger.com/doi/10.1159/000437372) for sample processing protocol.

Sample Size Estimation and Power Justification

Prevalence of stroke in Africa is up to 315 per 100,000 population [2]. Given the current prevalence and incidence rate of stroke in Africa, a sample size of 3,000 cases and 3,000 controls will pro-

Table 2. Socio-demographic characteristics of SIREN study sites

Siren sites	Population (number in 2011)	Tribes/ethnic groups	Hospital/university coordination site	Other
Nigeria[1]	177,155,754[1]	Nigeria, Africa's most populous country, is composed of more than 250 ethnic groups; the following are the most populous and politically influential: Hausa and Fulani 29%, Yoruba 21%, Igbo (Ibo) 18%, Ijaw 10%, Kanuri 4%, Ibibio 3.5%, Tiv 2.5%[1]	0.53 beds/1,000 population (2004)[1]	About 50% urban (2011)[1] Life expectancy at birth: 52.6 years[1] Languages: English (official), Hausa, Yoruba, Igbo (Ibo), Fulani, over 500 additional indigenous languages[1] Religion: Muslim 50%, Christian 40%, indigenous beliefs 10%[1] Literacy: 61.3%[1]
Southern Nigeria sites				
– Ibadan	2.949 million[1]	Yoruba	University college hospital Ibadan: 1st teaching hospital in Nigeria with 850 beds tertiary center with several community care centers and Blossom center for neuro-rehabilitation: 1st neuro-rehabilitation hospital in East, West and Central Africa	Capital of Oyo State and 3rd largest city in Nigeria
– Abeokuta		Yoruba	Federal medical center: 250 bed tertiary center with strong relationships to community care centers within and outside Abeokuta	Capital of Ogun State
Northern Nigeria sites				
– Kano	3.375 million[1]	Hausa and Fulani tribes are dominant	Aminu Kano teaching hospital and tertiary referral center with 78 beds	Largest commercial center in northern Nigeria with fertile plains for agriculture
– Zaria		Hausa and Fulani tribes are dominant	Ahmadu Bello university teaching hospital with 768 beds	
Ghana[2]	25,758,108[2]	Akan 47.5%, Mole-Dagbon 16.6%, Ewe 13.9%, Ga-Dangme 7.4%, Gurma 5.7%, Guan 3.7%, Grusi 2.5%, Mande-Busanga 1.1%, other 1.6% (2010 census)[2]	0.9 beds/1,000 population (2011)[2]	Low middle income economy. About 52% urban (2011)[2] Life expectancy: 65.75 years[2] Languages: English (official), Asante 14.8%, Ewe 12.7%, Fante 9.9%, Boron (Brong) 4.6%, Dagomba 4.3%, Dangme 4.3%, Dagarte (Dagaba) 3.7%, Akyem 3.4%, Ga 3.4%, Akuapem 2.9%, other (includes English) 36.1% (2,000 census)[2] Religion: Christian 71.2% (Pentecostal/Charismatic 28.3%, Protestant 18.4%, Catholic 13.1%, other 11.4%), Muslim 17.6%, traditional 5.2%, other 0.8%, none 5.2% (2010 census)[2] Literacy: 71.5%[2]
Southern Ghana site				
– Accra	2.573 million (2011)[2]	Akan, Ga-Dangme and Ewe are dominant	Korle Bu teaching hospital, a tertiary referral center with 1,500 beds. Stroke cases are referred from all levels of health care	Dedicated 20 bed multidisciplinary stroke unit
Northern Ghana site				
– Kumasi	2.019 million (2011)[2]	Akan sub-group of Ashanti are dominant	Komfo Anokye teaching hospital, a tertiary referral center with 1,000 beds	Dedicated 10 bed multidisciplinary stroke unit

[1] Index Mundi with data from CIA World Factbook Data: Nigeria Demographic Profile 2014. Retrieved on February 14, 2014 from: http://www.indexmundi.com/nigeria/demographics_profile.html.

[2] Index Mundi with data from CIA World Factbook Data: Ghana Demographic Profile 2014. Retrieved on February 14, 2014 from: http://www.indexmundi.com/ghana/demographics_profile.html.

