Core Phenotypes v2.0

Instruments

Purpose

A collection of measures to capture the core phenotypes related to African health and biomedical research, as determined by the H3Africa Phenotype Harmonisation Working Group (https://h3africa.org/index.php/consortium/working-groups/phenotype-harmonization/).

Guideline Description

The H3Africa core phenotypes can be used to collect essential phenotypes related to African health and biomedical research. These phenotypes were prioritised based on the overlap in phenotype data collection across multiple research projects being conducted across the African continent, in various disease fields. Phenotypes include information related to a participant's demographics, anthropometrics, tobacco and alcohol use, and disease history. The following document establishes guidelines (particularly applicable in Africa) on how to use the Core Phenotypes to collect detailed, relevant and harmonized phenotype and exposure data for research.

As listed below, the Core Phenotypes consist of 19 Instruments, labelled A to S:

Instrument	**Phenotypes**	**Instrument**	**Phenotypes**
A	Demographics	K	Diabetes History
B	Anthropometrics	L	Asthma
C	Blood Pressure	M	Dyslipidemia
D	Urine Test Results	N	Cancer History
E	Smoking Exposure	O	Kidney Disease History
F	Smoking Status	P	Stroke History
G	Alcohol Exposure	Q	CVD History
H	Alcohol Use	R	HIV
I	Drug Use	S	Infectious Disease History
J	Medication Log		

Important Notes

1. The toolkit employs branching logic, therefore, we recommend that it is completed in order, as some variables may or may not appear OR accept input based on the input of previously listed variables.

2. Some branching logic (specifically related to date of birth/age and current pregnancy) affects the display of items relevant to adult or paediatric participants across multiple instruments.

3. Any addition or removal of variables may also affect branching logic so editing of variables should be carefully positioned so as not to interrupt branching logic conditions with related variables.

4. Although not highlighted below, each instrument requires a collection date, which can be collected either manually or automatically.

5. Consistent codes are recommended for the identification of missing data, and these are incorporated into all Instruments discussed below. We recommend the use of ‘Temporarily unavailable’ for pending results in Instrument D and R.

6. Codes for Missing Data are specified below:

Code	Value Label
-991	No information
-992	Asked but unknown
-993	Temporarily unavailable
-994	Not asked
-995	Refused
-998	Not applicable

7. We recommend that when a participant responds with an "I don't know" to a question that the interviewer firstly ensures that the participant understands the question clearly and secondly is gently encouraged to reconsider their response if possible. If "I don't know" is still the response we make use of the 'Asked but unknown' missing code. Questions where "I don't know" is a highly anticipated and valid response will have a checkbox for Unknown included - it should be noted that this will not be recognised as missing data in statistical software.

Recommendations

Instrument A: Demographics

The instrument enables the collection of essential participant demographics such as age, gender, and language.

<table>
<tr><td>Questions</td><td>What is the participant's date of birth?
Age*:
What was is the participant's biological sex at birth?
Response Options: Male; Female; Other
In which country was the participant born?
What is the participant's native language?
(If Other) Specify other native language:
What is the participant's ethnic or tribal affiliation?
(If Other) Specify other ethnic or tribal affiliation:</td></tr>
<tr><td>Notes</td><td>- Date of birth should be collected in following format - DD-MM-YYYY
- Whenever possible, a participant’s date or birth should be captured and verified with official documentation.
- Age is collected in 2 ways: as an approximate estimate, if the date of birth is not known, or calculated automatically if the date of birth is provided.
- The sex field collects the biological sex of a participant, and should not be confused with gender identification. The option “Other” is used to identify Intersex participants who may be sensitive due to stigma about their biological sexual status.
- Country of birth field makes use of ISO numeric codes for countries coding
- The participant’s home/native language that they were raised with should be completed in the language field. If the participant has multiple native languages, complete the field with the most commonly used native tongue – the language they consider their home language
- The participant’s original ethnic or tribal affiliation should be collected in the ethnic affiliation field. If the participant identifies with multiple ethnic tribes, document the primary one or collect the one the participant first identified with growing up.
- Language and ethnic affiliations included in the instrument have been applied from the African Ethnolinguistic Ontology (ELO), developed by H3Africa, and is therefore specifically geared towards African use..
- If ELO is not suitable, local or relevant alternatives should be used instead.</td></tr>
<tr><td rowspan="2">Questions</td><td>What is the participant's biological father's country of birth?
What is the participant's biological father's native language?
(If Other) Specify the participant's biological father's native language:
What is the participant's biological father's ethnic or tribal affiliation?
(If Other) Specify the participant's biological father's ethnic or tribal affiliation:</td></tr>
<tr><td>What is the participant's biological mother's country of birth?
What is the participant's biological mother's native language?
(If Other) Specify the participant's biological mother’s native language:
What is the participant's biological mother’s ethnic or tribal affiliation?
(If Other) Specify the participant's biological mother’s ethnic or tribal affiliation:</td></tr>
</table>

