World Health Organization's Study on Global Ageing and Adult Health (SAGE) Wave 1: Health states and objective health measures in six low and middle income countries

Chatterji, Somnath
Kowal, Paul
Verdes, Emese

World Health Organization, Multi-Country Studies unit. Geneva, Switzerland.

ABSTRACT

Over 60% of the world's population aged 60 years and older are currently living in less developed countries (UN 2009), with an average annual growth rate almost three times that of more developed countries (2.5% versus 0.9%). However, data on health outcomes of older adults that can be directly compared in these countries is lacking. A key question, as these populations age, will be their health status over time and across cohorts. The implications for health policy planning and resource allocation decisions will be critical depending on their needs. WHO's Study of Global Ageing and Adult Health (SAGE) provides an important data collection platform to monitor health and health-related outcomes and their determinants in these populations over time. We report data from SAGE Wave 1 and present results comparing self-reported and measured health status in six low and middle income countries. We also describe strategies to adjust for systematic reporting biases and evaluate the performance of these methods. There were identifiable reporting biases, and the methods used to correct for these biases adjusted mainly for country-reporting biases. Clear declines in health by age and differentials by socioeconomic status, gender and residence were evident. As these cohorts are followed, issues related to the compression of morbidity can be addressed more systematically.

INTRODUCTION

Over 60% of the world's population aged 60 years and older are currently living in less developed countries (UN 2009), with an average annual growth rate almost three times that of more developed countries (2.5% versus 0.9%). This growth in the older population is occurring in parallel with increasing income inequality, disparities in access to health care and social support systems, and widening health gaps as a result of complex disease burden patterns and globalization of health risks. In many less developed countries, these issues are compounded for individuals by a lifetime of accumulated health risks associated with poverty and inadequate access to health care. Yet, few countries have the age-specific health and functioning data necessary to determine basic health parameters, much less which morbidity trajectory their respective aging populations are following: expansion (Gruenberg 1977, Scheider and Brody 1983), compression (Fries 1980, Fries 2003) or dynamic equilibrium (Manton 1982, Manton 2006).

The World Health Organization, with support from the US National Institute on Aging, has created a longitudinal data collection platform to generate cross-nationally comparable health data for the purposes of monitoring population health over time. WHO's Study on Global Ageing and Adult Health (SAGE) is a nationally representative longitudinal cohort study in six low and middle income countries (LMICs), China, Ghana, India, Mexico, Russia and South Africa, to measure health and health-related outcomes, and their determinants, and to understand the relationship between health and well-being over time.

One of the major challenges comparing self-reported health status across populations is that respondents often have systematic reporting biases. Typically, the measurement of health state relies on self-reported responses in surveys and the resulting data take the form of ordered categorical (ordinal) responses. However, the way people report their own health varies

systematically with factors such as education, sex, age, or other cultural factors. Various people use different response category cut-points across cultures or population sub-groups, and this *'response shift'* implies that self-report categorical data are not comparable across individuals (Murray 2003). Consequently, these responses on ordinal scales cannot be directly used to measure health states without adjustment. In order to ensure that data from self-reported interview surveys are truly comparable, four essential steps need to be followed: 1) agreement on a common conceptual framework for the measurement; 2) agreement on a measurement strategy that identifies a parsimonious set of domains and items for measurement; 3) *ex-ante* harmonization of questions, response scales and calibration strategies across languages and population groups; and, 4) *ex-post* harmonization of data using the calibrated responses. The first three of these approaches are most often employed in international efforts at data collection but, though necessary, are often not sufficient. The focus of this paper is strategies that can be used for *ex-post* harmonization.

Statistical methods must be devised that detect systematic reporting biases, determine factors that are responsible for these systematic biases and correct for these in order to approximate the true underlying level in the latent quantity of interest. In recognition of this, the WHO World Health Surveys (WHS), used a set of questions across a core set of domains to measure health states, and employed vignettes to detect and correct for biases in self-report in order to adjust for response category cut-point shifts (Ustun 2003).

An anchoring vignette is a brief description of a concrete level on a given health domain, with multiple vignettes used per health domain reflecting different levels on a latent scale. Respondents are asked to respond to a vignette using the same questions and response scales applied to self-assessments on that domain, as if the person described in the vignette is like the respondent herself / himself. Vignettes fix the level of ability on a domain so that variation in

categorical responses is attributable to variation in response category cut-points. There are two key requirements for the use of anchoring vignettes: *response consistency:* this requires that an individual will have the same cognitive mind set when evaluating hypothetical scenarios as when providing a self-assessment; and *vignette equivalence:* the requirement that the underlying domain levels represented in each vignette are understood in approximately the same way by all respondents, irrespective of their age, sex, income, education, country of residence or other characteristics.

