UNITED NATIONS UNIVERSITY
UNU-WIDER

WIDER Working Paper 2023/27

Differential health reporting error among older adults in India

Anna Choi,[1] Arnab K. Basu,[2] Nancy H. Chau,[3] and T.V. Sekher[4]

February 2023

Abstract: This paper studies the education gradient associated with health reporting errors for two highly prevalent non-communicable diseases among older adults in India. We leverage a novel data set—the Longitudinal Aging Study in India (2017–18) panel survey—to unpack the sources of health reporting error in a developing-country context for the first time. Our analysis points to a statistically significant level of false negative reporting (or over-reporting) for both high blood pressure and depression by those with no schooling and less than middle school education relative to their more highly educated peers. Interestingly, this result is driven by those whose income is above the poverty line. We further find false negative reporting for high blood pressure to be more prevalent among women with no schooling.

Key words: high blood pressure, depression, reporting error, education

JEL classification: I1, O5

Acknowledgements: This paper has benefited from comments by seminar participants at UNU-WIDER, Helsinki, and session participants at the Southern Economics Association Conference and European Health Economics Association Conference. We are particularly grateful to Omar McDoom, Rachel Gisselquist, Rahul Lohati, Kunal Sen, Enrico Nichelatti, and Rodrigo Oliveira for their insightful suggestions. All remaining errors and omissions are ours.

[1] Sejong University, Seoul, South Korea, anna.choi@sejong.ac.kr; [2] Cornell University, Ithaca, NY, USA, and IZA, Bonn, Germany: ab362@cornell.edu; [3] Cornell University, Ithaca, NY, USA, and IZA, Bonn, Germany, hyc3@cornell.edu; [4] International Institute for Population Sciences (IIPS), Mumbai, India, tvsekher@gmail.com

This study has been prepared within the UNU-WIDER project Academic excellence.

Information and requests: publications@wider.unu.edu

ISSN 1798-7237 ISBN 978-92-9267-335-2

https://doi.org/10.35188/UNU-WIDER/2023/335-2

Typescript prepared by Luke Finley.

United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Finland and Sweden, as well as earmarked contributions for specific projects from a variety of donors.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

1 Introduction

Target 3.4 of the 2030 Sustainable Development Goals (SDGs) adopted by all UN member countries benchmarks a reduction in premature mortality from non-communicable diseases by one-third through prevention and treatment, and an improvement in mental health and wellbeing. The purpose of these targets is to set supply-side guidelines for governments across the globe to invest resources in critical areas of healthcare delivery, including those for the betterment of mental health. After all, government spending on health and health-related activities in the developing world can be notoriously low—around 3 per cent of gross domestic product (GDP) for the South Asian group of countries (World Bank 2019)[1]—and the targets provided by the SDGs act as signals to increase spending in key areas.

Accompanying the supply-side push, several demand-side interventions have been introduced in developing countries to raise health awareness and effect proactive health-seeking behaviour. These include campaigns to change behavioural attitudes towards cleanliness and disease prevention via inducing more handwashing, smoking cessation campaigns, promoting healthy eating habits, and instilling the virtues of regular physical activity. However, bottlenecks are plenty—from a rise in chronic non-communicable diseases like hypertension and diabetes (older adults suffer more from chronic rather than acute illnesses) to income traps (low income associated with lower ability to access preventative care) to gender bias in health-seeking behaviour (women are less likely to seek medical care especially where the majority of the doctors are male) and societal norms and mistrust of modern medicine (reliance on traditional doctors and vaccine hesitancy).[2]

Demand- or supply-side interventions notwithstanding, health-seeking behaviour is guided by awareness of one's health status and therein lies the contribution of this paper. Leveraging a unique survey containing the self-reported and objectively measured health conditions of adults aged 45 years or older in India, we contribute to the literature on differential health reporting error by education levels in a developing-country setting.[3] The Longitudinal Aging Study in India (LASI) conducted over 2017–18 catalogues self-reports and objective measures of high blood pressure, depression, and physical activity, allowing for unique insights into the magnitude of reporting error for a widely prevalent non-communicable disease—hypertension—and a neglected health condition in developing countries—depression. Notable here is the fact that LASI is the only nationally representative dataset available from a developing country that captures both self-reported and objective measures of depression.

While studies on differential health reporting error are more common in developed countries, data paucity makes such analysis impossible to conduct for developing countries. Our paper is thus unique in this regard. This is also the first paper to our knowledge to investigate the differential reporting error by education levels in self-reported depression diagnosis compared with depressive

---

[1] For North America, health spending was roughly 16 per cent and for OECD countries it was 12.5 per cent of GDP in 2019 (World Bank 2019).

[2] For an analysis of the success of smoking cessation campaigns in South Asia see Iqbal et al. (2022), for an analysis a large-scale handwashing intervention and its impact on respiratory illness in Bangladesh see Najnin et al. (2019), for inadequate number of facilities to treat chronic non-communicable diseases in India see Dey et al. (2012), for how the low availability of women physicians in rural areas of India is negatively affecting health-seeking behaviour see Bhan et al. (2020), and for the origins of mistrust of modern vaccines in Africa see Lowes and Montero (2021).

