Impact of social isolation on mortality and morbidity in 20 high-income, middle-income and low-income countries in five continents

Ryo Naito,1,2 Darryl P Leong,1,2 Shrikant Ishver Bangdiwala,1,3 Martin McKee,1,4 S V Subramanian,5 Sumathy Rangarajan,1,2 Shofiqul Islam,1,2 Alvaro Avezum,6 Karen E Yeates,1 Scott A Lear,1 Rajeev Gupta,7 Afzalhussein Yusufali,10 Antonio L Dans,11 Andrzej Szuba,12 Khalid F Alhabib,13 Manmeet Kaur,14 Omar Rahman,15 Pamela Seron,16 Rafael Diaz,17 Thandi Puoane,18 Weida Liu,19 Yibing Zhu,20 Yundong Sheng,21 Patricio Lopez-Jaramillo,22 Jephath Chifamba,23 Ismail Rosnah,24 Kubilay Karsidag,25 Roya Kelishadi,26 Annika Rosengren,27 Rasha Khatib,28,29 Leela Itty Amma K R,30,31 Syed Iqbal Azam,32 Koon Teo,1,2 Salim Yusuf1,2

ABSTRACT

Objective To examine the association between social isolation and mortality and incident diseases in middle-aged adults in urban and rural communities from high-income, middle-income and low-income countries.

Design Population-based prospective observational study.

Setting Urban and rural communities in 20 high income, middle income and low income.

Participants 119894 community-dwelling middle-aged adults.

Main outcome measures Associations of social isolation with mortality, cardiovascular death, non-cardiovascular death and incident diseases.

Results Social isolation was more common in middle-income and high-income countries compared with low-income countries, in urban areas than rural areas, in older individuals and among women, those with less education and the unemployed. It was more frequent among smokers and those with a poorer diet. Social isolation was associated with greater risk of mortality (HR of 1.26, 95% CI: 1.17 to 1.36), incident stroke (HR: 1.23, 95% CI: 1.07 to 1.40), cardiovascular disease (HR: 1.15, 95% CI: 1.05 to 1.25) and pneumonia (HR: 1.22, 95% CI: 1.09 to 1.37), but not cancer. The associations between social isolation and mortality were observed in populations in high-income, middle-income and low-income countries (HR: 95% CI: 1.69 (1.32 to 2.17), 1.27 (1.15 to 1.40) and 1.47 (1.25 to 1.73), respectively, interaction p=0.02). The HR associated with social isolation was greater in men than women and in younger than older individuals. Mediation analyses for the association between social isolation and mortality showed that unhealthy behaviours and comorbidities may account for about one-fifth of the association.

Conclusion Social isolation is associated with increased risk of mortality in countries at different economic levels. The increasing share of older people in populations in many countries argues for targeted strategies to mitigate its adverse effects.

Key questions

What is already known?

► With ageing populations, urbanisation and fewer extended families, social isolation is becoming more common.

► Social isolation is associated with negative health consequences.

► Prior studies on social isolation are mainly from high-income countries, primarily from urban populations.

What are the new findings?

► For the first time, we investigate associations between social isolation and health outcomes in middle-aged community-dwelling adults from urban and rural sites in high-income, middle-income and low-income countries.

► The mortality risk of social isolation is consistently observed among diverse populations regardless of residence area (rural or urban) and country income level.

► The mortality risk was partly explained by unhealthy behaviours and baseline comorbidities.

► Social isolation is associated with increased risk of incident stroke, cardiovascular disease and pneumonia.

What do the new findings imply?

► Our study shows that the risk of mortality associated with social isolation is observed consistently among diverse populations regardless of residence area (rural or urban) and country income level. Healthcare workers and policy-makers should consider social isolation as an added risk factor for premature death.