Notes	- If the participant is adopted or orphaned and unable to supply this information, this should be recorded and it may be worth including a field on the demographics form to accommodate this possibility.
Questions	Is the participant currently attending school? **Response Options**: Yes; No (If Yes) What type of school? **Response Options**: Early Learning Centre / Nursery / Creche; Preschool / Kindergarten / Grade R Public Primary school (Grade 1 - 7); Private Primary school (Grade 1 - 7); Public Secondary school (Grade 8 - 12); Private Secondary school (Grade 8 - 12); Special school / School for children with disabilities; Homeschool / Cottage school Tertiary school / College / Technikon /University (If No) What is the highest level of education achieved by the participant? **Response Options**: No formal education; Primary; Secondary; Tertiary What is the highest level of education achieved by the participant's primary caregiver? **Response Options**: No formal education; Primary; Secondary; Tertiary What is the employment status of the participant's primary caregiver? **Response Options**: Self employed; Formal full-time employment by someone else; Part-time employment by someone else; Informal employment (dependent on availability of work) Unemployed
Notes	- Grades associated with different schools may differ regionally. These should be adapted to the local context. - Questions related to the participant's primary caregiver are limited to research participants under the age of 18.

Instrument B: Anthropometrics

The instrument enables the collection of anthropometric data, including height and weight, as well as waist and head circumference.

Questions	Head Circumference Measurement #1: Head Circumference Measurement #2: Head Circumference Measurement #3: Head Circumference Average Measurement:
Notes	- **Head circumference** should only be collected for participants ranging from the Infancy to Childhood Life Stages (see **Life Stage**). - Head circumference, measured in cm using a flexible, non-stretch measuring tape,

	is a measurement taken around the largest area of the head.
Questions	Height measurement #1: Height measurement #2: Height measurement #3: Average height: How tall are you?*

Notes	- **Height** is the distance from the top of the participant's head to the heels of his or her feet (i.e., the vertical length). Three separate height measurements need to be taken in the same session and then averaged to get an accurate height measurement.
	- Height is measured in cm using a stadiometer, if possible.
	- Participant's should only be asked to provide their height if measuring is not possible at all. Self-reported height is considered to be less accurate and should only be used if measured height could not be obtained.
	- Detailed instruments for measuring height can be found in PhenX:
	  - **Standing Height Instrument:** https://www.phenxtoolkit.org/index.php?pageLink=browse.protocoldetails&id=20703
	  - **Measuring height in seated position** for participants unable to stand: https://www.phenxtoolkit.org/index.php?pageLink=browse.protocoldetails&id=20701

Questions	Weight measurement #1: Weight measurement #2: Weight measurement #3 Average weight (kg): Is the participant wearing a cast or medical prosthesis? Location of cast or medical prosthesis: Is the participant wearing street clothes during the weight measurements?

Notes	- **Weight** is measured in kg using a floor scale. The instrument should be calibrated daily using standardized weights, and a log of calibration results should be maintained. Repeat this 3 times and each time record the weight in the 3 separate measurement boxes.
	- **Special Situations**:
	  - Small children: Infants and toddlers who cannot stand alone on the scale will be weighed with an adult, or with an infant's scale. If an adult is holding the child, then the adult guardian or the health technician will stand alone on the scale so the scale can be tared. This sets the scale readout to zero. The child is then handed to the adult and the child's weight is measured.
	  - Note that special consideration may be needed for participants whose weight exceeds the capacity of the study scale. For example, weight can be obtained using two portable scales.
	- Participants should only be asked to provide their weight if measuring is not possible at all. Self-reported weight is considered to be less accurate and should only be used if measured weight could not be obtained.

Questions	Mid-Upper Arm Circumference Measurement #1: Mid-Upper Arm Circumference Measurement #2: Mid-Upper Arm Circumference Measurement #3: Average Mid-Upper Arm Circumference:

Notes	- **Mid-upper arm circumference** (MUAC) is measured on a straight left arm, mid-way between the tip of the shoulder and the tip of the elbow. - MUAC should be measured in cm using a flexible, non-stretch measuring tape.
Questions	Hip Circumference Measurement #1: Hip Circumference Measurement #2: Hip Circumference Measurement #3: Hip Circumference Average Measurement:
Notes	- **Hip circumference** should only be collected for participants ranging from the Childhood to Adult Life Stages (see **Life Stage**). - Hip circumference, measured in cm using a flexible, non-stretch measuring tape, is a measurement taken around the largest part of your hips — the widest part of your buttocks.
Questions	Waist Circumference Measurement #1: Waist Circumference Measurement #2: Waist Circumference Measurement #3: Waist Circumference Average Measurement:
Notes	- **Waist circumference** should only be collected for participants ranging from the Childhood to Adult Life Stages (see **Life Stage**). - Waist circumference, measured in cm using a flexible, non-stretch measuring tape, is a measurement taken around the abdomen at the level of the belly button.