[3] In effect, our analysis here follows the methodology of Choi and Cawley (2018), undertaken for the United States.

symptoms as measured by the Center for Epidemiological Studies-Depression (CES-D) scores. In what follows, we discuss in turn the importance of analysing the health–education gradient, the magnitude of high blood pressure and depression among the Indian population, and earlier research conducted via the use of the LASI pilot study in 2010.

1.2 The health status reporting–education gradient

There are at least three pathways through which the reporting of health status and education can be related. First, health and education may be causally related—higher levels of education lead to better health. Second, selection may occur whereby the healthy become better educated. And finally, health and education may both be correlated with a confounder. For example, good health in later life and high educational attainment are both a result of early life interventions such as better nutrition and vaccinations. Our focus is on the selection problem given the cross-section data at our disposal, which precludes establishing a causal relationship between educational attainment and health. In addition, variables such as income are endogenous with current health status whereas education, completed several years prior for adults aged 45 years or older, is not. It is thus unsurprising that the health–education gradient is a popular subject for research. Several studies based in developed countries, primarily the US, have established that adults with higher educational attainment have better health and lifespans compared with their less-educated peers. For instance, Grossman (2008) finds that education (in terms of years of formal schooling) is one of the most important positive correlates of health. Different mechanisms and factors could drive this relationship. Cutler and Lleras-Muney (2010) find that income, health insurance, and family background account for 30 per cent, knowledge and cognitive ability measures account for another 30 per cent, and social networks account for 10 per cent of the relationship between education and health.

However, most of these studies that explore the health status reporting–education gradient rely on self-reported health measures, which could be prone to various biases and errors. It is possible that the better educated are more accurate in reporting their health conditions and behaviours, as evidenced in the studies undertaken by Choi and Cawley (2018) and Johnston et al. (2009). Social desirability bias may also influence how individuals respond to survey questions on health status or behaviours, especially when a more socially acceptable response is deemed to be expected. It is also possible that education is correlated with other skills and knowledge, such as health literacy. Van der Heide et al. (2013) find evidence of health literacy possibly mediating poor self-reported health status among those with lower levels of schooling.

For our geographical context, India, studies on the health status reporting–education gradient are sparse. Cramm et al. (2015) find low education levels to be a significant predictor of smoking status and lower levels of physical activity. Raghupathi and Raghupathi (2020) find that tertiary education is critical in positively influencing infant mortality, life expectancy, child vaccination, and healthcare enrolment rates.

1.2 High blood pressure and depression in India

The annual number of deaths from cardiovascular disease (CVD) in India was projected to rise from 2.26 million (1990) to 4.77 million (2020), according to Huffman et al. (2012). This is not surprising given that India has one of the highest burdens of CVD worldwide. Prevalence rates of CVD have been estimated to have ranged from 1.6 per cent to 7.4 per cent in rural populations and from 1 per cent to 13.2 per cent in urban populations over the past several decades (Huffman et al. 2011). Geldsetzer et al. (2018) estimate an even higher prevalence of hypertension in the urban areas of India, and compared with the poorest quintile, the richest quintile had only a moderately higher probability of being diagnosed with hypertension. Interestingly, Geldsetzer et al. (2018) do not find any difference in the probability of hypertension across educational attainment. However, Pednekar et al. (2011) find that CVD mortality is significantly greater among Indian men with low educational status, while the association is not clear in women.

The prevalence of depression, and even more so its causes and consequences, is an under-researched area, particularly in a developing-country context. Two recent studies by Bernadine (2021) and Dandona et al. (2020) offer some key insights into the extent of the issue. Bernadine (2021) reiterates the World Health Organization (WHO) finding that the economic loss associated with the prevalence of mental health disorders in India is extremely high, at US$1.03 trillion. In addition, data from the National Mental Health Survey (NMHS, 2015–16) point to mental health disorders disproportionately affecting households with lower income, less education, and under-employment (cited in Bernadine 2021). Further, median out-of-pocket expenditure by families on treatment and travel to access care amount to 1,000–1,500 Indian rupees (INR) per month, equivalent to approximately US$13–20: a large sum for poorer families. Dandona et al. (2020) estimate that 45.7 million people had depressive disorders in India in 2017. The highest prevalence was observed in the Southern Indian states of Tamil Nadu, Kerala, Goa, and Telangana (which rank high on the Social Development Index/SDI), Andhra Pradesh (which ranks among the middle SDI group of Indian states), and Odisha (which ranks among the low SDI group of Indian states). Dandona et al. (2020) further find no difference in overall prevalence between men and women and that depressive disorders increase with age, with the highest prevalence observed in older adults, especially among women starting at 45 years. Neither of these two studies, however, analyse the prevalence of mental disorders by educational attainment or income level.

1.3 Research using LASI pilot study (2010) and Wave 1 (2017–18)

The pilot Longitudinal Aging Study (LASI) was undertaken in 2010, and a few studies based on this survey have been published. Cramm et al. (2015) is the first such study to find that subjective (compared with objective) health measures underestimate the health status of older adults, and, as mentioned earlier, also finds a negative relationship between educational attainment and both smoking status and physical activity. Vellakkal et al. (2013) find that self-reports (compared with objective measures) underestimate the prevalence of hypertension among the poor, while Onur and Velamuri (2018) find that measurement error in self-reported hypertension and lung disease is over 80 per cent—larger than what is found in high-income countries. However, there are no LASI pilot-based or Wave 1 studies that have analysed either the differential health reporting status or the health–education gradient of high blood pressure/hypertension and depression prevalence among older adults. Lee et al. (2022) analyse the prevalence of hypertension using data from LASI Wave 1 (2017–18) and the pilot study (2010) and find that almost 46 per cent of adults aged 45 or older are hypertensive. Of those diagnosed with hypertension, 39 per cent take anti-hypertensive medication and access to public healthcare is a key predictor of hypertension treatment among the lowest income group.