INTRODUCTION

Social isolation is characterised as the absence of social relationships1 in the forms of social
contacts, social resources and participation in social or religious activities.²³ Ageing populations, urbanisation and fewer extended families are increasing levels of social isolation in many countries. In meta-analyses, social isolation was associated with a 29% increase in the risk of death⁴ and a 29% increase in the risk of coronary heart disease and 32% increase in the risk of stroke.⁵ However, most these studies are from high-income countries (HICs), primarily from urban populations, and with a focus on older people⁶⁻⁹ with few studying the general adult population. None examined whether there are differences in observed associations in countries at different levels of economic development. This last point is important because while social networks in poor countries may be stronger, social services provided by governments or other organisations may be weaker.¹⁰ Furthermore, family and social structures in rural communities may differ from that in urban communities. Here, we examine the relationship between social isolation and health outcomes in middle-aged community-dwelling adults from urban and rural sites in several HIC, middle-income country (MIC) and low-income country (LIC).

METHODS

The objective of this study was to examine the association between social isolation and both mortality and certain incident diseases in middle-aged adults from HIC, MIC and LIC. We hypothesised that social isolation is associated with increased risk of mortality in populations everywhere but the magnitude of associations between social isolation and outcomes may vary by the economic development of a country, by urban or rural residence, between men and women and by age group.

This is a secondary analysis of the Prospective Urban Rural Epidemiology (PURE) study, which is a prospective, population-based cohort study that has recruited community-dwelling adults aged 35–70 years old from both urban and rural areas.¹¹ Details of the sampling methods, response rates, documentation of events and their adjudication have been published previously and are summarised in the online supplemental appendix 1. Countries selected were classified according to the World Bank scheme as HIC, MIC and LIC at the beginning of the study in 2006. The HIC include Canada, Saudi Arabia, Sweden and the United Arab Emirates. The MIC include Argentina, Brazil, Chile, China, Colombia, Iran, Malaysia, Palestine, Philippines, Poland, South Africa and Turkey. The LIC include Bangladesh, India, Pakistan, Tanzania and Zimbabwe. Although some of the countries have subsequently been reclassified as their economic status has changed, for consistency with previous publications from PURE, we have retained the countries within their original economic categories assigned at the beginning of the study. Information on lifestyles, health-related risk factors, the presence of chronic disease and outcomes were collected using standardised methods (online supplemental appendix 2). Sampling aimed to achieve a broadly representative sample of adults living in each community. Although different sites used varying methods of approaching households depending on what was practical in each setting, all enrol individuals irrespective of the presence of pre-existing conditions. At least three attempts to contact individuals were made in all approaches. If a household was eligible (at least one member was between the ages of 35 and 70 years), then all consenting and eligible individuals were enrolled. Once recruited, all participants were invited to a clinic to complete a standardised set of questionnaires and measurements. Participants or other family members were contacted in person or by telephone at least every 3 years to document deaths and key non-fatal events. Information on medically certified death was accessed through administrative registries, where available. Otherwise, event documentation was obtained from household interviews, medical records, death certificates, verbal autopsies¹² and other sources. Participants from China (n=47927) are not included in this study since local ethics committees did not allow some of the key questions on social isolation.

Patient and public involvement

The study participants and the public were not involved in the design and conduct of this study.

Assessment of social isolation

A single measure of social isolation has not yet been agreed. Berkman and Syme constructed their social network index (The Social Network Index; SNI) to assess social relations in 1979.² Subsequent studies have used the SNI or variations of it as measures of social isolation. The SNI comprises four domains; information on partnership, contact with family members or friends, engagement in religious activities and membership in organisations or clubs. In this analysis, we measured social isolation using an adaptation of the SNI. The social isolation scale was constructed using five items from the PURE baseline questionnaire relevant to the SNI:

1. Marital status (scored as 1 for any of the following: never married, widowed, separated, or divorced and 0 otherwise).
2. ‘Can you count on your family members in a difficult situation?’ (possible responses include none, little, moderate/average and a great deal)—scored as 1 for ‘none’ or ‘little’ and 0 for ‘moderate’ or ‘a great deal’.³
3. ‘Can you count on any organization in a difficult situation?’ (possible responses include none, little, moderate/average and a great deal—scored as 1 for ‘none’ or ‘little’ and 0 for ‘moderate’ or ‘a great deal’).
4. ‘Are you a member of any religious group?’ (yes=0, no=1).
5. ‘Are you a member of any social group?’ (yes=0, no=1).