Instrument C: Blood Pressure

The instrument enables the participant's self-reported history of hypertension, as well as current blood pressure estimates.

Questions	Has a healthcare worker ever said the participant has high blood pressure or hypertension? **Response Options**: Yes; No (If Yes) At what age was the participant first told this?
	(FOR Women): Was this during pregnancy only? **Response Options**: Yes; No Has the participant ever taken medication for hypertension/high blood pressure? **Response Options**: Yes, now; Yes, not now; No (If Yes) At what age did they begin taking medicine for hypertension/high blood pressure?
Notes	- **Hypertension** - Persistently high systemic arterial BLOOD PRESSURE. Based on multiple readings (BLOOD PRESSURE DETERMINATION), hypertension is currently defined as when SYSTOLIC PRESSURE is consistently greater than 140 mm Hg or when DIASTOLIC PRESSURE is consistently 90 mm Hg or more. - **Hypertension medication** is also known as antihypertensives; including a range of medication classes such as Diuretics, Beta-blockers, ACE inhibitors, Angiotensin II receptor blockers, Calcium channel blockers, Alpha blockers, Alpha-2 Receptor Agonists, Combined alpha and beta-blockers, Central agonists, Peripheral adrenergic inhibitors, and Vasodilators

Questions	For blood pressure measurements, specify Aneroid sphygmomanometers name and model: Blood pressure cuffs sizes: **Response Options**: S; M; L; XL Systolic Pressure measurement #1 Diastolic Pressure measurement #1 Systolic Pressure measurement #2 Diastolic Pressure measurement #2 Systolic Pressure measurement #3 Diastolic Pressure measurement #3
Notes	- **Systolic Blood Pressure** indicates the blood pressure associated with heart contraction, usually indicated by the numerator on the measurement instrument. - **Diastolic Blood Pressure** indicates the blood pressure associated with heart relaxation, usually indicated by the denominator on the measurement instrument. - Blood pressure is measured in mmHg. - Average systolic and diastolic blood pressure measurements are calculated from the 3 measurements taken.

Instrument D: Urine Test Results

This Protocol enables the recording of a research participant's laboratory results with regards to urine indicators. The information recorded in this Protocol must be gained from qualified medical laboratory facilities with trained and qualified laboratory staff.

Questions	Urinary albumin: Urinary creatinine: Urinary total protein:
Notes	- **Urinary albumin** indicates albumin protein in urine samples, which is a marker for kidney dysfunction. - **Urinary creatinine** indicates creatinine protein in urine samples, which is a marker for kidney dysfunction. - **Urinary total protein** indicates total protein in urine samples, which is a marker for kidney dysfunction.

Instrument E: Smoking Exposure

The instrument is strictly for use in participants younger than 18 and enables the self-reported collection of household smoking exposure.

Questions	Does anyone in the participant's household smoke cigarettes or other tobacco-based products? **Response Options**: Yes; No (If YES) How many members in the household smoke? (If YES) How frequently do they smoke? **Response Options**: Once a day, More than once a day, once or twice a week, once or twice a month

Notes	- The fields collect information regarding any smoking which occurs within the participant's household. - This instrument is strictly applicable for the Infancy to Adolescence Life Stages.

Instrument F: Smoking Status

The instrument enables the self-reported collection of smoking use in participants in the Adolescence and Adult life stages.

Questions	Has the participant smoked at least 1/100 cigarette(s) in your entire life?* **Response Options**: Yes; No How old was the participant when they first started smoking cigarettes?
Notes	- For adolescents, the first question should be specified as 1 cigarette. - For adults, the first question should be specified as 100 cigarettes. - For adults, It is assumed that people who have smoked less than 100 cigarettes in their lifetime do not have a significant smoking status to investigate. - This instrument is strictly applicable for the Adolescence and Adult Life Stages. - Interviewers need to be sensitive to the participant's culture and religion and be aware that some participants may be reluctant to answer these questions truthfully. - Participants should be reassured that their answers will be kept confidential.
Questions	What type of smoker would you currently say you are: **Response Options**: An EVERY day smoker; A FAIRLY REGULAR (some days) smoker; A FORMER smoker; Don't Know; Refused Has the participant EVER smoked cigarettes EVERY DAY for at least 6 months? **Response Options**: Yes; No
Notes	- Non-smokers should be excluded from questions related to frequency of use.
Questions	On the days that they smoke, on average, how many cigarettes does the participant smoke? **OR** If the participant is a former smoker, on the days that they smoked, on average, how many cigarettes did the participant smoke? Over the past 30 days, on how many days did the participant smoke? **OR** If you are a former smoker, on average, on how many days did the participant smoke in a month? (FOR Former Smokers): About how long has it been since the participant COMPLETELY quit smoking cigarettes?
Notes	- The above questions are relevant for An EVERY day smoker; A FAIRLY REGULAR (some days) smoker; and A FORMER smoker.
Questions	In his/her lifetime, has the participant .... [Smoked at least 50 cigars? Smoked a pipe at least 50 times? Used an e-cigarette or vape at least 50 times? used snuff (such as Skoal, Skoal Bandit or Copenhagen) at least 20 times?