With the background in place, we discuss our descriptive statistics, empirical strategy, results, and robustness checks in the next sections.

2 Data and descriptive statistics

We use the first wave (2017–18) data from the LASI, which contains self-reported and objective measures of high blood pressure, depression, and physical activity (walking). The LASI contains a nationally and state representative sample of adults aged 45 and older in India and covers 30 states and six Union Territories of India (the Sikkim survey was undertaken in 2021).[4] Given the widespread prevalence of CVD in India and the paucity of studies on depression in developing countries, we focus on the education gradient of reporting error of these two health conditions.

In the survey, respondents report whether any health professional has ever diagnosed them with high blood pressure and any neurological or psychiatric problems such as depression.[5] High blood pressure is defined as self-reported physician diagnosis or elevated blood pressure (BP) on measurement of systolic BP $\geq$ 140 mmHg or diastolic BP $\geq$ 90 mmHg. For depression, respondents' answers on the ten-item Center for Epidemiological Studies-Depression (CES-D) questionnaire are used. The CES-D is a standard evaluation tool for mental health status; the scores from the questions can serve as a close proxy for depression screening and have good predictive accuracy when compared with the 20-item scale among older adults (Andresen et al. 1994; Gellis 2010). If the respondent has a score of ten or above, we consider them to have depressive symptoms (Andresen et al. 1994; Gellis 2010; Radloff 1977). Later in the paper, we run robustness checks using different thresholds of CES-D (25th, median, 75th, and 90th percentile) to categorize someone as suffering from depression. We consider the respondent to have high blood pressure with the following objective measurements: average systolic 140 mmHg or higher, and average diastolic 90 mmHg or higher. After removing observations with missing data on education, we have 64,866 older adults (45 and older) in our analysis sample.

We next turn to the descriptive statistics. As Table 1 shows, roughly 50 per cent of respondents in our sample do not have any formal education, the mean age is 60 years, 45 per cent are male, and 91 per cent are married. A large percentage—45 per cent—are actively employed and, perhaps not surprisingly for a country like India, only 20 per cent have health insurance coverage. Nevertheless, a high percentage of respondents—72 per cent—have visited a healthcare facility within the past year of the survey.

---

[4] Perianayagam et al. (2022) provides a comprehensive overview of the cohorts surveyed in LASI Wave 1 (2017–18) while Bloom et al. (2021) explain how the LASI survey can advance ageing research in response to the challenges posed by population ageing.

[5] The questionnaire wording is: 'Has any health professional ever diagnosed you with any neurological, or psychiatric problems such as depression, Alzheimer's/dementia, unipolar/bipolar disorders, convulsions, Parkinson's etc.?' There is a subsequent question that asks respondents to choose which neurological or psychiatric problem they were diagnosed with, including depression.

Table 1: Descriptive statistics

Variable	N	Mean	Standard deviation
No education/schooling	64,866	50.4%	0.50
Less than middle school	64,866	23.0%	0.42
Middle or high school education	64,866	20.9%	0.41
College education	64,866	3.6%	0.19
Graduate education	64,866	2.1%	0.14
Age	64,866	60.17	10.68
Male	64,866	45.95%	0.50
Married	64,662	92.7%	0.26
Hospitalized in the past year	63,375	9.1%	0.44
Insurance coverage	64,211	20.3%	0.40
Visited a healthcare facility in the past year	64,245	72.2%	0.45
Current work status	64,851	46.7%	0.50
Annual earnings (INR)	64,204	41,688.40	89,627.14

Note: includes adults 45 and older; excludes responses if reported by a proxy or had missing values for the education variable.

Source: authors' calculations based on LASI Wave I (weighted) (IIPS 2023).

Our sample, constructed of only those respondents who report their education levels, compares well with the overall LASI sample as reported by the International Institute for Population Sciences (IIPS) et al. (2020). According to the IIPS study, across the 45–59 age group, 13 per cent of men and 9 per cent of women have less than primary school education, 15 per cent of men and 11 per cent of women have completed primary schooling, 23 per cent of men and 13 per cent of women have completed secondary schooling, 7 per cent of men and 3 per cent of women have completed higher secondary schooling, while 9 per cent of men and 4 per cent of women have completed college or higher degrees. Across the 60+ age group the percentage distribution of educational attainment is similar, with 15 per cent of men and 9 per cent of women having less than primary school education, 15 per cent of men and 8 per cent of women having completed primary schooling, 21 per cent of men and 8 per cent of women having completed secondary schooling, 5 per cent of men and 1 per cent of women having completed higher secondary schooling, and 5 per cent of men and 1 per cent of women having completed college or higher degrees. Notably, 34 per cent of men and 61 per cent of women in the overall LASI sample (45 years and older) have no schooling. Our sample also compares favourably with the overall 45+ LASI sample in terms of the demographic characteristics of the respondents (see IIPS et al. 2020): 54 per cent of the sample consists of adults in the 45–59 age group, 46 per cent of adults aged 60 and over; 42 per cent of the sample aged 45 and over are men while 58% are women; and 68 per cent reside in rural areas while 32 per cent reside in urban areas.