Table 3. Variables to be studied in the SIBS-Phenomics (adapted from the INTERSTROKE study) [10]

Variable	Description
Personal and lifestyle factors	Ethnicity, tobacco use, exposure to environmental tobacco, alcohol use, dietary habits (for estimation of intake of fruits, vegetables, fried foods, meats, fish, added salt, etc.), physical activity, sleep patterns, substance abuse (e.g. cocaine, amphetamine), neck manipulation/injury and pollution
Socioeconomic factors	Education, occupation, objects owned, total income
Psychosocial factors	Stress at work and home, social support, locus of control, depression, life events
Medical history	Hypertension, diabetes mellitus, smoking, hypercholesterolemia, atrial fibrillation, CVD, cancer, current medication use, other vascular disease, oral contraceptive use, hormone replacement, family history, migraine, menstrual and pregnancy history. Pre-morbid modified-Rankin Scale score. Herbal medication use
Physical measurements	Blood pressure measurement after stroke may be a biased assessment of premorbid blood pressure. Therefore, we will use a self-reported history of hypertension. However, we will also record blood pressure at baseline, heart rate, height, weight, waist and hip circumferences at the time of the interview
Neurological assessment	Clinical assessment, neuroimaging, location of lesion, results of diagnostic tests and confirmation of stroke sub-type
Stroke subtype	Based on clinical assessment and neuroimaging (e.g. CT of brain) – information on from neurovascular, carotid Doppler and echocardiographic investigations
Neuroimaging	CT or MRI of brain (for all cases), ECG (for all cases), echocardiography (for all cases of suspected cardioembolic stroke), carotid Doppler USS (for all cases of ischemic stroke, anterior circulation)
Course in hospital and status at 30 days	Treatments given, outcomes in hospital, modified-Rankin Scale at 30 days
Blood tests	Lipid profile (including lipoprotein (a), HbA1c), fasting blood sugar, 2-hour post prandial, WBC, markers of infection and inflammation (chlamydia pneumonia, hs-CRP, ESR), homocysteine and folate, fibrinogen
Genetic markers	As done in SIBS genomics, DNA to be extracted

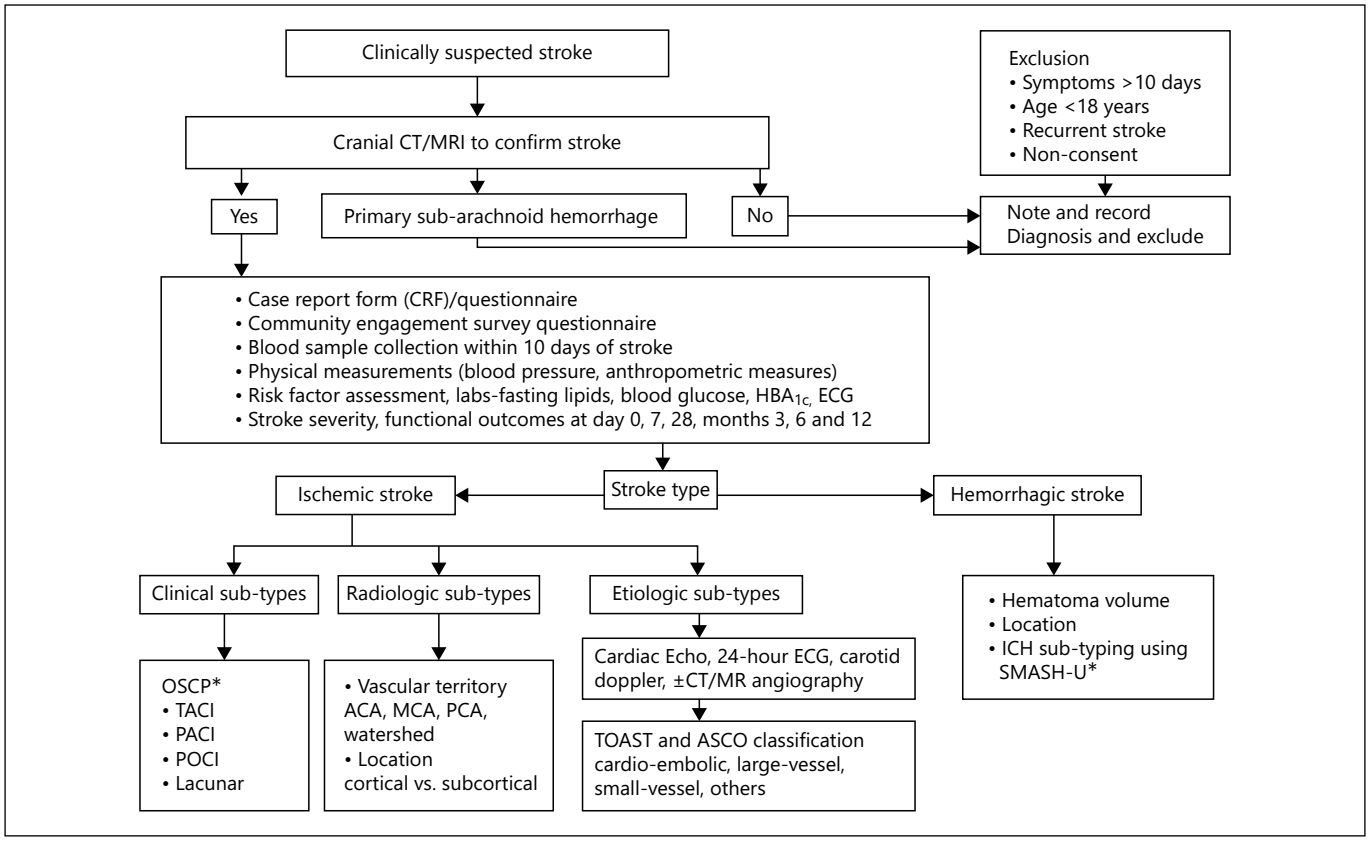
vide sufficient sample size for the intended study to get a power of at least 80% for detecting an effect size of odds ratio at least 1.4, allowing for several types of exposures (categorical, count, or continuous) and the interactions among these exposures. A conditional logistic regression model will be used for analysis with a binary response variable (stroke status) and discrete or continuous covariates. To detect a change in the response variable from a proportion at baseline of 10–14%, which corresponds to an odds ratio of 1.4; a total sample size of 6,000 (3,000 cases and 3,000 controls) will lead to at least 80% power at a 0.05 significance level. An R-squared of 20% [30] was assumed to account for covariate adjustment in the multiple logistic regression model. This sample size is sufficient to detect odds ratios of at least 1.4 or more in subgroup analysis, for instance, to test an interaction between sub-type of stroke and exposures. Thus, a sample size of 6,000 will be sufficient to detect a range of effect sizes accounting for missing data.