The social isolation scale ranges from 0 to 5. Individuals who score 0 are defined as having the most social support and those who score 5 are defined as having maximum
social isolation. In preparatory work, we explored using the scale as a continuous or binary predictor and confirmed the relationship with mortality was non-linear and it was more appropriate to treat social isolation as a binary variable. We therefore considered individuals with a score of 4 or 5 as being socially isolated.

Outcome
The outcomes of interest for this analyses were all-cause mortality, cardiovascular mortality, non-cardiovascular mortality and incidence of selected diseases (myocardial infarction, stroke, heart failure, cardiovascular disease (CVD), cancer, pneumonia, chronic obstructive pulmonary disease (COPD) and injury). CVD included fatal or non-fatal myocardial infarction, stroke, heart failure and other fatal CVD events.

Statistical analysis
The characteristics of participants in each of the two groups (social isolation vs no social isolation) were compared using χ² tests for categorical variables and student t-test or the Mann-Whitney U test for continuous variables. Multivariable logistic regression analyses with social isolation as a dependent variable were conducted to evaluate factors associated with social isolation. We used Cox proportional-hazard regression models to evaluate the relationship between social isolation and mortality. To account for the clustered nature of the data, we used shared frailty models in which the community to which each individual belongs served as the clustering variable. In the models, those with a social isolation score of 0–3, served as the reference group. The adjusted model included following baseline variables: age, sex, education attainment (presecondary, secondary or postsecondary education), residence area (rural or urban), country income (LIC, MIC or HIC), smoking, alcohol use, physical inactivity, diet score, hypertension, diabetes, coronary artery disease, depression and disabilities. We also performed Cox regression analyses using the adjusted model to evaluate the relationship between social isolation and incident disease (myocardial infarction, stroke, heart failure, CVD, cancer, pneumonia, COPD and injury). The incidence rates of each outcome were expressed in person-years (per thousand). The CIs are calculated using the quadratic approximation to the Poisson log likelihood for the log-rate parameter. Definitions and values of baseline participant characteristics are reported in online supplemental appendix 3. To quantify the contribution of risk factors to mortality, population attributable fractions were calculated from a Cox proportional regression model, in which social isolation, education attainment, smoking, smoking, alcohol, physical inactivity, diet quality, hypertension, diabetes and depression were included. Mediation analyses were performed to examine what factors mediate any relationship between social isolation and mortality. Factors chosen as the candidate potential mediators were behavioural factors (current smoking, current alcohol, physical inactivity and low diet quality) and comorbidities (hypertension, diabetes, abdominal obesity, coronary artery disease, stroke, cancer, disabilities and depression). The analytical methods are provided in online supplemental appendix 4. To minimise the potential for reverse causality, we performed sensitivity analyses in which participants with diseases at baseline or those who developed clinical outcomes within the first 2 years of follow-up were excluded. STATA V.15.1 was used for statistical analyses and graphs.

RESULTS
Characteristics of participants with and without social isolation
A flow chart describing the selection of the study population is provided in online supplemental figure 1. A total of 119894 individuals were enrolled between 6 July 2005 and 2 June 2016, of whom 118764 with the social isolation scale recorded were included in this study. The proportions of participants from LIC, MIC and HIC were 31.9% (n=37863), 52.9% (n=62855) and 15.2% (n=18046), respectively. The prevalence of social isolation (social isolation scale of 4 or 5) was 10.9% (n=12992). Socially isolated participants were older and more likely to be women (table 1). They had higher prevalence of baseline comorbid conditions including hypertension, diabetes mellitus, coronary artery disease, stroke, cancer, COPD, as well as depression. Table 2 shows participants’ characteristics associated with social isolation. Older age, being female, with a low level of education and unemployed were associated with increased odds of being socially isolated. Social isolation was more common in urban than rural areas; and in MICs and HICs compared with LICs. Current smoking, poor diet and disabilities were associated with social isolation.

Social isolation by country income
The age-sex adjusted prevalence of social isolation in LIC, MIC and HIC were 7.7%, 13.1% and 12.0%, respectively (figure 1A). Table 3 shows the participant characteristics of participants associated with social isolation. Women were more likely to be socially isolated, consistently across countries at all income levels. The directions of the association between other factors and social isolation were inconsistent across different country income levels. For example, low education was associated with social isolation in LICs and MICs, while no association was observed in HICs. Unemployment was strongly associated with social isolation in LICs, while similar associations were not observed in MICs or HICs. Higher age was associated with social isolation in LICs and MICs but not in HICs.