	Used chewing tobacco (such as Redman, Levi Garrett or Beechnut) at least 20 times?] **Response Options**: Yes; No
Notes	- Questions related to tobacco use (not in cigarettes) must be completed by all participants regardless of specified cigarette usage.

Instrument G: Alcohol Exposure

The instrument is strictly for use in participants younger than 18 and enables the self-reported collection of household alcohol exposure.

Questions	Does anyone in the participant's household drink alcohol? **Response Options**: Yes; No (If Yes) How many members in the household drink alcohol? (If Yes) How frequently do they drink alcohol? **Response Options**: Once, daily; More than once, daily; Weekly; Infrequently
Notes	- The fields collect information regarding any alcohol use which occurs within the participant's household. - This instrument is strictly applicable for the Infancy to Adolescence Life Stages.

Instrument H: Alcohol Consumption

The instrument enables the self-reported collection of alcohol use in participants in the Adolescence and Adult life stages.

Questions	In his/her entire life, has the participant had at least 1 drink of any kind of alcohol? **Response Options**: Yes; No About how old was the participant when he/she first started drinking alcohol?
Notes	- This instrument is strictly applicable for the Adolescence and Adult Life Stages. - Interviewers need to be sensitive to the participant's culture and religion and be aware that some participants may be reluctant to answer these questions truthfully. - Participants should be reassured that their answers will be kept confidential. - Studies should have a defined quantity or way to measure drinks as examples for participants. One "standard" drink (or one alcoholic drink equivalent) contains roughly 14 grams of pure alcohol. In practice, the total volume differs between types of alcohol. For more information, see: https://www.phenxtoolkit.org/protocols/view/30301
Questions	During the past 30 days, on how many days did the participant drink one or more drinks of an alcoholic beverage? On the days that the participant drank during the past 30 days, how many drinks

	did they usually have each day? What was the LARGEST number of drinks that the participant ever drank in a single day? What is the participant's preferred alcoholic beverage? **Response Options**: Beer; Brandy; Cider; Gin; Rum; Tequila; Vodka; Whiskey; Wine; Other (If Other) Specify preferred alcoholic beverage:
Notes	- Non-drinkers should be excluded from questions related to frequency of use and alcoholic preference.

Instrument I: Drug Use

The instrument enables the self-reported collection of drug use in participants in the Adolescence and Adult life stages.

Questions	Has the participant ever used any of the following substances... [Club Drugs Cocaine / Crack Cocaine? Hallucinogens? Heroin? Inhalents / Solvents? Marijuana? Painkillers? Sedatives / Tranqulizers? Stimulants? Any other non-prescribed substances?] **Response Options**: Yes; No
Notes	- This instrument is strictly applicable for the Adolescence and Adult Life Stages. - Interviewers need to be sensitive to the participant's culture and religion and be aware that some participants may be reluctant to answer these questions truthfully. - Participants should be reassured that their answers will be kept confidential. - Examples: - **Club drugs** - MDMA, ecstasy, GHB, Rohypnol, ketamine, Special K, XTC, roofies - **Cocaine or crack** - blow, rock, snow - **Hallucinogens** - LSD, acid, PCP, mescaline, peyote, psilocybin, mushrooms, angel dust, cactus - **Heroin** - smack, black tar, poppy - **Inhalants or solvents** - nitrous oxide, lighter fluid, gasoline, cleaning fluid, glue, poppers, whippets - **Marijuana** - weed, pot, dope, hashish, Mary Jane, joint, blunt (including THC) - **Painkillers** - methadone, codeine, Demerol, Vicodin, OxyContin, opium, oxy, Percocet, Dilaudid, Percodan, morphine - **Sedatives or tranquilizers** - barbs, downers, Ambien, Lunesta, phenobarbital, pentobarbital, Halcion, Tuinal, Nembutal, Seconal, Librium, Valium, Xanax, benzodiazepines, tranks, Ativan. - **Stimulants** - Adderall, Concerta, Cylert, Provigil, Ritalin or Dexedrine, speed, amphetamine, methamphetamine, uppers, bennies, pep pills, crystal, crank, tik
Questions	For each drug…

	[Age of First Use? Frequency Used in Past 30 Days?]
Notes	- Age of first use and 30 day frequency are limited to those drugs which the participants responded yes too.