3 Measures of reporting error

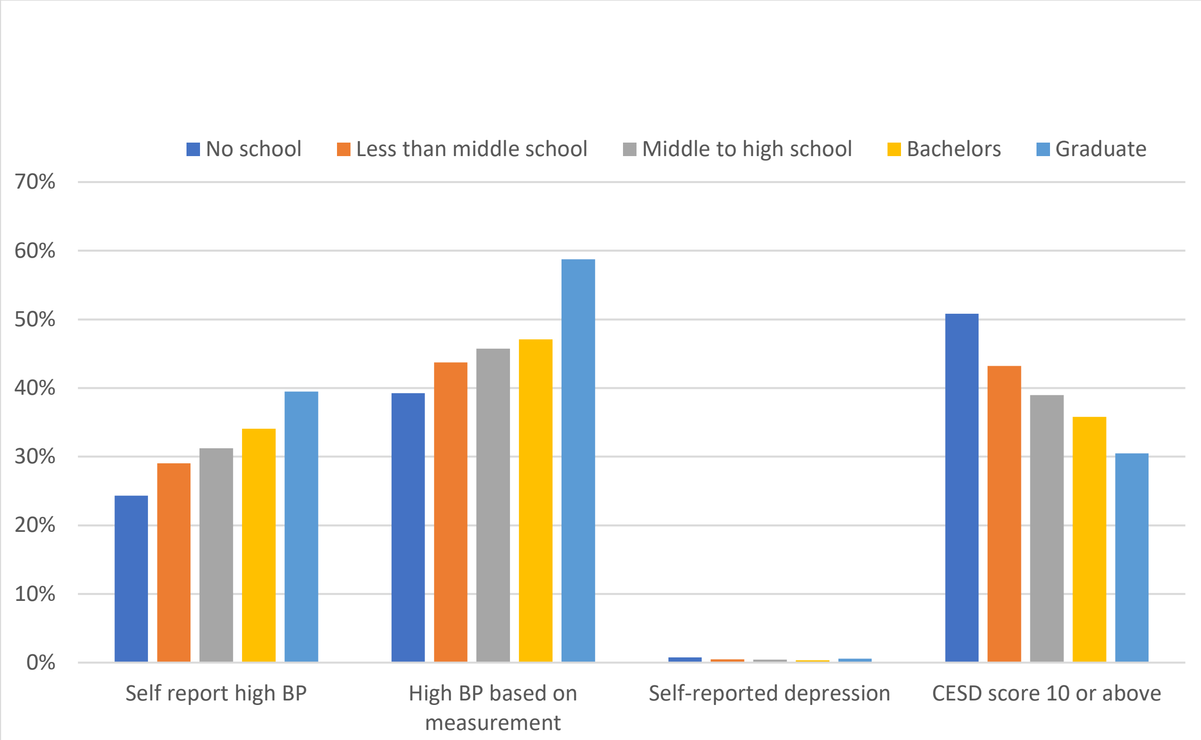
We construct three dichotomous measures of reporting error:

- **Accuracy**: If the self-reports of health condition or behaviours match objective measurements, then we code the respondent as providing an accurate report (equal to 1 if this is the case and 0 otherwise).[6] For example, if a respondent reports having high blood pressure and the objective measurements also indicate high blood pressure based on the readings, we code this response as an accurate self-report of high blood pressure. On the other hand, if the respondent mentions not having been diagnosed with high blood pressure but objective measurements show otherwise, we code the response as 0 for accurate reporting of high blood pressure. If a respondent is taking medication for high blood pressure, it is possible that they will have blood pressure measurements lower than the threshold for hypertension. We take this into account and code respondents who report taking medication for high blood pressure but have blood pressure measurements lower than the threshold as providing accurate self-reports of high blood pressure.
- **False negative**: When respondents report not having high blood pressure/depression but the blood pressure reading/CES-D score shows otherwise, the respondent is coded as 1 (0 otherwise). For example, when a respondent reports not having depression but the CES-D score shows otherwise, the respondent is coded as providing a false negative report of depression.
- **False positive**: When respondents report having high blood pressure/depression but the blood pressure readings/CES-D score indicate otherwise, the respondent is coded as 1 (0 otherwise). For example, when a respondent reports having high blood pressure, but the blood pressure readings indicate otherwise, then we code this as 1 for true positive reports.

Figure 1 shows the gap between self-reported and objectively measured incidence of high blood pressure and depression by education levels in our sample. A rough comparison shows that, on average, the gap between self-reported and objectively measured high blood pressure is 10–15 per cent for all education levels except those with graduate education, where the gap is around 20 per cent. However, for depression the gap between self-reported and objectively measured incidence by education levels is reversed. Self-reporting is low across the board, but the self-reported and objectively measured gap is greatest for those with no education, at around 50 per cent, and lowest for those with graduate education, at around 30 per cent.

[6] There are also 'true positive' (self-reported sick and correctly identified objectively) and 'true negative' (self-reported healthy and correctly identified objectively) measures which are identical to the accuracy measure.