Social isolation in urban and rural populations
The age-sex adjusted prevalence of social isolation in rural areas was 9.3% compared with 12.4% in urban areas (figure 1B). The participant characteristics stratified urban and rural residence associated with social
isolation are shown in online supplemental appendix 5. The patterns of all variables except for employment status and current smoking were similar irrespective of rural or urban area of residence. Older age, women, low education, poor diet, current alcohol use and disabilities were associated with increased odds of being socially isolated. Unemployment was associated with social isolation in rural but not urban residence. Current smoking was associated with social isolation in urban but not in rural residence.

Association between social isolation and mortality
Survival analyses were conducted in 115,816 (97.5%) individuals whose vital status was available. During the mean follow-up of 9.0 years, we observed 9,487 (8.2%) deaths (2,695 cardiovascular and 6,794 non-cardiovascular). The adjusted HR of mortality for social isolation was 1.26 (95% CI: 1.17 to 1.36) (figure 2A). The adjusted HR for cardiovascular and non-cardiovascular mortality were 1.30 (95% CI: 1.13 to 1.50), and 1.25 (95% CI: 1.14 to 1.37), respectively. The magnitude of mortality risks associated with social isolation was greatest in HICs (figure 2B). While social isolation was consistently associated with increased risk of all-cause mortality regardless of age, sex, area of residence and country income level, the magnitude of the association was greater in younger adults and men.

Population attributable fractions were examined to quantify the contribution of social isolation to all-cause mortality. These are compared with similar data for other risk factors. The population attributable fraction of social isolation was 2.4%, which is modest compared with the other risk factors (figure 3). Regional variations in mortality risks associated with social isolation are shown

### Table 1 Baseline characteristics of study participants with and without social isolation

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>No social isolation (n=105,772)</th>
<th>Social isolation (n=12,992)</th>
<th>P value Univariate analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, year (mean±SD)</td>
<td>50.1±9.8</td>
<td>52.2±10.2</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Women (%)</td>
<td>59,567 (56.3)</td>
<td>9402 (72.4)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Hypertension (%)</td>
<td>22,682 (21.5)</td>
<td>3372 (26.0)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Diabetes mellitus (%)</td>
<td>10,066 (9.5)</td>
<td>1357 (10.5)</td>
<td>0.001</td>
</tr>
<tr>
<td>Coronary artery disease (%)</td>
<td>3115 (3.0)</td>
<td>443 (3.4)</td>
<td>0.003</td>
</tr>
<tr>
<td>Stroke (%)</td>
<td>1396 (1.3)</td>
<td>271 (2.1)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Cancer (%)</td>
<td>1646 (1.6)</td>
<td>255 (2.0)</td>
<td>0.001</td>
</tr>
<tr>
<td>COPD (%)</td>
<td>845 (0.9)</td>
<td>159 (1.2)</td>
<td>0.001</td>
</tr>
<tr>
<td>Depression (%)</td>
<td>15,570 (14.8)</td>
<td>2670 (20.6)</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

COPD, chronic obstructive pulmonary disease.

### Table 2 Factors associated with social isolation using multivariable logistic regression analyses

<table>
<thead>
<tr>
<th></th>
<th>Adjusted OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, 10-year increase</td>
<td>1.06 (1.03 to 1.09)</td>
</tr>
<tr>
<td>Women (vs men)</td>
<td>2.17 (2.06 to 2.29)</td>
</tr>
<tr>
<td>Education attainment level</td>
<td>1.37 (1.29 to 1.45)</td>
</tr>
<tr>
<td>Presecondary (vs secondary or postsecondary)</td>
<td>1.64 (1.55 to 1.73)</td>
</tr>
<tr>
<td>Unemployed vs employed</td>
<td>1.16 (1.10 to 1.22)</td>
</tr>
<tr>
<td>Residence in urban (vs rural)</td>
<td>2.41 (2.25 to 2.57)</td>
</tr>
<tr>
<td>Country income level (low as reference)</td>
<td>2.03 (1.85 to 2.22)</td>
</tr>
<tr>
<td>Middle vs low</td>
<td>1.33 (1.25 to 1.41)</td>
</tr>
<tr>
<td>High vs low</td>
<td>1.04 (0.98 to 1.10)</td>
</tr>
<tr>
<td>Current smoking (vs former or never smoking)</td>
<td>1.04 (0.97 to 1.11)</td>
</tr>
<tr>
<td>Low diet score (lowest tertile of diet score) (vs the other two tertiles)</td>
<td>1.27 (1.07 to 1.18)</td>
</tr>
<tr>
<td>Current alcohol use (vs former or never drinking)</td>
<td>1.06 (0.97 to 1.15)</td>
</tr>
<tr>
<td>Physical inactivity (vs WHO recommended physical activity)</td>
<td>1.27 (1.20 to 1.35)</td>
</tr>
</tbody>
</table>