We use data from SAGE to assess declines in health with age in six LMICs and the impact of systematic reporting biases. We compare results across age, socioeconomic variables and countries.

DATA AND METHODS

Data from SAGE Wave 1 on adults aged 50 years and older, from China, Ghana, India, Mexico, Russia and South Africa, were used in this analysis.

SAGE is a household survey that interviewed respondents drawn from a nationally representative frame. All respondents had a known non-zero probability of selection.

The survey instrument collected information on sociodemographic variables such as age, sex, place of residence, education, marital and economic status, and employment. Health status was assessed in eight domains of functioning: affect, cognition, interpersonal relationships, mobility, pain, self-care, sleep and energy, and vision. Respondents were asked the extent to which they have difficulty in, or experience problems with, carrying out a task or action. Two questions are asked about each health domain.

In addition, respondents were also asked to respond to a set of vignettes about the health domains where they were asked to imagine they are the person described in the vignette, and then answer the same questions that were asked about their own health status for that particular domain. Responses were ordered in the same manner as for the self-report from no difficulty (or problem) to complete difficulty (or problem).

In addition to self-reported health status, we also used data from the performance tests used in SAGE. Tests of cognition included a measure of verbal fluency assessed using a category naming task of animals, a word list for immediate and delayed recall, and digit span forwards and backwards for assessment of working memory. Normal and rapid walking speed and grip strength were also measured. A combined score on these tests was created using factor analysis to get an overall combined measure of performance.

A composite health status score was created using a partial-credit model using graded response based on Item Response Theory. Data from the vignette responses were used to adjust for the self-report and to identify shifts in cut-points. A score on the latent trait was created using the Binormal Hierarchical Ordered Probit based on the Compound Hierarchical Probit model (Tandon 2003).

RESULTS

Background characteristics of the sample are shown in Table 1.

Vignette results by age groups (18-49, 50-64, 65+) are presented in Figure 1, showing robust results with no systematic reporting differences by age. The expected trends in vignette-adjusted health scores by sex, age, education level and income level are evident across all countries.

Performance measures in three domains (grip strength, timed walk and cognitive function) are presented. Initial analyses suggest a correlation between the vignette-adjusted health score and each performance test, but curiously not significantly different by age. A weak negative correlation was found between one health domain (mobility) and timed regular-paced walk at all ages, with a moderate positive correlation with grip strength (0.2-0.3) and cognitive function (0.3-0.4). These correlations were not significantly different by age group.

The presence of depression, however, does create systematic reporting biases in health by age and education levels. More educated depressed respondents and younger respondents systematically underestimated their health, whereas, older depressed respondents systematically overestimated their health.

Figure 2 shows the estimated median cut-points from the Binormal Hierarchical Probit model (BiHOPIT) using the vignettes data and controlling for a range of covariates such as age, sex, education, income, and other self-reported health variables such as stress, satisfaction with health, depression and country of residence. The median cut-points by country are clearly different, suggesting that different standards are used by respondents in these countries when reporting their mobility.

Figure 3 shows the distribution in the five categories of the scale for the unadjusted self-reported data, as well as the adjusted results from the BiHOPIT model, using China as the reference category. Once again, there are clear shifts in distribution. For example, while nearly 10% of respondents in India reported severe difficulties with moving around, the adjusted model would predict only about 2% of respondents in this category, thereby narrowing the difference with China as compared to the unadjusted self-report.

Figure 4 shows the results of the adjusted and unadjusted scores comparing China with India. Using the China respondents cut-points as the reference, it is evident that respondents from India in their self-reports systematically overestimate their difficulties in health when the reporting biases are not adjusted, as the score derived from the BIHOPIT model systematically shifts these respondents scores upwards.

Figure 5 shows the adjusted and unadjusted scores by age for each of the six countries. Clear declines by age are evident across all countries with very similar patterns of decline by age.

Figure 6 shows the health status across the pooled sample by income quintiles. As expected, health status is considerably better in richer as compared to poorer quintiles with a clear gradient.

DISCUSSION

Our analyses show clear declines in health status with age but with some differences in the trajectory of decline across countries. The self-reported measure of health status that measures health status as a vector of functioning on a chosen parsimonious set of domains seems to capture variations by age and socioeconomic status well. It is weakly correlated with the measured tests overall with a fair correlation between the individual domains and the respective performance tests.