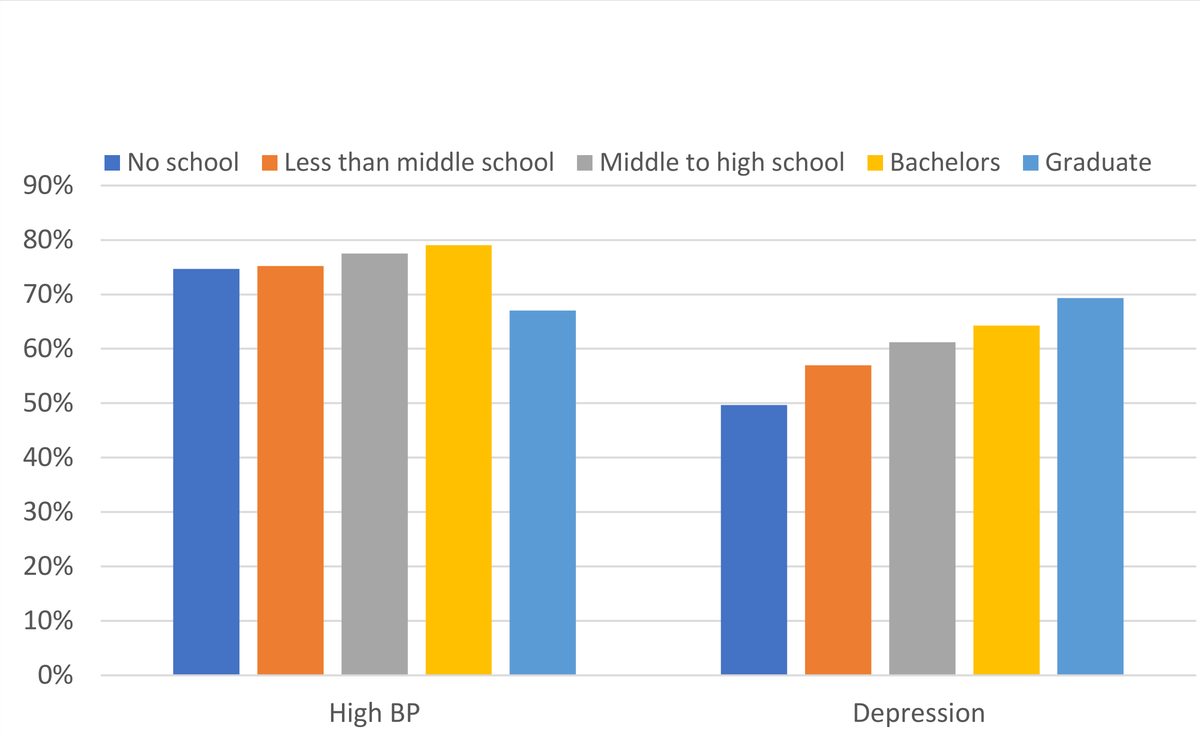
Figure 1: Self-reported and objectively measured blood pressure and depression by education levels (%)



Source: authors' illustration based on LASI Wave I.

Figure 2 shows the percentage of older adults who accurately predict high blood pressure and depression. Almost 75 per cent of older adults with no education report high blood pressure accurately; the figure is consistently this high or higher for all education levels, reaching 79 per cent for those with a bachelor's degree, except the highest—graduate education—for which it is around 66 per cent. We observe the opposite pattern with depression: the least educated group have the lowest percentage of accurate reporting (49 per cent) and the most educated group have the highest (69 per cent).

Figure 2: Accuracy of reporting of high blood pressure and mental depression by education levels (%)



Source: authors' illustration based on LASI Wave I.

4 Methods

We estimate probit regression models for the three types of reporting (accurate, false negative, and false positive) as the dependent variables. In the reported marginal effects in the tables that follow, we use individual-level weights to account for the multistage stratified area probability cluster sampling design with stage stratification and sample selection stages.[7] The main specification is as follows:

$$Reporting\ error = \alpha + \beta Education_i + \gamma X_i + \delta_{state} + \varepsilon_i$$

The main explanatory variable is level of education. We create an indicator variable for no schooling or education, less than middle school, middle or high school graduate, college graduate, and post-graduate degree. About half of the sample report no schooling or education (Table 1). The share of older adults reporting having a college or graduate degree is less than 10 per cent. The omitted reference category is middle or high school graduates. We also control for correct recall of words as a proxy for cognitive ability: after the interviewer read a set of ten words, respondents were asked to recall as many as possible in two minutes. On average, respondents recalled about five words correctly. Number of words recalled is negatively associated with no schooling or education but positively associated with increasing levels of education. X is a vector of individual characteristics including gender, age, marital status, first language, religion, and

[7] We exclude from the sample those who responded via a proxy rather than themselves, and those with missing values for education.

whether the respondent lives in a rural area. We also include state fixed effects (δ_{state}) in our baseline model.

5 Results

As Table 2 shows, older adults with no schooling or with less than middle school education report high blood pressure and depression less accurately than those with middle or high school education. This remains true (with larger point estimates) when we add different control variables such as earnings, healthcare utilization, and community-based identifiers like scheduled caste later in our robustness checks. Those with no schooling are 3.2 percentage points (4.2 per cent) less likely to accurately report high blood pressure and 5.3 percentage points (9.7 per cent) less likely to accurately report depression than those with middle or high school education. Similarly, those with less than middle school education are 2.1 percentage points (2.8 per cent) less likely to accurately report high blood pressure and 3.2 percentage points (5.9 per cent) less likely to accurately report depression than those with middle to high school education. Those with graduate education are 10.7 percentage points (19.6 per cent) more likely to accurately report depression in the survey than those with middle and high school education only.

