

Appendices

Appendix 1: PURE Study Participant Selection Methodology as Excerpted from Teo et al. Am Heart J. 2009 Jul;158(1):1-7

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 (see Guidelines for Selection of Countries, Communities, Households, and Individuals Recruited to PURE). 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 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 (e.g., 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 (e.g., 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 (e.g. 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. 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.

For each approach, at least 3 attempts at contact were made. All individuals within these households between 35 and 70 years providing written informed consent were enrolled. When an eligible household or eligible individual in a household refused to participate, demographics and self-reported data about CVD risk factors, education, and history of CVD, cancers and deaths in the households within the 2 previous years were recorded.

To ensure standardization and high data quality, we used comprehensive operations manual, training workshops, DVDs, regular communication with study personnel and standardized report forms. We entered all data in a customized database programmed with range and consistency checks, which was transmitted, electronically to the Population Health Research Institute in Hamilton (Ontario, Canada) where further quality checks were implemented.

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, e.g., villages in rural developing countries should be within 45-min drive of an appropriate facility.
3. Define community to a geographical area, e.g., 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, e.g., 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, e.g., 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 (i.e., 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.

Appendix 2: Standardized Event Definitions in PURE

Prospective Follow-up for Cardiovascular Events and Mortality: History of disease was collected at baseline from every participant with standardized questionnaires regarding history of a) hypertension, b) diabetes c) stroke d) angina/myocardial infarction/coronary artery disease e) heart failure f) other heart disease.

Information on specific events (death, myocardial infarction, stroke, heart failure, cancer, hospitalizations, new diabetes, injury, tuberculosis, human immunodeficiency viral infections, malaria, pneumonia, asthma, chronic obstructive pulmonary disease) were obtained from participants or their family members (events were reported by the participants if alive or by a relative if the individual had died). This information was adjudicated centrally in each country by trained physicians using standardized definitions. Because the PURE study involves urban and rural areas from middle- and low-income countries, supporting documents to confirm cause of death and/or event varied in degrees of completion and availability. In most of middle- and low-income countries there was no central system of death or event registration. Therefore, information was obtained about prior medical illness and medically certified cause of death where available, and, second, best available information was captured 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. Verbal autopsies were also used 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 middle- and low-income countries.

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 3: Definitions and assessments of factors included in baseline participant characteristics

Physical inactivity was assessed using the long-form International Physical Activity Questionnaire (IPAQ) (1) and calculated as the total of occupation, transportation, housework, and recreational activity reported in metabolic equivalents (MET) \times minutes per week. Physical activity was also reported in minutes per week of moderate intensity physical activity using the equation where minutes reported in each physical activity domain on the IPAQ by the participant are weighted relative to moderate intensity physical activity. Physical inactivity was defined as physical activity level < 600 MET \times minutes per week, which corresponds to < 150 minutes of moderate intensity physical activity per week (2). Diet quality was assessed by the Alternate Healthy Eating Index (AHEI) (3). Poor diet is defined as having the lowest tertile of the AHEI score. Disability was assessed using the PURE baseline questionnaire which recorded information on trouble grasping/handling with fingers, walking (requiring a walking stick cane/walker), bending down and picking up objects, reading, seeing a person from across the room (with glasses worn), speaking, or hearing in a normal conversation. Individuals having none of those items were scored as 0, those having one item were scored as 1, and those having two and more were scored as 2, indicating presence of multiple disabilities. Depression was assessed using the PURE baseline questionnaire which documented whether a participant felt sad, blue, or depressed for two weeks or longer in the previous year, and if so, whether they experienced loss of interest in pleasurable activities, tiredness, unintentional weight changes, difficulty sleeping or concentrating, feeling of worthless or thoughts about death during the same period. We classified participants with 4 or more of those 7 symptoms as having probable depression, consistent with previous validation studies (4, 5).

