APPENDIX 1

Design of the Prospective Urban Rural Epidemiological Study (PURE Study) Design

The Prospective Urban Rural Epidemiological Study (PURE Study) enrolled 168,067 individuals between 35 and 70 years of age from 21 low, middle and high-income countries. The study includes population samples from 664 communities from 21 countries from 5 continents representing a broad range of economic and social circumstances. PURE includes countries in four income strata based on World Bank classification in 2006: five low-income countries (Bangladesh, India, Pakistan, Tanzania, and Zimbabwe), five lower middle-income countries (China, Colombia, Iran, Occupied Palestinian Territory, and the Philippines), seven upper middle-income countries (Argentina, Brazil, Chile, Malaysia, Poland, South Africa, and Turkey), and four high-income countries (Canada, Saudi Arabia, Sweden, and United Arab Emirates). Recruitment began on Jan 1, 2003, and was completed in the first wave of 18 countries by March 31, 2013. The second wave of three countries (Philippines, Saudi Arabia, and Tanzania) began on January 1, 2012 and was completed by June 2014. The study is coordinated by the Population Health Research Institute, Hamilton Health Sciences and McMaster University, Canada.

Participant Selection Methodology

Selection of Countries

The choice and number of countries selected in PURE reflects a balance between involving a large number of communities in countries at different economic levels, with substantial heterogeneity in social and economic circumstances and policies, and the feasibility of centers to successfully achieve long-term follow-up. Thus, PURE included sites in which investigators

are committed to collecting good-quality data for a low-budget study over the planned 10-year follow-up period and did not aim for a strict proportionate sampling of the entire world.

Selection of Communities

Within each country, urban and rural communities were selected based on broad guidelines. A common definition for "community" that is applicable globally is difficult to establish. In PURE, a community was defined as a group of people who have common characteristics and reside in a defined geographic area. A city or large town was not usually considered to be a single community, rather communities from low-, middle-, and high-income areas were selected from sections of the city and the community area defined according to a geographical measure (eg, a set of contiguous postal code areas or a group of streets or a village). The primary sampling unit for rural areas in many countries was the village. The reason for inclusion of both urban and rural communities is that for many countries, urban and rural environments exhibit distinct characteristics in social and physical environment, and hence, by sampling both, we ensured considerable variation in societal factors across PURE communities. The number of communities selected in each country varied, with the aim to recruit communities with substantial heterogeneity in social and economic circumstances balanced against the capacity of local investigators to maintain follow-up. In some countries (eg, India, China, Canada, and Colombia), communities from several states/provinces were included to capture regional diversity, in policy, socioeconomic status, culture, and physical environment. In other countries (eg, Iran, Poland, Sweden, and Zimbabwe), fewer communities were selected.

Selections of Households and Individuals

Within each community, sampling was designed to achieve a broadly representative sample of that community of adults aged between 35 and 70 years. The choice of sampling frame within

each center was based on both “representativeness” and feasibility of long-term follow-up, following broad study guidelines. Once a community was identified, where possible, common and standardized approaches were applied to the enumeration of households, identification of individuals, recruitment procedures, and data collection. The method of approaching households differed between regions. For example, in rural areas of India and China, a community announcement was made to the village through contact of a community leader, followed by in-person door-to-door visits of all households. In contrast in Canada, initial contact was by mail followed by telephone inviting members of the households to a central clinic. For each approach, at least 3 attempts at contact were made. Households were eligible if at least 1 member of the household was between the ages of 35 and 70 years and the household members intended to continue living in their current home for a further 4 years. All individuals within these households between 35 and 70 years providing written informed consent were enrolled. When a household refused to participate, demographics and simple self-report risk factor data were recorded in a non-responder form.

Guidelines for Selection of Countries, Communities, Households, and Individuals Recruited to PURE

Countries

1. High-income countries, middle-income countries, and low-income countries, with the bulk of the recruitment from low- and middle-income regions.

2. Committed local investigators with experience in recruiting for population studies.

Communities

1. Select both urban and rural communities. Use the national definition of the country to determine urban and rural communities.

2. Select rural communities that are isolated (distance of >50 km or lack easy access to commuter transportation) from urban centers. However, consider ability to process bloods samples, eg, villages in rural developing countries should be within 45-min drive of an appropriate facility.

3. Define community to a geographical area, eg, using postal codes, catchment area of health service/clinics, census tracts, areas bordered by specific streets or natural borders such as a river bank.

4. Consider feasibility for long-term follow-up, eg, for urban communities, choose sites that have a stable population such as residential colonies related to specific work sites in developing countries. In rural areas, choose villages that have a stable population. Villages at greater distance from urban centers are less susceptible to large migration to urban centers.