Statistical Analysis

We will assess the association between exposure variables (e.g. routine physical activity (yes/no), stress score (validated INTERSTROKE instrument), depression (yes/no), cigarette smoking (yes/no) and alcohol use (yes/no)) and case-control status using conditional logistic regression [31] with and without adjustment to potential confounders that are not used in the matching. The adjusted model will include selected covariates and their interactions depending on whether or not covariates are confounders, by whether they did not improve the model fit, or increased the standard error of the parameter estimate of the primary exposure. The final model will be assessed for collinearity and goodness of fit using residual analysis.

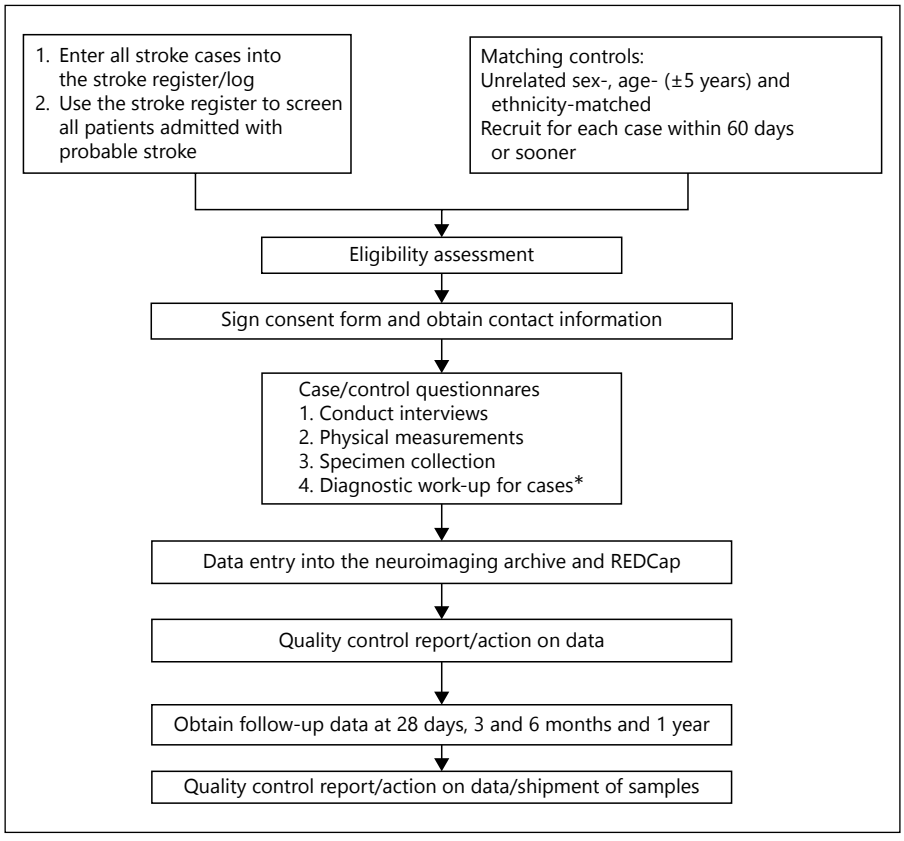
Calculation of Population Attributable Risks (PARs). We will calculate both attributable risk (AR) and PAR for each of the exposure variables listed in table 3. For instance, we will calculate the risk attributable to a unit increase in the stress score in the cases compared to the controls. We will use a model-based regression approach [32] to calculate the adjusted AR. The standard error of the adjusted AR will be estimated using the delta method or using the bootstrap. This calculation will be useful to infer the association between stroke and these exposure variables. Given that the final results are going to be based on the adjusted AR regression model, there are not issues of multiple comparisons that arise for exposures with more than 2 categories. We will combine data from Africa (SIBS) and the United States (REGARDS) [22] and we will fit logistic regression models with the outcome case-control status as a dependent variable and, each of the common exposures in both studies as independent variables with an interaction term by study type. The interaction term will allow us to estimate and test whether the odds ratio estimates corresponding to the common exposure variables are the same between the 2 studies. We will use the Wald test [33] for this purpose.

Missing Data and Recall Bias. We will handle missing data using several techniques including multiple imputation and maximum likelihood [31, 34] using SAS System version 9.4 (Cary, N.C.,



3

Fig. 3. Algorithm for investigating stroke patients. OCSP = Oxford Community Stroke Project; TACI = total anterior circulation infarction; PACI = partial anterior circulation infarction; POCI = posterior circulation infarction; ASCO = atherosclerosis-small vessel disease-cardiac source-other; TOAST = Trial of Org 10172 in Acute Stroke Treatment; ACA = anterior cerebral artery; MCA = middle cerebral artery; PCA = posterior cerebral artery; SMASH-U = structural, medication, amyloid angiopathy, systemic diseases, hypertension, underdetermined.