The BIHOPIT model used in this paper, besides using the covariates of country, age, income and sex, also included other covariates such as overall self-reported global health, stress, depression, satisfaction with health, overall quality of life and self-reported happiness. In addition, the measured tests were included to predict the latent trait of health adjusted for reporting biases with the best accuracy, given the data collected. Though clear reporting biases are evident mainly by country, there are no systematic biased detected by age, income status or

sex. This could possibly be because the study is adults aged 50 years and older. Another limitation could be that vignettes are not particularly well understood by older adults in LMICs and hence do not have the properties of response consistency in this population. The present study has not systematically examined this possibility. Another consideration is the possible need for better descriptions in the vignettes that more clearly describe levels of functioning within a domain. The cognitive demand of the vignette task may impose burdens on older adult populations in LMICs that may make the task not very useful in large scale national surveys. Alternatively, small focused surveys could be used to identify these biases if they exist and then develop strategies to generalize the results to national surveys.

The incorporation of measured tests in large national surveys is a strength that needs to be further developed. Their utility in the development of a measure of frailty to identify vulnerable populations, combined with self-reported health status, is likely to be borne out in these populations in LMICs. Declines in health status in these cohorts needs to be monitored in subsequent waves of SAGE to assess the velocity of decline in health in these populations. We can then begin to address questions around compression of morbidity in this rapidly growing segment of the world's population. Health and social systems will need to be prepared to deal with this next public health challenge in LMICs.

Works Cited
Fries JF. Aging, natural death, and the compression of morbidity. *NEJM.* 1980;303:130-5.

Fries JF. Measuring and monitoring success in compressing morbidity. *Ann Intern Med*. 2003;139:455-9.

Gruenberg EM. The failures of success. *Milbank Memorial Fund Quarterly: Health and Society*. 1977;55:3-24.

Manton KG. Changing concepts of morbidity and mortality in the elderly population. *Milbank Mem Fund Q Health Soc.* 1982;60:183-244.

Manton KG, Gu Xiliang, Lamb VL. Change in chronic disability from 1982 to 2004/2005 as measured by long term changes in function and health in the U.S. elderly population. *PNAS*. 2006;103:18374-9.

Murray CJL, Ozaltin E, Tandon A, Salomon JA, Sadana R, Chatterji S. Empirical evaluation of the anchoring vignette approach in health surveys. In: Murray CJL and Evans D (eds.). *Health Systems Performance Assessment: Debates, Methods and Empricism.* World Health Organization, Geneva, 2003.

Schneider EL, Brody JA. Aging, natural death and the compression of morbidity: another view. *N Engl J Med*. 1983;309:854–7.

Tandon A, Murray CJL, Salomon JA, King G. Statistical models for enhancing cross-population comparability. In: Murray CJL, Evans DB, eds. *Health systems performance assessment: debates, methods and empiricism.* Geneva, World Health Organization, 2003.

Ustun TB, Chatterji S, Mechbal A, Murray CJL, WHS Collaborating Groups. The World Health Surveys In: Murray CJL, Evans DB, eds. *Health systems performance assessment: debates, methods and empiricism.* Geneva, World Health Organization, 2003.

Table 1. Sociodemographic characteristics, SAGE Wave 1, by country.

Characteristics	China	India	Mexico	Russia	South Africa	Ghana
Total subject (N)	13,265	7,527	2,734	3,163	3,842	4,768
Sex (%)						
Men	47.0	38.7	38.3	37.1	42.5	52.8
Women	53.0	61.3	61.7	62.9	57.5	47.3
Age Group (%)						
18-49	11.0	41.6	15.6	8.5	9.1	15.8
50-59	39.5	26.2	15.8	34.1	40.1	33.1
60-69	26.3	19.9	34.2	24.9	29.2	23.6
70-79	18.5	9.4	22.6	23.7	15.7	19.3
80+	4.7	2.9	11.8	8.8	5.9	8.3
Highest education level (%)						
Completed college/university	6.2	10.8	6.4	20.8	7.3	7.4
Completed high school	17.8	19.5	4.8	50.7	10.7	34.7
Completed secondary	28.0	22.7	13.4	18.9	19.5	11.2
Completed primary	25.4	27.9	28.6	7.8	31.1	25.2
Less than primary	22.6	19.1	46.8	1.9	31.4	21.6
Current marital status (%)						
Not married	15.8	22.4	36.5	42.5	47.0	40.0
Married	84.2	77.6	63.5	57.5	53.0	60.0
Always live in the area (%)	55.7	56.7	75.7	53.6	71.3	65.9
Currently working (%)	50.6	64.2	53.5	36.8	32.3	73.2