Instrument J: Medication Log

The instrument enables the self-reported collection of participant medication use. It is recommended for the collection of prescribed medication, but can also be employed to collect non-prescription medication use.

Questions	Medication name: Medication Coded Name: Indication/reason for medication? Start date: Ongoing? **Response Options**: Yes; No Stop date: Dose amount: Dose units: **Response Options**: mg; ml; spray or puff; tablet; pill; softgel; capsule; application Dose frequency: **Response Options**: once per day (QD); twice a day (BID); three times a day (TID); four times a day (QID); nightly (NOCT); as needed (PRN) Route of administration: **Response Options**: Orally; Per rectum; Intravenous; Per vaginal; Inhaled; Intramuscularly; Nasogastric; Subcutaneously; Sublingually; Topical
Notes	- Participants are particularly encouraged to bring along prescription medication in use to facilitate instrument data collection.

Instrument K: Diabetes History

The instrument enables the retrospective, self-reported collection of a research participant's diabetes history, including the treatment thereof.

Questions	Has a doctor or healthcare worker ever told the participant that he/she has diabetes (sugar in blood)? **Response Options**: Yes; No [(FOR Women, if Yes): Did the doctor diagnose the participant with gestational diabetes (diabetes occurring only during pregnancy)? (If Yes) Did the doctor diagnose the participant with type 1 diabetes? (If Yes) Did the doctor diagnose the participant with type 2 diabetes?} **Response Options**: Yes; No; Don't Know [(If Yes) Is the participant currently taking medication for diabetes?

	(If Yes) Is the participant taking insulin?] **Response Options**: Yes; No At what age was the participant's diabetes first treated? [Was insulin the participant's first diabetes medicine? Has anyone in the participant's family been diagnosed with diabetes (any type)?] **Response Options**: Yes; No
Notes	- **Type 1 Diabetes** - A chronic condition in which the pancreas produces little or no insulin. Type I diabetes mellitus is manifested by the sudden onset of severe hyperglycemia with rapid progression to diabetic ketoacidosis unless treated with insulin. - **Type 2 Diabetes** - A type of diabetes mellitus initially characterized by insulin resistance and hyperinsulinemia and subsequently by glucose intolerance and hyperglycemia. - **Gestational Diabetes** - Carbohydrate intolerance first diagnosed during pregnancy. - **Insulin** - A protein hormone formed from proinsulin in the beta cells of the pancreatic islets of Langerhans. The major fuel-regulating hormone, it is secreted into the blood in response to a rise in concentration of blood glucose or amino acids. Insulin promotes the storage of glucose and the uptake of amino acids, increases protein and lipid synthesis, and inhibits lipolysis and gluconeogenesis.

Instrument L: Asthma

The instrument enables the retrospective, self-reported collection of a research participant's asthma status, including symptoms thereof.

Questions	Has a clinician or a doctor diagnosed the participant as an asthmatic? (If Yes) Is the participant on any chronic medication Yes for management / treatment of asthma? **Response Options**: Yes; No
Notes	- **Asthma** - A chronic respiratory disease manifested as difficulty breathing due to the narrowing of bronchial passageways. - **Asthmatic** - An individual diagnosed with asthma.
Questions	[(If No) Does the participant experience any of the following signs and symptoms: - Frequent coughing spells during play, at night or while laughing or crying? - A chronic cough in the past 3 months or more? - Appear listless / unusually low in energy after physical play? - Rapid breathing sessions from time to time? - Complaining of a tight chest or sore chest? - Complain of difficulty breathing OR (if infant) working harder to breathe (nostrils flaring, skin is sucking in around and between ribs or above the sternum, or exaggerated belly movement)? - Making a whistling sound when breathing in or out? - Difficulty eating OR (Infant) difficult sucking? - Has the participant ever required emergency medical care or hospitalisation for difficulty breathing?] **Response Options**: Yes; No

Notes	- The aforementioned questions are related to asthma symptoms, and should only be collected if a participant has not been diagnosed with asthma.

Instrument M: Dyslipidemia

The instrument enables the retrospective, self-reported collection of a research participant's dyslipidemia history, including the treatment thereof.

Questions	Has a doctor or healthcare worker ever told the participant that he/she has dyslipidemia? **Response Options:** Yes; No (If Yes) At what age was the participant first told this? Was it confirmed with a laboratory test? **Response Options:** Yes; No Has the participant ever taken medication for dyslipidemia? **Response Options:** Yes now; Yes not now; No (If Yes) At what age did the participant start taking medication for dyslipidemia?
Notes	- **Dyslipidemia** - A lipoprotein metabolism disorder characterized by decreased levels of high-density lipoproteins, or elevated levels of plasma cholesterol, low-density lipoproteins and/or triglycerides.