A closer look at reporting error shows false negative reporting to be dominant (where respondents over-report their health, for example stating an absence of high blood pressure in the self-report where the objective measurement indicates otherwise). As with the results on accurate reporting, the point estimates are larger for those with no education, who are more likely to give a false negative report of high blood pressure than those with middle or high school education. For reports of high blood pressure, we observe a monotonic decrease in the degree of false negative reporting based on education levels. College-educated respondents are 5.9 percentage points (12.5 per cent) less likely to give a false negative report of high blood pressure than those with middle or high school education.

Table 2: Probit regression results for accurate, false negative, and false positive reporting of health conditions and behaviours (overall sample)

	Accurate reporting		False negative reporting		False positive reporting	
Variables	High blood pressure	Mental depression	High blood pressure	Mental depression	High blood pressure	Mental depression
No education	−0.032**	−0.053***	0.166***	0.000	−0.005	0.001
	(0.013)	(0.014)	(0.022)	(0.001)	(0.007)	(0.001)
Less than middle school	−0.021*	−0.032**	0.066***	0.001	0.015*	0.001
	(0.012)	(0.014)	(0.02)	(0.001)	(0.008)	(0.001)
Bachelor’s	0.014	0.012	−0.059*	0.001	0.011	0.000
	(0.02)	(0.027)	(0.032)	(0.002)	(0.015)	(0.002)
Graduate	−0.098	0.107*	0.039	−0.001	0.012	0.003
	(0.073)	(0.061)	(0.081)	(0.005)	(0.021)	(0.004)
Observations	59,116	63,241	26,392	24,354	32,647	36,035

Note: includes adults 45 and older; excluded if responses were reported by a proxy or had missing values for education variable; regressions control for number of words recalled correctly as well as indicator variables of age, gender, marital status, religion, first language, whether respondent lives in a rural area, and state fixed effects; *** p<0.01, ** p<0.05, * p<0.10; standard errors in parentheses

Source: authors’ construction based on LASI Wave I (weighted).

The results on false positive reporting of health suggest that for an asymptomatic condition like high blood pressure, those with less than middle school are more likely to under-report their health (for example saying that they have high blood pressure when objective measurements do not support this) than those with middle or high school education.

6 Robustness checks: income, healthcare utilization status, and different thresholds for depression

We further estimate the main baseline model controlling for additional variables related to income and healthcare utilization, such as (1) income levels including current work status and annual earnings (Table 3) and (2) respondent's health insurance coverage and healthcare utilization in the past year (whether respondents have any health insurance coverage, whether they visited a healthcare facility or were hospitalized in the past year) (Table 4). We control for these variables as they may influence the likelihood of discovering certain health conditions such as high blood pressure, which is asymptomatic.

Table 3: Probit regression results for accurate, false negative, and false positive reporting of health condition (overall sample), controlling for income-related variables (current work status and annual earnings)

	Accurate reporting		False negative reporting		False positive reporting	
Variables	High blood pressure	Mental depression	High blood pressure	Mental depression	High blood pressure	Mental depression (LPM)
No education	−0.030**	−0.054***	0.151***	0.000	−0.003	−0.032***
	(0.013)	(0.014)	(0.023)	(0.001)	(0.007)	(0.005)
Less than middle school	−0.020*	−0.032**	0.061***	0.001	0.015*	−0.016***
	(0.012)	(0.014)	(0.021)	(0.001)	(0.009)	(0.005)
Bachelor's	0.021	0.006	−0.058*	0.001	0.009	0.008
	(0.02)	(0.028)	(0.033)	(0.002)	(0.015)	(0.01)
Graduate	−0.077	0.088	0.022	−0.001	0.015	0.016
	(0.067)	(0.061)	(0.077)	(0.005)	(0.022)	(0.016)
Observations	58,583	62,665	26,205	24,208	32,301	33,950

Note: includes adults 45 and older; excluded if responses were reported by a proxy or had missing values for education variable; LPM indicates linear probability models (used for only the false positive reporting of depression, since convergence is unobtainable via the Probit specification); regressions control for number of words recalled correctly as well as indicator variables of age, gender, marital status, religion, first language, whether respondent lives in a rural area, state fixed effects, work status, and annual earnings in rupees; *** p<0.01, ** p<0.05, * p<0.10; standard errors in parentheses.

Source: authors' construction based on LASI Wave I (weighted).

Furthermore, we estimate the main specification with all the control variables used in the baseline model as well as whether or not the respondent belongs to a scheduled caste,[8] which may influence their health through eating habits or other health-seeking behaviours (Table 5).

The findings are remarkably consistent when additional income- and health-related variables are added. Adding scheduled caste in addition to the income and health variables also does not change our baseline findings (Table 5). The magnitude of the point estimates is smaller than in our baseline model (Table 2), but the patterns are mostly the same. Those with no schooling or who did not graduate from middle school are significantly less likely to accurately report both high blood pressure and depression. Similarly, the less educated are more likely to over-report high blood pressure, whereas those with a bachelor's degree are less likely to do so in each of our robustness checks (Table 3, 4, and 5).