Figure 1. Vignette and self-reported mobility ratings* by age groups, pooled SAGE Wave 1 data.
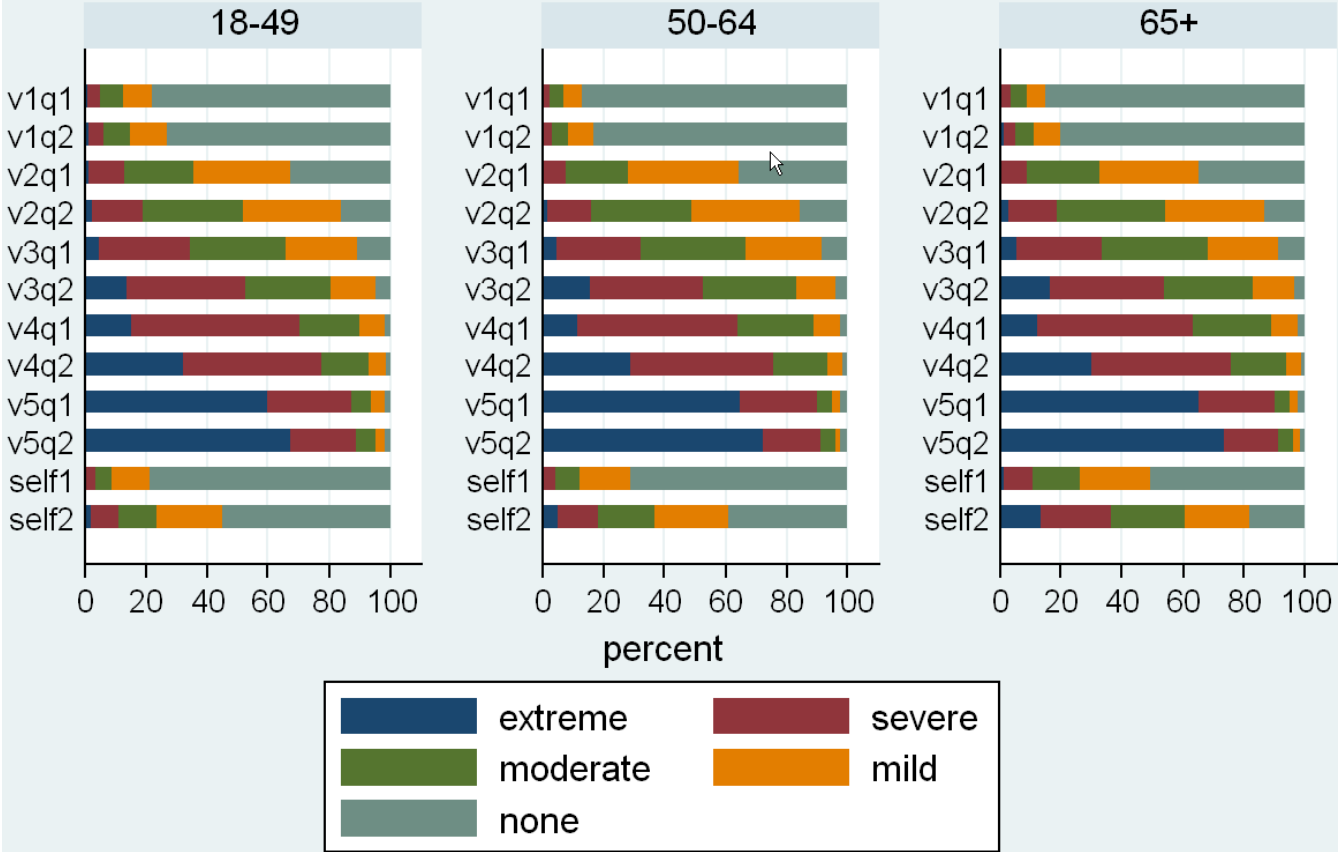
*Note: Self 1 and Self 2 are respondent responses about their own mobility levels, with v1-v5 increasing levels of difficulty on a latent scale of mobility from "easiest" to "hardest" vignette.