ORs were adjusted for age, sex, education attainment, employment status, residence area, country income level, smoking, alcohol, presence of physical inactivity, diet score, presence of comorbidities and presence of disabilities.
Overall, mortality rates were higher among the socially isolated and although with some variations across regions. The magnitude of the association between social isolation and mortality was greatest in South Asia, North America/Europe followed by Africa and South America. The association was not significant in Middle East and Southeast Asia.

Mediation analyses for the association between social isolation and mortality showed that unhealthy behaviours accounted for 18% of the association, while comorbidities explained 3% of the association. A model adjusted for both behavioural factors and comorbidities showed that these variables accounted for 21% of the association.

Association between social isolation and incident disease

During follow-up, a new myocardial infarction occurred in 3417 (3.0%), a new stroke in 2129 (1.8%), new onset heart failure in 827 (0.7%), a new cancer in 4377 (3.8%), pneumonia in 2578 (2.2%), a new diagnosis of COPD in 1423 (1.2%) and hospitalisation for injury in 13608 (11.7%). Figure 4 shows that after adjustment social isolation was associated with an increased risk of stroke (HR: 1.23, 95% CI: 1.07 to 1.40) and CVD (HR: 1.15, 95% CI: 1.05 to 1.25). For non-cardiovascular events, a significant association was observed only for pneumonia (HR: 1.22, 95% CI: 1.09 to 1.37). The associations stratified by country income levels are shown in online supplemental figure 2.

To further examine the impact of social isolation on mortality, case fatality rates within 90 days from the occurrence of new illnesses were assessed after a clinical event (myocardial infarction, stroke, heart failure, CVD, cancer, pneumonia, COPD and injury). Only for stroke was the case fatality rates higher in the socially isolated (online supplemental figure 3). The higher risk and case fatality rates of strokes could partly explain their increased risk of cardiovascular mortality.