Instrument N: Cancer History

The instrument enables the retrospective, self-reported collection of a brief overview of a research participant's cancer history.

Questions	Has the participant ever had cancer? **Response Options**: Yes; No (If Yes) Specify the number of unrelated cancer diagnoses: (If Yes) Specify the cancer type/s: (If Other) Specify the cancer type:
Notes	- **Cancer** - Cancer is a disease in which some of the body's cells grow uncontrollably and spread to other parts of the body. - Unrelated cancer diagnoses refers to the occurrence of independent cancer events. This does not include cancer which spread from one site to another (metastasized) or recurrence of cancer that was in remission.

Instrument O: Kidney Disease History

The instrument enables the retrospective, self-reported collection of a research participant's kidney diseases history, including the treatment thereof.

Questions	Has a doctor or healthcare worker ever told the participant that he/she had kidney failure? **Response Options**: Yes; No How old was the participant when this was first noted? (If Yes) Are one or both kidneys working well now? **Response Options**: Yes, both kidneys working well; Yes, one kidney working well; No, neither kidneys working well; Don't Know [Has the participant previously been on renal dialysis? Has the participant ever had a kidney transplant?] **Response Options**: Yes; No; Don't Know
Notes	- **Kidney Failure** - failure of the kidneys to adequately filter toxins and waste products from the blood; classified as acute (as in acute kidney injury) or chronic (as in chronic kidney disease); a number of other diseases or health problems may cause either form of renal failure to occur - **Renal dialysis** - A procedure to remove toxic substances from the blood that is used in patients with end-stage chronic kidney disease or acute kidney failure. This includes hemodialysis and peritoneal dialysis. - **Kidney transplant** - A surgical procedure in which one or both kidneys from a donor organism are implanted into a recipient organism.
Questions	[Has anyone in the participant's family either had kidney disease or died from it? (If Yes) Is the type of kidney disease known?] **Response Options:** Yes; No; Don't Know (If Yes) Specify kidney disease: [Has a doctor ever told the participant that his/her kidneys have compromised function? Has a doctor or healthcare worker told the participant that he/she has a kidney disease?] **Response Options:** Yes; No
Notes	- **Kidney Disease** - A disease involving the kidney.

Instrument P: Stroke History

The instrument enables the retrospective, self-reported collection of a research participant's stroke history, including the symptoms thereof.

Questions	[Was the participant ever told by a doctor or healthcare worker he/she had a stroke? Was the participant ever told by a doctor or healthcare worker he/she had a TIA, ministroke, or transient ischemic attack?] **Response Options**: Yes; No If yes, how long did the weakness last? **Response Options**: A few minutes; Less than 15 minutes; Less than an hour; A few hours; More than a day

Notes	- **Stroke** - Sudden impairment of blood flow to a part of the brain due to occlusion or rupture of an artery to the brain. - **TIA** - Brief reversible episodes of focal, nonconvulsive ischemic dysfunction of the brain having a duration of less than 24 hours, and usually less than one hour, caused by transient thrombotic or embolic blood vessel occlusion or stenosis.
Questions	[Has the participant ever had a sudden painless weakness on one side of his/her body? Has the participant ever had a sudden numbness or a dead feeling on one side of his/her body? Has the participant ever had a sudden painless loss of vision in one or both eyes? Has the participant ever suddenly lost one half of his/her vision? Has the participant ever suddenly lost the ability to understand what people are saying? Has the participant ever suddenly lost the ability to express himself/herself verbally or in writing?] **Response Options**: Yes; No
Notes	- The above questions cover a range of symptoms associated with stroke, and this information should be collected regardless of the participant's self-reported incidence of stroke.

Instrument Q: Cardiovascular Disease History

The instrument enables the retrospective, self-reported collection of a research participant's cardiovascular disease history, including the treatment thereof. More specifically, the instrument focuses particularly on the collection of information related to arrhythmia and rheumatic fever history.

Questions	ARRYTHMIA Has a doctor or healthcare worker ever told the participant they have a heart rhythm problem called atrial fibrillation? **Response Options**: Yes; No (If Yes) Provide date of diagnosis: Did the participant go to a hospital/clinic or see a doctor? **Response Options**: Yes; No Has the participant had a permanent pacemaker inserted? **Response Options**: Yes; No (If Yes) What year was the participant's pacemaker inserted? Is the participant taking any of the cardiovascular medications below: - Anticoagulants - Antiarrhythmics **Response Options**: Yes, now; Yes, not now; No