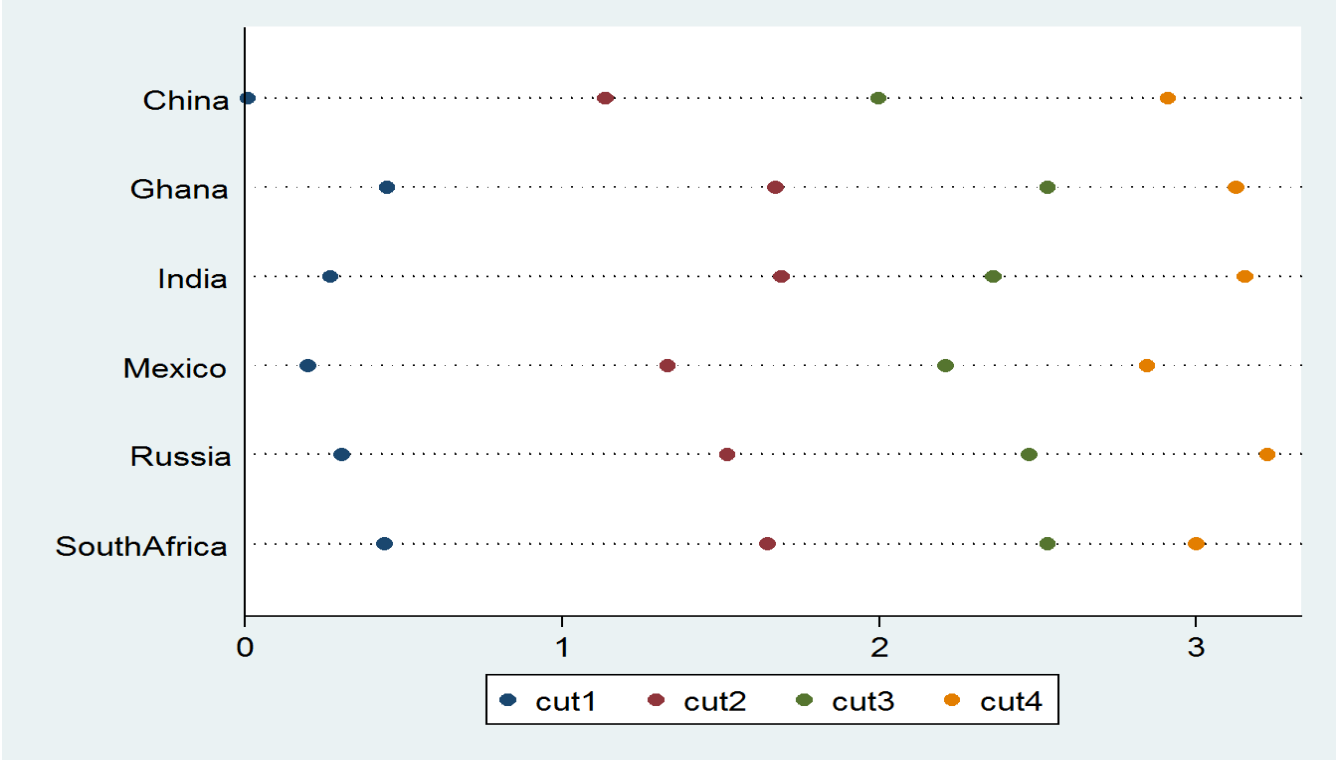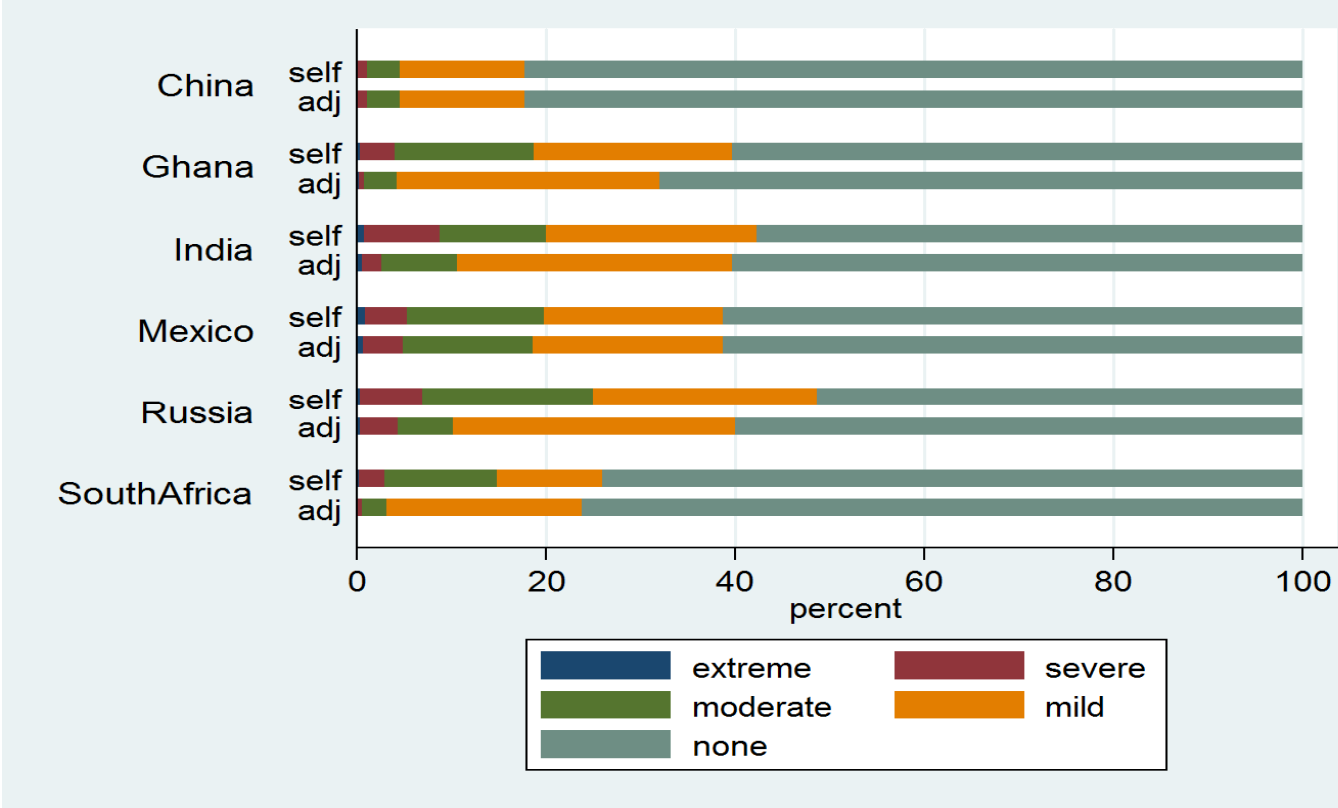Figure 2. Estimated median cut-points for mobility, by country. Adjusted results using BiHOPIT.