### DISCUSSION

**Principal findings**

There are four main findings from our study: (1) social isolation is more common among women, older individuals, those with low levels of education or unemployed, living in urban areas and in richer countries; (2) social isolation is independently associated with increased risk of mortality after adjusting for conventional risk factors for CVD and this is observed in HIC, MIC and LIC. The mortality risk was partly explained by unhealthy behaviours and baseline comorbidities; (3) social isolation is associated with increased risk of incident stroke, CVD and pneumonia and (4) the population attributable fraction of social isolation on mortality is 2.4%.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Low (n=37863)</th>
<th>Middle (n=62855)</th>
<th>High (n=18046)</th>
<th>P for interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, 10-year increase</td>
<td>1.08 (1.02 to 1.15)</td>
<td>1.18 (1.14 to 1.22)</td>
<td>0.92 (0.86 to 0.98)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Women (vs men)</td>
<td>1.42 (1.22 to 1.64)</td>
<td>2.65 (2.47 to 2.84)</td>
<td>1.46 (1.30 to 1.64)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Education attainment level</td>
<td>2.64 (2.30 to 3.02)</td>
<td>1.16 (1.08 to 1.24)</td>
<td>1.16 (0.97 to 1.38)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Unemployed vs employed</td>
<td>3.52 (3.05 to 4.05)</td>
<td>0.93 (0.87 to 0.99)</td>
<td>0.99 (0.87 to 1.14)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Residence area urban (vs rural)</td>
<td>1.12 (0.99 to 1.26)</td>
<td>1.67 (1.56 to 1.79)</td>
<td>1.72 (1.49 to 1.98)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Current smoking (vs former or never smoking)</td>
<td>0.73 (0.62 to 0.87)</td>
<td>1.37 (1.27 to 1.48)</td>
<td>1.91 (1.66 to 2.19)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Low diet score (lowest tertile of AHEI) (vs the other two tertiles)</td>
<td>0.44 (0.37 to 0.52)</td>
<td>1.35 (1.27 to 1.44)</td>
<td>1.04 (0.92 to 1.17)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Current alcohol use (vs former or never drinking)</td>
<td>1.54 (1.27 to 1.87)</td>
<td>1.34 (1.26 to 1.44)</td>
<td>0.48 (0.43 to 0.55)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Physical inactivity (vs WHO recommended physical activity)</td>
<td>0.78 (0.67 to 0.90)</td>
<td>0.84 (0.77 to 0.92)</td>
<td>1.84 (1.60 to 2.12)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Number of comorbidities≥2 (vs one or no comorbidities)</td>
<td>0.86 (0.66 to 1.13)</td>
<td>1.03 (0.93 to 1.15)</td>
<td>1.24 (1.03 to 1.49)</td>
<td>0.01</td>
</tr>
<tr>
<td>Number of disabilities≥2 (vs one or no disabilities)</td>
<td>1.03 (0.90 to 1.19)</td>
<td>1.21 (1.13 to 1.31)</td>
<td>1.26 (1.06 to 1.48)</td>
<td>0.2</td>
</tr>
</tbody>
</table>

**Table 3** Variations in the association of factors that are associated with social isolation by income level of countries using multivariable logistic regression analyses

**ORs 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.
Comparison with other studies
Factors associated with higher prevalence of social isolation
Previous studies have shown that older adults are at risk of being socially isolated due to limited mobility caused by chronic illnesses and disabilities or reduced social ties through life events such as retirement or loss of their spouses, family members or friends.\textsuperscript{4,14,15} Our findings are consistent with this. Previous studies have reported contradictory findings on gender difference in the prevalence of social isolation.\textsuperscript{16-18} In our study, social isolation was more common among women. This may be due to our finding that women had lower rates of education and lower employment outside the home. In addition, lack of social supports might limit interactions with other individuals or groups, particularly in LICs and MICs. However, the higher rates of social isolation among women were observed in countries at all economic levels although most prominent in LICs.

Current smoking was more common in socially isolated people in our study, consistent with prior reports that those who are socially isolated are more likely to engage in unhealthy behaviours.\textsuperscript{16,19,20} In a French cohort,\textsuperscript{21} socially isolated men were more likely to be smokers and to indulge in heavy episodic drinking. Smokers might be forced to move to the periphery of social networks under pressure to avoid interactions with non-smokers due to concerns about the health risks of secondhand smoke.\textsuperscript{19,22} Socially isolated individuals were less likely to

Figure 2  The mortality risk of social isolation. Social isolation is associated with increase in the risk of all-cause, cardiovascular and non-cardiovascular mortality (A). The mortality risk associated with social isolation is greatest in HICs (B). The incidence rates of death were higher among the socially isolated and the mortality risk of social isolation was observed across regions with some random variations (C). HICs, high-income countries.
services for older people in HICs may account for this pattern. (2) The association between social isolation and current smoking was strong in HICs, while no association was observed in LICs. It is possible that smoking might encourage social isolation in HICs where smoking is less socially acceptable and public smoking is more strictly restricted. (3) In LICs, low education and unemployment were strongly associated with social isolation, suggesting that social connectedness is more affected by economic opportunities in LICs. (4) Urban residence was strongly associated with social isolation particularly in MICs and HICs. Urban communities might have weaker social connectedness through having fewer opportunities for social contacts, especially in highly developed countries.24

Variation in social isolation in urban versus rural settings

We found that social isolation was more common in urban areas than in rural areas. This might reflect a lack of interest in remaining connected or lack of interpersonal relationships in urban areas perhaps shaped by different values that focus on personal success.24 Furthermore, women in urban communities were more likely to be socially isolated than their counterparts in rural settings.