References of appendix 3

1. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc.* 2003 Aug;35(8):1381–95.
2. World Health Organization. Global recommendations on physical activity for health. [Internet]. 2010 [cited 2019 Oct 4]. Available from: <http://www.ncbi.nlm.nih.gov/books/NBK305057/>
3. Shivappa N, Hebert JR, Kivimaki M, Akbaraly T. Alternative Healthy Eating Index 2010, Dietary Inflammatory Index and risk of mortality: results from the Whitehall II cohort study and meta-analysis of previous Dietary Inflammatory Index and mortality studies. *Br J Nutr.* 2017;118(3):210–21.
4. Kessler RC, Andrews G, Mroczek D, Ustun B, Wittchen H-U. The World Health Organization Composite International Diagnostic Interview short-form (CIDI-SF). *Int J Methods Psychiatr Res.* 1998;7(4):171–85.
5. Patten SB, Brandon-Christie J, Devji J, Sedmak B. Performance of the composite international diagnostic interview short form for major depression in a community sample. *Chronic Dis Can.* 2000;21(2):68–72.

Appendix 4: Detail in mediation analyses

To assess the extent to which these factors explain the association between social isolation and mortality, we assessed mediating effects by calculating the percentage of excess risk mediated (PERM) for the factors. The PERM was calculated using a following formula: $PERM = \frac{Hazard\ ratio_1 - Hazard\ ratio_2}{Hazard\ ratio_1} \times 100$ (1). Hazard ratio1 is hazard ratio adjusted for age, sex, education attainment (pre-secondary, secondary or post-secondary education), residence area (rural or urban), country income (LIC, MIC- or HIC). Hazard ratio2 is further adjusted for potential mediators.

Reference of appendix 4

1. Lin, D.Y., Fleming, T.R., De Gruttola, V., 1997. Estimating the proportion of treatment effect explained by a surrogate marker. *Stat. Med.* 1997; 16: 1515–1527.

Appendix 5: Variations in the association between social isolation and variables by residence areas

Variables	Odds ratio (95% CI)		
	Residence area		P for interaction
	Rural (N=53,958)	Urban (N=64,806)	
Age, 10-year increase	1.10 (1.06-1.15)	1.12 (1.08-1.15)	<0.0001
Women (vs. men)	1.58 (1.44-1.73)	2.66 (2.49-2.85)	<0.0001
Education attainment level Pre-secondary (vs. secondary or post-secondary)	1.45 (1.33-1.59)	1.16 (1.08-1.24)	0.0003
Unemployed vs. employed	1.35 (1.23-1.47)	1.06 (0.99-1.13)	0.0008
Current smoking (vs. former or never smoking)	1.03 (0.93-1.14)	1.45 (1.34-1.56)	0.01
Low diet score (lowest tertile of AHEI) (vs. the other two tertiles)	1.16 (1.07-1.27)	1.27 (1.20-1.36)	0.2
Current alcohol use (vs. former or never drinking)	1.40 (1.27-1.54)	1.09 (1.02-1.16)	0.0003
Physical inactivity (vs. WHO recommended physical activity)	0.99 (0.89-1.11)	1.03 (0.95-1.11)	0.6
Number of comorbidities ≥ 2 (vs. one or no comorbidities)	1.10 (0.94-1.29)	1.14 (1.03-1.26)	0.4
Number of disabilities ≥ 2 (vs. one or no disabilities)	1.41 (1.28-1.55)	1.19 (1.11-1.29)	0.004

Odds ratios were adjusted for age, sex, education attainment, employment status, residence area, smoking, alcohol, presence of physical inactivity, diet score, presence of comorbidities, and presence of disabilities. AHEI: alternative healthy eating index.

Appendix 6: Risk of mortality stratified by variables and social isolation

Subgroups	All-cause mortality	p for interaction
Age category		<0.0001
< 60 years	1.42 (1.27-1.60)	
≥ 60 years	1.24 (1.11-1.39)	
Sex		0.0003
Women	1.24 (1.11-1.40)	
Men	1.49 (1.33-1.67)	
Residence area		0.3
Urban	1.37 (1.23-1.54)	
Rural	1.29 (1.15-1.45)	