Figure 3. Adjusted and unadjusted distributions* for the mobility health domain, by country.

*Note: "self" refers to unadjusted self-report, and "adj" is the BiHOPIT adjusted result on the y-axis.

Figure 4. Adjusted (adjIRT) and unadjusted self-reports (origIRT), comparison between SAGE India and China.
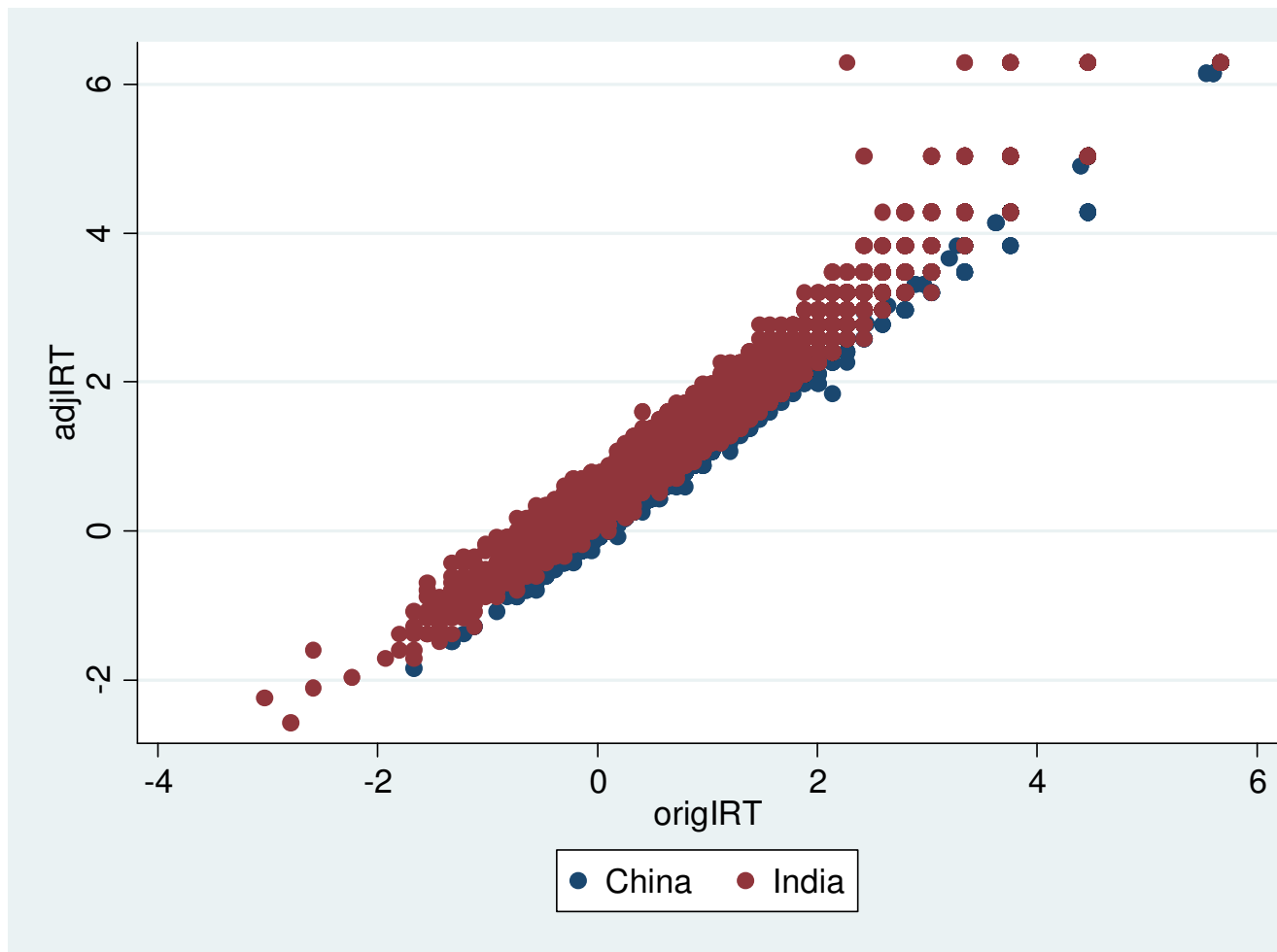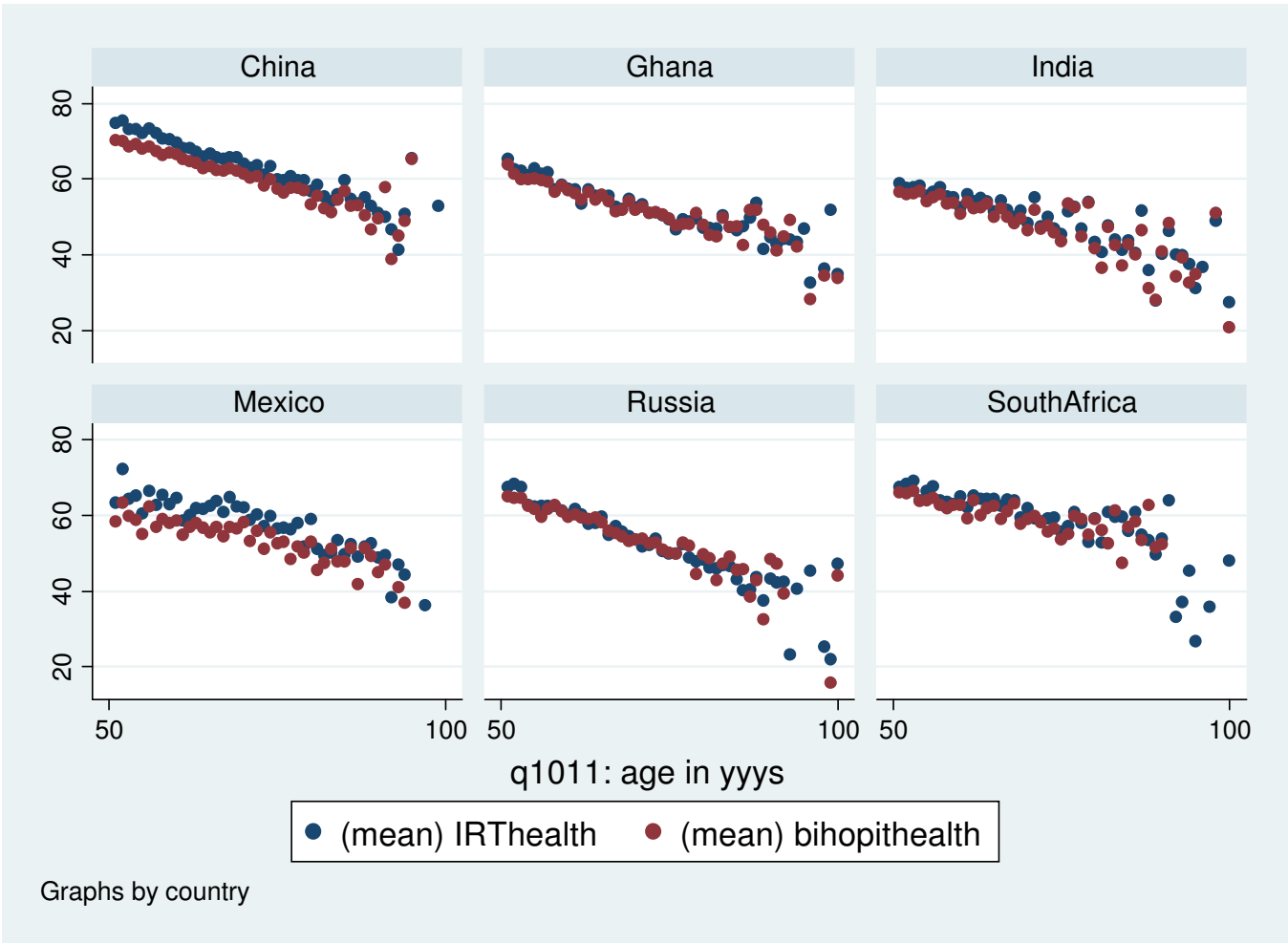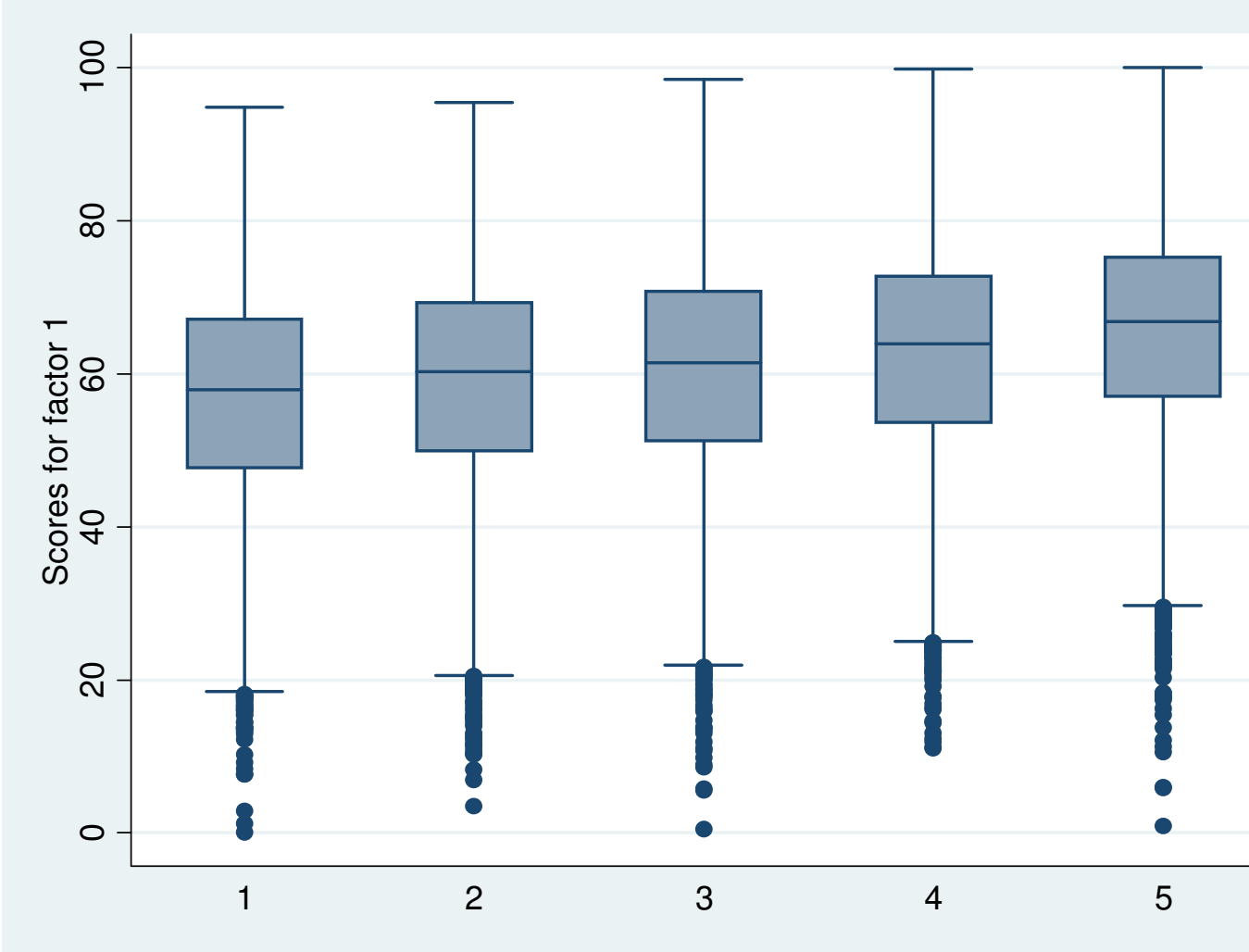
Figure 5. Declines in health status* by age, adjusted and unadjusted, by country

*Note: The composite health scores on the y-axis are scaled from 0 (worst) to 100 (best) health.

Figure 6. Health status* by wealth quintiles, pooled SAGE Wave 1 data from six countries.

*Note: The composite health scores on the y-axis are scaled from 0 (worst) to 100 (best) health.