5. Enlist a community organization to facilitate contact with the community, eg, in urban areas, large employers (government and private), insurance companies, clubs, religious organizations, clinic or hospital service regions. In rural areas, local authorities such as priests or community elders, hospital or clinic, village leader, or local politician.

Individual

1. Broadly representative sampling of adults 35 to 70 years within each community unit.

2. Consider feasibility for long-term follow-up when formulating community sampling framework, eg, small percentage random samples of large communities may be more difficult to follow-up because they are dispersed by distance. In rural areas of developing countries that are not connected by telephone, it may be better to sample entire community (ie, door-to-door systematic sampling).

3. The method of approach of households/individuals may differ between sites. In MIC and HIC, mail, followed up by phone contact may be the practical first means of contact. In LIC,

direct household contact through household visits may be the most appropriate means of first contact.

4. Once recruited, all individuals are invited to a study clinic to complete standardized questionnaires and have a standardized set of measurements.

Guidelines and approach to inclusion of communities and individuals (Selection of countries to ensure that each region of the world involving LIC, MIC and HIC are included)

	Urban	**Rural**
1) Definition	National definition of urban	National definition of rural + OECD definition of <150 people /sq km.
2) Location	Cities	In addition to above definition, rural centres must be reasonably isolated/ separate (distance of >50km or lack easy access to transportation) from urban centers. Note most rural sites are >100km from the urban site
3) Facilitating organization	Large employers (government and private), insurance companies, clubs, religious organizations, neighbourhoods and slums, clinic or hospital service regions	Local authorities such as priests or community elders, hospital or clinic, village leader or local politician.
4) sampling frame	Catchment area of clinics, lists from insurance plans and employers, neighbourhood census data, etc.	All villages that are within a 45-minute drain to a medical centre or appropriate facility (to enable all bloods to be processed within 2-hours).
a) Feasibility For Long-Term Follow-Up	Choose employers or sites that have a stable population. Consider approaches such as cluster-sampling to improve efficiency	Choose villages that have a stable population. Avoid communities that may be susceptible to large migration to urban centres.
b) Choosing the sampling site	From each strata, sites are chosen at random. Other factors pertaining to feasibility must be considered. Specific sites may have to be included to cover full SES spectrum	Ideally, lists of small, medium and large villages are compiled and from each, random locations are selected for sampling. Other local factors pertaining to feasibility must be considered (e.g. distance to blood processing facility).
5) Approaches to Recruitment	Door to door contact or through employer	Door to door recruitment of all households in selected village.

Source: PURE Protocol

Collection of Demographics, Risk Factors and Outcome Events

We collected data at national, community, household, and individual levels with standardized questionnaires. Questions about age, sex, education, smoking status, hypertension, diabetes, and obesity were identical to those in the INTERHEART and INTERSTROKE studies. We obtained BP measurements in individuals and so hypertension was defined as those with a BP >140/>90 or those who were already on treatment. Fasting glucose was available in most individuals (76%) and so diabetes was defined as those who were reported as having diabetes and those with a fasting glucose >7.0 mmol/L. (Sensitivity analyses indicate a very high correlation between self report of diabetes alone versus self-report and fasting glucose >7.0 in the 110,000 people with both measures, and so self-report is a reasonable surrogate for the prevalence of diabetes) Total cholesterol was available in 122,640 individuals and a value of >5.2 mmol/L was considered to be elevated.

In most of the LIC and MIC there was no central system of death or event registration. We therefore; 1) obtained information on prior medical illness and medically certified cause of death where available, 2) captured best available information from reliable sources in those instances where medical information was not available in order to be able to arrive at a probable diagnosis or cause of death. Event documentation was based on information from household interviews and medical records, death certificates and other sources. We also used Verbal Autopsies to ascertain cause of death in addition to medical records which were reviewed by a health professional. This approach has been used in several studies conducted in LIC and MIC.

To ensure a standard approach and accuracy for classification of events across all countries and over time, the first 100 CVD events (deaths, MI, strokes, heart failure or cancers) for China and India, and 50 cases for other countries were adjudicated both locally and also by the

adjudication chair, and if necessary further training was provided. Thereafter, every year, 50 cases for China and India and 25 cases for each of the remaining countries were adjudicated as above.