Despite the similarities in factors associated with social isolation, the magnitude of certain associations was different between rural and urban areas. For example, a stronger association was observed in women in urban areas, which might related to smaller social networks than their rural peers possibly due to greater economic independence or barriers of personal relationships with individuals in their communities which may hinder building social relationships. We also found that disability was more closely related to social isolation in rural areas as compared with urban areas. We speculate that larger interpersonal distances in rural settings may have a disproportionately large isolating effect on those with disabilities.26

Adverse health consequences associated with social isolation

In line with previous research, our study shows that social isolation is associated with a 26% increased risk of mortality with similar results for cardiovascular and non-cardiovascular mortality. The increased mortality risk was observed in different age groups, in men and women, in those living in rural or urban areas or in countries at different income levels. There has been a paucity of data regarding underlying mechanisms through which social isolation has influence on mortality. A study in the UK of people with a mean age of 57 years and mean follow-up of 6.5 years found that lifestyle behaviours, socioeconomic factors and mental health could explain 64% of the mortality risk associated with social isolation.27 In our study, mediation analyses of the association between social isolation and mortality attributed 21 of the association to unhealthy behaviours and baseline comorbidities. Previous studies have produced conflicting results on the association between social isolation and incident...
CVD. A meta-analysis reported an increased risk of coronary heart disease and stroke in individuals reporting social isolation, but more recent studies did not report support this conclusion. These inconsistent results could be explained by differences in study populations or definitions of social isolation. Our study found that social isolation was associated with increased risk of CVD and particularly stroke. The higher case fatality rates from stroke, along with their higher incidence of strokes, could partly explain the increased risk of cardiovascular mortality.

Evidence of an association between social isolation and non-CVD is scarce. We only found a significant association for pneumonia despite non-cardiovascular mortality being higher in people with social isolation. The increased risk of non-cardiovascular mortality might be explained by self-harm, substance abuse or suicide associated with social isolation though detailed information on these events were not available in this study. Previous studies showed that socially isolated individuals are at high risk of cancer, pulmonary disease as well as infection. Fewer social ties were associated with a higher risk of developing respiratory disease, which were consistent to our finding. The link between social isolation and non-CVD could be explained by older age, unhealthy lifestyle behaviours and pre-existing chronic illnesses, which may make them vulnerable to death. Also, social networks might play a role in resisting infection through regulation of the immune system.

Strengths and weaknesses of the study
Our study is the first study to examine the associations of social isolation with health outcomes in 20 countries from five continents, including LIC, MIC and HIC and from urban and rural communities. The large size of our study and the diversity of the populations has allowed examination of the consistency or heterogeneity of associations in different settings and in different subgroups of the population.

Our study has some potential limitations. First, it is not possible to exclude unmeasured confounding factors such as feelings of loneliness, history of substance abuse and criminal records or victimhood in this observational study, although a wide range of potential explanatory factors were studied. Second, reverse causality could be a concern. We conducted sensitivity analyses to address this concern. In those analyses, individuals with disease at baseline or those who developed clinical outcomes within the first 2 years of follow-up were excluded, which did not alter our results. Third, covariates adjusted for the analyses were assessed only at baseline, but demographic data including country income levels and other socioeconomic factors could have changed over the study period. However, since the association between social isolation and clinical outcomes is similar in HIC, MIC and LIC, a shift in the categorisation of countries from one economic group to another would not be expected to materially alter our results. Regarding our mediation analyses, potential mediators were not independent of each other since some factors (ie, smoking, alcohol) may mediate not only mortality but also some comorbidities (ie, CVD, cancer). Thus, the estimates for the mediation effects may be affected by other factors. Besides that, since potential mediators were only assessed at baseline, causal relationship among social isolation, potential mediators and outcomes cannot be reliably derived using our current study design. In our analyses, only high alcohol intake defined as >14 drinks/week for women or 21 drinks/week for men was associated with increased risk of mortality in a multivariable Cox regression analysis (HR: 1.48, 95% CI: 1.31 to 1.67), pointing to the particular importance of socially isolated people taking care about excessive alcohol intake. Finally, our social isolation scale did not include information on living alone, subjective social isolation (ie, loneliness), or social network size that may provide a more nuanced reflection of social isolation. Future studies should include such information as well as new concepts which emerge in this field of scholarship.