Notes	- **Arrythmia** - Any disturbances of the normal rhythmic beating of the heart or MYOCARDIAL CONTRACTION. Cardiac arrhythmias can be classified by the abnormalities in HEART RATE, disorders of electrical impulse generation, or impulse conduction. - **Pacemaker** - A medical device that generates electrical impulses delivered by electrodes to contract the heart muscles and regulate the electrical conduction system of the heart. - Examples: - **Anticoagulants** - Coumadin; Warfarin; etc. - **Antiarrhythmics** - Quinidine; Procainamide; Norpace; Disopyramide; etc.
Questions	RHEUMATIC FEVER/RHEUMATIC HEART DISEASE Has a doctor or healthcare worker ever told the participant they have rheumatic fever (inflammatory rheumatisim)? **Response Options**: Yes; No (If Yes) Has the participant had it in the past 12 months? **Response Options**: Yes; No Is the participant taking any medication for it? **Response Options**: Yes; No (If Yes) Specify the medication being used:
Notes	- **Rheumatic Fever** - A disease which causes inflammation, especially of the heart, blood vessels and joints. - **Rheumatic Heart Disease** - Cardiac manifestation of systemic rheumatological conditions, such as RHEUMATIC FEVER. Rheumatic heart disease can involve any part of the heart, most often the HEART VALVES and the ENDOCARDIUM.

Instrument R: HIV

The instrument enables the retrospective, self-reported collection of a research participant's HIV testing and treatment history.

Questions	Has the participant ever been tested for HIV? **Response Options**: Yes; No; Refused When did the participant have his/her most recent HIV test? What was the result of the participant's most recent HIV test? **Response Options**: Positive Negative Indeterminate Never obtained results Don't know Refused to answer Is the participant currently on HIV treatment? **Response Options**: Yes; No (If No) Has the participant recently (past 6 months) been on antiretroviral (ARV) treatment for HIV?

	Response Options: Yes; No
Notes	- **HIV** - Includes the spectrum of human immunodeficiency virus infections that range from asymptomatic seropositivity, thru AIDS-related complex, to acquired immunodeficiency syndrome (AIDS). - Interviewers need to be sensitive to the participant's culture and religion and be aware that some participants may be reluctant to answer these questions truthfully. - Participants should be reassured that their answers will be kept confidential.

Instrument S: Infectious Disease History

The instrument enables the retrospective, self-reported collection of a research participant's infectious disease history. The instrument includes a range of diseases, some of which depend on participant Age branching logic.

<table>
<tr><td>Questions</td><td>For each of the specified diseases…
Has a doctor or healthworker ever told the participant that they have 'specified disease'?
Response Options: COVID-19; MIS-C; Other
How old was the participant when diagnosed with 'specified disease'?

Specified Diseases:
- Tuberculosis (TB)
- Malaria
- Sleeping Sickness
- Hepatitis A
- Hepatitis B
- Hepatitis C
- Measles
- Chickenpox
- Meningitis
- COVID-19</td></tr>
<tr><td>Notes</td><td>- Complete descriptions of specified infectious diseases can be found below:
<table>
<tr><td>Intervention</td><td>Description</td></tr>
<tr><td>Tuberculosis</td><td>A chronic, recurrent infection caused by the bacterium Mycobacterium tuberculosis. Tuberculosis (TB) may affect almost any tissue or organ of the body with the lungs being the most common site of infection.</td></tr>
<tr><td>Malaria</td><td>A life-threatening parasitic disease caused by Plasmodium parasites that are transmitted by Anophles mosquito bites to humans and is typically clinically characterized by attacks of fever, headache, chills and vomiting.</td></tr>
<tr><td>Sleeping</td><td>A parasitic disorder caused by protozoa of the Trypanosoma brucei</td></tr>
</table></td></tr>
</table>

Sickness	species. It is transmitted by flies and is endemic in various regions of Sub-Saharan Africa. Signs and symptoms include fever, joint pain, headache, and significant swelling of the lymph nodes. If left untreated, the parasitic infection causes anemia, heart, kidney, and endocrine failure, and neurologic damage. Subsequently patients develop confusion, disruption of the sleep cycle, and mental deterioration.
Hepatitis A	Acute inflammation of the liver caused by the hepatitis A virus. It is highly contagious and usually contracted through close contact with an infected individual or their feces, contaminated food or water.
Hepatitis B	An infection by the hepatitis B virus, which can be transmitted by direct contact of infected blood with mucous membranes or open areas of the skin. Signs and symptoms may include loss of appetite, joint and muscle pain, low-grade fever and stomach pain.
Hepatitis C	A Hepacivirus infectious disease which is a chronic bloodborne infectious disease caused by Hepatitis C virus. The symptoms include fever, fatigue, loss of appetite, nausea, vomiting, abdominal pain, clay-colored bowel movements, joint pain and jaundice.
Measles	A highly contagious viral infection caused by the measles virus. Symptoms appear 8-12 days after exposure and include a rash, cough, fever and muscle pains that can last 4-7 days. Measles vaccines are available to provide prophylaxis, usually combined with mumps and rubella vaccines.
Chickenpox	A contagious childhood disorder caused by the varicella zoster virus. It is transmitted via respiratory secretions and contact with chickenpox blister contents. It presents with a vesicular skin rash, usually associated with fever, headache, and myalgias.
Meningitis	A disorder characterized by acute inflammation of the meninges of the brain and/or spinal cord.
COVID-19	An acute infection of the respiratory tract that is caused by the severe acute respiratory syndrome coronavirus 2. COVID-19 is thought to mainly spread from person to person through respiratory droplets. Typically, there is a two- to 14-day incubation period and infected persons can present with no symptoms or mild to severe fever, dry cough, fatigue, and difficulty breathing. Dysgeusia, anosmia, and gastrointestinal and flu-like symptoms have also been reported.