4

Fig. 4. SIREN protocol workflow for cases and controls.

USA). We will use missing data methods suggested for matched case-control data with missing covariate [35]. We will also use imputation techniques such as multiple imputation [35] and midpoint imputation [36] if the missing data mechanism is missing at random. To deal with missing categorical variables, we will use latent class multiple imputation [34]. Likelihood techniques will be used if the missing mechanism deviates from missing at random. Missing data assumptions will be determined based on the understanding of the missing data mechanism and tests to determine whether the missing data depends on observed data and sensitivity analysis [34] will be performed. To enable the controls to represent the same population as the cases, and thus reduce sampling bias, matching will be used. Matching enhances power, given the smaller sample size. Recall bias is ameliorated as data used in this study will be collected routinely prior to the knowledge of case/control definition.

Discussion

There are several innovative aspects to SIBS-Phenomics. First, SIBS-Phenomics will be the first cross-national and sub-regional investigation aimed at fully exploring the environmental factors that contribute to stroke in SSA. Second, SIBS-Phenomics will explore all risk factors for stroke: traditional (such as hypertension, dyslipidemia, smoking, among others – observed mostly among people of White race in developed countries), emerging (such as inflammation, infections, sleep habits among others identified in developed countries), and environmental (such as both traditions and lifestyle- unique to Black people in SSA). Third, SIBS-Phenomics, will for the first time compare stroke in Negros in SSA (in Ghana and Nigeria) with Americans of African ancestry participating in the REGARDS study [22]. Fourth, SIBS-Phenomics will ensure community sensitization to facilitate early presentation of stroke cases to hospital, thereby greatly minimizing referral bias as well as recruitment of controls from the community. This community engagement core within SIBS-Phenomics will be the first-of-its kind, a public outreach engagement initiative to evaluate and address perceptions about stroke and genomics by patients, caregivers, and local leaders. Doing this will help promote better uptake of medical regimen, adherence, behavioural modification, and participation in research studies by at-risk persons and the communities at large. It will also facilitate dissemination of results, as well as research uptake and impact. Finally, SIBS-Phenomics will be part of a central strategy in SSA to create intra-continental collaborations and build infrastructure, skills and capacity for research into stroke and cerebral disorders.