Table 4: Probit regression results for accurate, false negative, and false positive reporting of health condition (overall sample), controlling for health-related variables (health insurance coverage and healthcare utilization)

	Accurate reporting		False negative reporting		False positive reporting	
Variables	High blood pressure	Mental depression	High blood pressure	Mental depression	High blood pressure	Mental depression (LPM)
No education	−0.027**	−0.058***	0.159***	0.001	−0.005	0.001
	(0.013)	(0.014)	(0.021)	(0.001)	(0.007)	(0.001)
Less than middle school	−0.021*	−0.033**	0.074***	0.002	0.014	0.001
	(0.012)	(0.014)	(0.019)	(0.001)	(0.009)	(0.001)
Bachelor's	0.02	0.002	−0.082**	0.001	0.013	0.000
	(0.02)	(0.028)	(0.034)	(0.002)	(0.016)	(0.002)
Graduate	−0.089	0.094	−0.007	−0.001	0.014	0.003
	(0.069)	(0.06)	(0.071)	(0.004)	(0.021)	(0.004)
Observations	58,109	62,231	25,977	23,600	32,050	35,598

Note: includes adults 45 and older; excluded if responses were reported by a proxy or had missing values for education variable; LPM indicates linear probability models; regressions control for number of words recalled correctly as well as indicator variables of age, gender, marital status, religion, first language, whether respondent lives in a rural area, state fixed effects, whether respondent has insurance coverage, hospitalization in past year, and whether respondent visited a healthcare facility in past year; *** p<0.01, ** p<0.05, * p<0.10; standard errors in parentheses.

Source: authors' construction based on LASI Wave I (weighted).

[8] We control for whether the respondent belongs to a scheduled caste because of previous literature suggesting significant association between caste membership and health behaviours and mortality rates (Coelho et al. 2016). Our results remain the same with a larger point estimate without this control.

Table 5: Probit regression results for accurate, false negative, and false positive reporting of health condition (overall sample), controlling for income- and health-related variables and scheduled caste

	Accurate reporting		False negative reporting		False positive reporting	
Variables	High blood pressure	Mental depression	High blood pressure	Mental depression	High blood pressure	Mental depression (LPM)
No education	−0.025**	−0.056***	0.143***	0.000	−0.003	0.001
	(0.013)	(0.014)	(0.022)	(0.001)	(0.007)	(0.001)
Less than middle school	−0.019	−0.030**	0.070***	0.001	0.014	0.001
	(0.012)	(0.014)	(0.02)	(0.001)	(0.009)	(0.001)
Bachelor's	0.028	−0.001	−0.078**	0.001	0.01	0.000
	(0.02)	(0.028)	(0.035)	(0.002)	(0.016)	(0.002)
Graduate	−0.067	0.08	−0.022	−0.001	0.018	0.004
	(0.063)	(0.059)	(0.068)	(0.004)	(0.022)	(0.004)
Observations	57,424	61,479	25,718	23,354	31,624	35,147

Note: includes adults 45 and older; excluded if responses were reported by a proxy or had missing values for education variable; LPM indicates linear probability models; regressions control for number of words recalled correctly as well as indicator variables of age, gender, marital status, religion, first language, scheduled caste, whether respondent lives in a rural area, state fixed effects, annual earnings in rupees, current work status, whether respondent has insurance coverage, hospitalization in the past year, and whether respondent visited a healthcare facility in past year; *** p<0.01, ** p<0.05, * p<0.10; standard errors in parentheses.

Source: authors' construction based on LASI Wave I (weighted).

Furthermore, we conduct additional sensitivity analysis for depression. We use the threshold of score 10 from the CES-D ten-item questionnaires to define someone as suffering from depression based on objective measurement. To check whether our results are sensitive to the threshold we use, we estimate the baseline model with additional controls (income- and health-related variables as in Tables 3 and 4) for depression using four different thresholds (the 25th percentile of the sample, which is CES-D score of 7 or higher; the 50th, score of 9 or higher; the 75th, score of 12 or higher; and the 90th, score of 15 or higher). We re-create the three outcome variables for reporting error for depression separately based on these different thresholds. The main results remain robust, as shown in Table 6: the less educated are less likely to accurately report depression at all thresholds than those who have middle or high school education. Even at the 90th percentile this remains true: the better educated are more likely to accurately report depression than those with middle or high school education.

Table 6: Probit regression results for accurate, false negative, and false positive reporting of depression using different CES-D thresholds, controlling for income- and health-related variables

Accurate reporting				
	25th percentile	50th	75th	90th
No education	−0.063***	−0.060***	−0.064***	−0.030***
	(0.011)	(0.013)	(0.011)	(0.007)
Less than middle school	−0.021**	−0.030**	−0.048***	−0.019**
	(0.010)	(0.013)	(0.011)	(0.007)
Bachelor’s	0.026	0.009	0.015	0.021*
	(0.020)	(0.027)	(0.019)	(0.011)
Graduate	0.118**	0.116**	0.075**	0.038***
	(0.057)	(0.058)	(0.029)	(0.015)
Observations	61,648	61,668	61,663	61,655
False negative reporting				
	25th percentile	50th	75th	90th
No education	−0.000	−0.001	0.001	0.003
	(0.001)	(0.001)	(0.002)	(0.003)
Less than middle school	0.001	0.000	0.003**	0.004**
	(0.001)	(0.001)	(0.001)	(0.002)
Bachelor’s	−0.000	0.000	0.002	
	(0.002)	(0.002)	(0.002)	
Graduate	−0.001	−0.002	−0.001	−0.033
	(0.003)	(0.004)	(0.006)	(0.037)
Observations	43,970	33,046	14,360	5,627
False positive reporting				
	25th percentile †	50th	75th	90th
No education	−0.000	−0.001	0.001	0.001
	(0.001)	(0.000)	(0.000)	(0.000)
Less than middle school	0.001	0.000	0.001	0.001
	(0.001)	(0.000)	(0.001)	(0.001)
Bachelor’s	−0.001	0.000	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Graduate	0.005	0.003	0.003	0.002
	(0.005)	(0.003)	(0.003)	(0.003)