Abbreviations

ACE: Angiotensin-converting enzyme
ARV: Antiretroviral
CVD: Cardiovascular Disease
COVID-19: Coronavirus Disease 2019
ELO: Ethnolinguistic Ontology
GHB: Gamma hydroxybutyrate
HIV/AIDS: Human Immunodeficiency Virus/Acquired Immune Deficiency Syndrome

LSD: Lysergic acid diethylamide
MDMA: Methylenedioxymethamphetamine
MUAC: Mid-upper arm circumference
PCP: Phencyclidine
TB: Tuberculosis s
THC: Tetrahydrocannabinol
TIA: Transient Ischemic Attack
XTC: Ecstasy

Administration

Mode of Administration

	Instruments																		
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
Interview OR Self-administered questionnaire	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Clinical assessment		X	X																
Bioassay/Lab-based assessment				X															

Life Stage

	Instruments																		
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
Infancy (0 - 12 months)	X	X		X	X		X			X	X	X		X	X		X	X	X
Toddler (13 - 24 months)	X	X		X	X		X			X	X	X		X	X		X	X	X
Childhood (2-11 years)	X	X		X	X		X			X	X	X		X	X		X	X	X
Adolescence (12 - 18 years)	X	X	X	X		X		X	X	X	X	X	X	X	X	X	X	X	X
Adult (18 and older)	X	X	X	X		X		X	X	X	X	X	X	X	X	X	X	X	X

Personnel and Training Required

All instruments (excluding D) may be implemented as either self-reported questionnaires or interviewer- administered questionnaires. If interviewer-administered, interviews should be conducted by trained or study coordinators or data collectors who speak the native/local language of the target population. If available, **All instruments (excluding E to I)** may be recorded from or cross-checked with hospital and(or) patient records. Information recorded in **Protocol C** needs to be collected by qualified healthcare professionals, while information recorded in **Protocol D** needs to be gained from qualified medical laboratory facilities, with trained and qualified laboratory staff.

References

The Core Phenotypes are based on and aligned with several existing standards, to facilitate data harmonisation. These resources are listed below:

1. AWI-Gen Collaborative Centre - Cardiometabolic Disease Research Protocols
2. Protocol - Blood Pressure (https://www.phenxtoolkit.org/protocols/view/40301)
3. Protocol - Cigarette Smoking Status (https://www.phenxtoolkit.org/protocols/view/30604)
4. Protocol - Tobacco - 30-Day Quantity and Frequency (https://www.phenxtoolkit.org/protocols/view/30804)
5. Protocol - Alcohol - Lifetime Use (https://www.phenxtoolkit.org/protocols/view/30101)
6. Protocol - Substances - 30-Day Frequency (https://www.phenxtoolkit.org/protocols/view/31302)
7. Protocol - Medication Inventory (https://www.phenxtoolkit.org/protocols/view/140301)
8. Protocol - Personal History of Type I and Type II Diabetes (https://www.phenxtoolkit.org/protocols/view/140501)
9. Protocol - Personal History of Kidney Failure (https://www.phenxtoolkit.org/protocols/view/140601)
10. Protocol - Arrhythmia (https://www.phenxtoolkit.org/protocols/view/41101)
11. Protocol - Rheumatic Fever/Rheumatic Heart Disease (https://www.phenxtoolkit.org/protocols/view/41401)
12. Protocol - History of Stroke (https://www.phenxtoolkit.org/protocols/view/130301)

Contributors

Phenotype Harmonisation WG - H3Africa
Joint Phenotype Harmonisation Project - H3ABioNet
Phenotype Standardisation Project - H3ABioNet
Alia Benkahla
Fouzia Radouani
Jeff Streuwing
Judit Kumuthini
Katherine Johnston
Lyndon Zass

Melek Chaouch
Vicky Nembaware

Contact Us

For queries related to this standard and guideline, users can log a ticket to the Phenotypes Standards queue in the H3ABioNet Helpdesk. User feedback and improvements on the current toolkit are welcome and encouraged. These can also be submitted through the Helpdesk, or on our GitHub Issues page.