This undertaking may be challenged by the following: first, the recruitment of stroke cases within the hospital setting would mean that potentially patients with severe stroke may be recruited excluding patients with less severe strokes and those who may seek alternative therapies for stroke outside of hospital settings. Our community engagement strategy is designed to eliminate this bias. Second, even though all participating sites have at least one CT or MRI scanner in place, these investigative tools may break down and could limit our ability to accurately identify the type of stroke. However, privately operated CT/MRI scanners are available at all sites and would be resorted to should the need arise. Third, the funding provided for evaluation of a case-control pair is inadequate to thoroughly investigate all stroke patients. However, prudent fund management will be undertaken such that patients who can afford some of the investigations will be encouraged to pay, while support from the study will be targeted for those who cannot afford to pay for investigations to ensure recruitment and inclusion of all eligible stroke subjects and controls. Finally, storage of samples may be challenged by inadequate electrical power supply, a challenge we have approached by procuring alternative power supply mechanisms including solar powered inverter systems for our freezers.

Funding

The SIREN study is supported by U01 NS079179, and U54 HG007479 from the National Institutes of Health as part of the H3Africa Consortium.

Disclosure Statement

The authors have no conflict of interest to declare.

References

1 Feigin VL, Lawes CM, Bennett DA, Barker-Collo SL, Parag V: Worldwide stroke incidence and early case fatality reported in 56 population-based studies: a systematic review. Lancet Neurol 2009;8:355–369.
2 Owolabi MO: Taming the burgeoning stroke epidemic in Africa: stroke quadrangle to the rescue. West Indian Med J 2011;60:412–421.
3 Yusuf S, Reddy S, Ounpuu S, Anand S: Global burden of cardiovascular diseases: part II: variations in cardiovascular disease by specific ethnic groups and geographic regions and prevention strategies. Circulation 2001;104:2855–2864.