APPENDIX 2

Major Cardiovascular Disease: This included myocardial Infarction (ECG with new ischemic changes new ST elevation/depression or T wave inversion ≥2 mm), stroke (an acute focal neurological deficit diagnosed by a physician and thought to be of vascular origin (without other causes such as brain tumor) with signs and symptoms lasting ≥24 hrs), congestive heart failure (the diagnosis of congestive heart failure requires signs - rales, increased jugular venous pressure or ankle edema or symptoms - nocturnal paroxysmal dyspnea, dyspnea at rest or ankle edema of congestive heart failure and one or both of the following: radiological signs of pulmonary congestion and treatment of heart failure with diuretics), effort angina with documented Ischemia (stress test with ECG with new ST depression >1 mm or positive imaging – ECHO/scan compatible with ischemia) and unstable angina (hospitalization for typical symptoms with new ECG changes – T wave inversion <2mm or Coronary revascularization within one week of admission, and treated as unstable angina).

i. Cardiovascular mortality: include sudden unexpected cardiovascular death, fatal myocardial infarction, fatal stroke, fatal congestive heart failure, death due to other cardiovascular diseases.

ii. All cause mortality: included CV death and non CV death – infections, cancers, respiratory disease, pregnancy/delivery/puerperium, injury or other causes (eg. nervous system, digestive system, genitourinary system).

Event Definitions

FATAL EVENTS

Death due to cardiovascular events

Sudden unexpected cardiovascular death

Death that occurred suddenly and unexpectedly without evidence of other cause of death (examples: witnessed collapse, persons resuscitated from cardiac arrest who later died) or persons seen alive less than 12 hours prior to discovery of death (example persons found dead in his/her bed).

Non-sudden unexpected cardiovascular death.

Death that occurred unexpectedly without evidence of other cause of death in persons seen alive more than 12 hours but less than 24 hours.

Fatal myocardial infarction (one of the following)

• Autopsy demonstrating fresh myocardial infarction and/or recent coronary occlusion, or

• ECG showing new and definite sign of MI (Minnesota code 1.11), or

• Symptoms typical or atypical or inadequately described but attributed to cardiac origin lasting more than 10 minutes and cardiac enzymes at least twice above the upper limit of normal or troponin at least at the lower level of necrosis, or

• ECG with new ischemic changes (new ST depression or T wave inversion ≥ 2 mm) and cardiac enzymes at least twice above the upper limit of normal or troponin at least at the lower level of necrosis.

Fatal stroke

Diagnosis of stroke by a physician based on sudden neurological deficit of vascular origin with or without neuroimaging studies (CT scan/MR scan/angiography/Doppler) lasting 24 hours and more, occurring within 30 days of signs or symptoms of stroke or autopsy evidence of a recent stroke. *If death occurred within 24 hours of onset of stroke signs, this will be considered a definite death due to stroke.*

Congestive heart failure

Death due to heart failure in absence of myocardial infarction or other causes was attributed to fatal heart failure.

The diagnosis of congestive heart failure required 2 of the 3 following criteria:

• Signs (rales, increased jugular venous pressure or ankle edema) or symptoms (nocturnal paroxysmal dyspnea, dyspnea at rest of ankle edema) of congestive heart failure,

• Radiological signs of pulmonary congestion,

• Treatment of heart failure with diuretics

Other cardiovascular death *(other causes having been excluded)*

Arterial rupture of aneurysm, Pulmonary embolism, Arrhythmic death (A-V block, sustained ventricular tachycardia in absence of other causes), Death after invasive cardiovascular intervention, Congenital heart disease, Heart valve disease (including rheumatic heart disease, Endocarditis, Myocarditis, Tamponade.

Non-fatal cardiovascular events

Non-periprocedural myocardial infarction

• ECG showing new and definite sigh of MI (Minnesota code 1.11), or

• Symptoms typical or atypical or inadequately described but attributed to cardiac origin lasting more than 10 minutes and cardiac enzymes at least twice above the upper limit of normal (ULN) or troponin at least at the lower level of necrosis, or

• ECG with new ischemic changes (new ST depression or T wave inversion ≥ 2 mm) and cardiac enzymes at least twice above the upper limit of normal or troponin at least at the lower level of necrosis.

Periprocedural myocardial infarction

ECG showing new and definite sign of MI (Minnesota code 1.11), OR cardiac marker values:

• Percutaneous coronary intervention, CKMB should be ≥ 3X ULN of troponin ≥ 5 X above lower level of necrosis,

• Coronary surgery cardiac markers CKMB should be ≥ 10X ULN or troponin ≥ 10 X above lower limit of necrosis.

Stroke

Diagnosis of stroke by a physician based on sudden neurological deficit of vascular origin with or without neuroimaging studies (CT scan/MRI scan/angiography/Doppler) lasting 24 hours and more.

Congestive heart failure

The diagnosis of congestive heart failure requires 2 of the 3 following criteria:

• Signs (rales, increased jugular venous pressure or ankle edema) of symptoms (nocturnal paroxysmal dyspnea, dyspnea at rest of ankle edema) of congestive heart failure,

• Radiological signs of pulmonary congestion,

• Treatment of heart failure with diuretics

Non-Major CVD

Includes all hospitalizations for CVD (other than those due to myocardial infarction, stroke, or heart failure) or related investigations or procedures.