Implications for clinicians and policy-makers
Our findings support for strategies to address several factors (lower socioeconomic status and unhealthy lifestyles) and consequences of social isolation. Healthcare workers and policy-makers may wish to consider social isolation as an added factor in identifying individuals at higher risk who may benefit from specific measures that go beyond the usual preventive and treatment strategies, to mitigate their higher risk.

CONCLUSIONS
Our study is the first to demonstrate the associations between social isolation and health outcomes in middle-aged community-dwelling adults from urban and rural sites in HIC, MIC and LIC. Social isolation is associated with increased risk of death and morbidity among diverse populations across the world. It should be considered as an added risk factor to that conferred by conventional risk factors. The best ways of addressing this issue, whether intensive use of proven therapies and lifestyle modification or measures to improve social support remain unclear and may be context dependent.

Author affiliations
1Population Health Research Institute, Hamilton, Ontario, Canada
2Department of Medicine, McMaster University and Hamilton Health Sciences, Hamilton, Ontario, Canada
3Department of Health Research Methods, Evidence and Impact, McMaster University and Hamilton Health Sciences, Hamilton, Ontario, Canada
4ECOHOST, LSHTM, London, UK
5Harvard Center for Population and Development Studies and Department of Society and Human Development, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA
6International Research Center, Hospital Alemão Oswaldo Cruz, São Paulo, Brazil
7Department of Medicine, Queen’s University, Kingston, Ontario, Canada
8Faculty of Health Sciences, Simon Fraser University, Vancouver, British Columbia, Canada
9Eternal Heart Care Centre & Research Institute, Jaipur, India
Appendices


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, e.g., 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) × 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 × 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


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 the following formula: \[ \text{PERM} = \frac{\text{Hazard ratio}_1 - \text{Hazard ratio}_2}{\text{Hazard ratio}_1} \times 100 \] (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

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

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

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

<table>
<thead>
<tr>
<th>Subgroups</th>
<th>All-cause mortality</th>
<th>p for interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age category</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 60 years</td>
<td>1.42 (1.27-1.60)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>≥ 60 years</td>
<td>1.24 (1.11-1.39)</td>
<td></td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>1.24 (1.11-1.40)</td>
<td>0.0003</td>
</tr>
<tr>
<td>Men</td>
<td>1.49 (1.33-1.67)</td>
<td></td>
</tr>
<tr>
<td><strong>Residence area</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>1.37 (1.23-1.54)</td>
<td>0.3</td>
</tr>
<tr>
<td>Rural</td>
<td>1.29 (1.15-1.45)</td>
<td></td>
</tr>
</tbody>
</table>
Supplemental figures legends

Supplementary Figure 1. A flow chart of the study population
Legend: The study sample was obtained from the PURE study dataset. Seven countries lacking follow-up information were excluded. China was excluded because relevant information on social isolation were not collected. Several individual sites* were excluded. Participants aged <35 or >70 were excluded. Finally, participants not providing information on social isolation were excluded.
*Sites from Argentina, Chile, Peru, and Poland participating in the PURE study at a later phase of the study

Supplementary Figure 2. Associations between social isolation and incident diseases stratified by county income levels
Legend: The magnitude of disease risks associated with social isolation is greatest in the LICs except for incident myocardial infarction and COPD.

Supplementary Figure 3. Case fatality rates (95% CI) over 90 days from the occurrence of a new illness in those with and without social isolation
Legend: The case fatality rate from stroke is higher in individuals with social isolation than those without while no significant differences are observed for those from the other illnesses.
MI=myocardial infarction. CVD=cardiovascular disease. COPD=chronic obstructive pulmonary disease.
N=202,524, 28 countries

N=14,557 exclude 7 countries

N=187,967

N=47,927 exclude China

N=140,040

N=19,087 exclude several sites*

N=120,953

N=1,059 exclude age<35 and age>70

N=119,894

N=1,130 without social isolation scale

N=118,764
BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance
placed on this supplemental material which has been supplied by the author(s)

BMJ Global Health
doi: 10.1136/bmjgh-2020-004124


Supplemental material