4 Owolabi MO, Akarolo-Anthony S, Akinyemi R, Arnett D, Gebregziabher M, Jenkins C, Tiwari H, Arulogun O, Akpalu A, Sarfo FS, Obiako R, Owolabi L, Sagoe K, Melikam S, Adeoye AM, Lackland D, Ovbiagele B; Members of the H3Africa Consortium: The burden of stroke in Africa: a glance at the present and a glimpse into the future. Cardiovasc J Afr 2015;26(2 suppl 1):S27–S38.
5 Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJ: Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. Lancet 2006;293:1747–1757.
6 Owolabi MO, Ugoya S, Platz T: Racial disparity in stroke risk factors: the Berlin-Ibadan experience; a retrospective study. Acta Neurol Scand 2009;119:81–87.
7 Owolabi MO: Determinants of health-related quality of life in Nigerian stroke survivors. Trans R Soc Trop Med Hyg 2008;102:1219–1225.
8 Owolabi MO: Health-related quality of life (HRQOL) measures: there are still many unanswered questions about human life. ScientificWorldJournal 2008;8:357–363.
9 Owolabi MO, Ogunniyi A: Profile of health-related quality of life in Nigerian stroke survivors. Eur J Neurol 2009;16:54–62.
10 Owolabi MO: What are the consistent predictors of generic and specific post-stroke health-related quality of life? Cerebrovasc Dis 2010;29:105–110.
11 Owolabi MO: Which is more valid for stroke patients: generic or stroke-specific quality of life measures? Neuroepidemiology 2010;34:8–12.
12 Owolabi MO: Psychometric properties of the HRQOLISP-40: a novel, shortened multiculturally valid holistic stroke measure. Neurorehabil Neural Repair 2010;24:814–825.
13 Owolabi MO, Bower JH, Ogunniyi A: Mapping Africa's way into prominence in the field of neurology. Arch Neurol 2007;64:1696–1700.
14 Mensah GA: Epidemiology of Stroke and high blood pressure in Africa. Heart 2008;94:697–705.
15 Walker R, Whiting D, Unwin N, Mugusi F, Swai M, et al: Stroke incidence in rural and urban Tanzania: a prospective, community-based study. Lancet Neurol 2010;9:786–792.
16 O'Donnell MJ, Xavier D, Liu L, et al: Risk factors for ischaemic and intracerebral haemorrhagic stroke in 22 countries (the INTERSTROKE study): a case-control study. Lancet 2010;376:112–123.
17 Breslow NE, Day NE: Statistical Methods in Cancer Research. Volume I – The Analysis of Case-Control Studies. Lyon, IARC Scientific Publications, 1980.
18 Oakes D: Survival times: aspects of partial likelihood (with discussion). Int Stat Rev 1981;49:235–264.
19 Thomas DC: Addendum to the paper by Liddel, McDonald and Thomas. J R Stat Soc Series B 1977;140:483–485.
20 Thomas DC: General relative-risk models for survival time and matched case-control analysis. Biometrics 1981;37:673–686.
21 Truelsen T, Heuschmann PU, Bonita R, et al: Standard method for developing stroke registers in low-income and middle-income countries: experiences from a feasibility study of a stepwise approach to stroke surveillance (STEPS stroke). Lancet Neurol 2007;6:134–139.
22 Howard VJ, Cushman M, Pulley L, et al: The reasons for geographic and racial differences in stroke study: objectives and design. Neuroepidemiology 2005;25:135–143.
23 Owolabi MO, Platz T: Proposing the stroke Levity Scale: a valid, reliable, simple, and time-saving measure of stroke severity. Eur J Neurol 2008;15:627–633.
24 Bamford J, Sandercock P, Dennis M, Burn J, Warlow C: Classification and natural history of clinically identifiable subtypes of cerebral infarction. Lancet 1991;337:1521–1526.
25 Kolominsky-Rabas PL, Weber M, Gefeller O, Neundoerfer B, Heuschmann PU: Epidemiology of ischemic stroke subtypes according to TOAST criteria: incidence, recurrence, and long-term survival in ischaemic stroke subtypes: a population-based study. Stroke 2001;32:2735–2740.
26 Amarenco P, Bogousslavsky J, Caplan LR, Donnan GA, Hennerici MG: New approach to stroke subtyping: the A-S-C-O (Phenotypic) classification of stroke. Cerebrovasc Dis 2009;27:502–508.
27 Meretoja A, Strbian D, Putaala J, Curtze S, Haapaniemi E, Mustanoja S, et al: SMASH-U: a proposal for etiologic classification of intracerebral hemorrhage. Stroke 2012;43:2592–2597.
28 Meschia JF, Brott TG, Chukwudelunzu FE, et al: Verifying the stroke-free phenotype by structured telephone interview. Stroke 2000;31:1076–1080.
29 Akinyemi RO, Allan L, Owolabi MO, Akinyemi JO, Ogbole G, Ajani A, Firbank M, Ogunniyi A, Kalaria RN: Profile and determinants of vascular cognitive impairment in African stroke survivors: the CogFAST Nigeria study. J Neurol Sci 2014;346:241–249.
30 Hsieh FY, Bloch DA, Larsen MD: A simple method of sample size calculation for linear and logistic regression. Stat Med 1998;17:1623–1634.
31 Gebregziabher M, Langholz B: A semiparametric missing-data-induced intensity method for missing covariate data in individually matched case-control studies. Biometrics 2010;66:845–854.
32 Little RJA, Rubin DB: Statistical Analysis with Missing Data. New York, John Wiley & Sons Inc., 2002.
33 Paik MC, Sacco RL: Matched case-control data analyses with missing covariates. Appl Statist 2000;49:145–156.
34 Gebregziabher M, DeSantis S: Latent class based multiple imputation approach for missing categorical data. J Statist Plann Inference 2010;140:3252–3262.
35 Greenland S: Model-based estimation of relative risks and other epidemiologic measures in studies of common outcomes and in case-control studies. Am J Epidemiol 2004;160:301–305.
36 Hosmer DW, Lemeshow S, et al: Applied Logistic Regression, New York, Wiley, 2000.