Observations	13,610	13,494	33,411	44,693

Note: includes adults 45 and older; CES-D score thresholds for 25th, 50th, 75th, and 90th percentiles are 7, 9, 12, and 15 respectively; excluded if responses were reported by a proxy or had missing values for education variable; regressions control for number of words recalled correctly as well as indicator variables of age, gender, marital status, religion, first language, whether respondent lives in a rural area, state fixed effects, annual earnings in rupees, current work status, whether respondent has insurance coverage, hospitalization in past year, and whether respondent visited a healthcare facility in past year; *** p<0.01, ** p<0.05, * p<0.10; standard errors in parentheses; † Linear probability model results reported.

Source: authors' construction based on LASI Wave I (weighted).

7 Subgroup analysis: gender, income levels, family history of high blood pressure, and accurate reporting of physical activity

In addition to different robustness checks, we conduct several subgroup analyses based on categories such as gender, poverty status, and whether the respondent's family has a history of high blood pressure. In Tables 7a and 7b we show the main results separately by gender, as studies have found a significant gender gap in health among adults in India (Lee et al. 2015). Although the main results remain the same, this is primarily driven by men, as shown in Table 7a. Among women (Table 7b), although the direction of the coefficient is the same for results on accurate, false negative, and false positive reporting of health, many coefficients are no longer significant even with a larger sample size than that of men. Women with no schooling are significantly more likely to over-report high blood pressure than those with middle or high school education.

Table 7a: Probit regression results for accurate, false negative, and false positive reporting of health condition, gender = male, controlling for income- and health-related variables

	Accurate reporting		False negative reporting		False positive reporting	
Variables	High blood pressure	Mental depression	High blood pressure	Mental depression	High blood pressure	Mental depression (LPM)
No education	−0.023*	−0.074***	0.181***	0.001	−0.016*	−0.052
	(0.013)	(0.014)	(0.022)	(0.04)	(0.01)	(0.013)
Less than middle school	−0.029**	−0.055***	0.104***	0.001	0.006	-0.031
	(0.012)	(0.013)	(0.02)	(0.053)	(0.011)	(0.013)
Bachelor's	0.028	−0.029	−0.070**	−0.001	0.000	0.028
	(0.022)	(0.028)	(0.034)	(0.084)	(0.018)	(0.026)
Graduate	−0.082	0.146***	−0.021	−0.002	0.01	0.009
	(0.065)	(0.038)	(0.057)	(0.12)	(0.025)	(0.034)
Observations	26,654	28,594	11,678	9,044	14,989	12,811

Note: includes adults 45 and older; excluded if responses were reported by a proxy or had missing values for education variable; LPM indicates linear probability models; regressions control for number of words recalled correctly as well as indicator variables of age, gender, marital status, religion, first language, whether respondent lives in a rural area, state fixed effects, current employment status, annual earnings, whether respondent has insurance coverage, hospitalization in past year, and whether respondent visited a healthcare facility in past year; *** p<0.01, ** p<0.05, * p<0.10; standard errors in parentheses.

Source: authors' construction based on LASI Wave I (weighted).

Table 7b: Probit regression results for accurate, false negative, and false positive reporting of health condition, gender = female, controlling for income- and health-related variables

	Accurate reporting		False negative reporting		False positive reporting	
Variables	High blood pressure	Mental depression	High blood pressure	Mental depression	High blood pressure	Mental depression (LPM)
No education	−0.017	−0.034	0.090***	−0.004	0.01	0.000
	(0.024)	(0.025)	(0.035)	(0.004)	(0.011)	(0.002)
Less than middle school	−0.005	0.000	0.007	0.000	0.024**	0.001
	(0.024)	(0.025)	(0.034)	(0.004)	(0.011)	(0.002)
Bachelor’s	0.034	0.06	−0.057	0.006	0.034	0.001
	(0.041)	(0.055)	(0.073)	(0.004)	(0.031)	(0.003)
Graduate	−0.028	−0.064	0.02	−0.007	0.044	−0.003*
	(0.048)	(0.098)	(0.078)	(0.007)	(0.045)	(0.002)
Observations	30,904	33,061	14,104	15,024	16,813	18,047

Note: includes adults 45 and older; excluded if responses were reported by a proxy or had missing values for education variable; LPM indicates linear probability models; regressions control for number of words recalled correctly as well as indicator variables of age, gender, marital status, religion, first language, whether respondent lives in a rural area, state fixed effects, current employment status, annual earnings, whether respondent has insurance coverage, hospitalization in past year, and whether respondent visited a healthcare facility in past year; *** p<0.01, ** p<0.05, * p<0.10; standard errors in parentheses.

Source: authors’ construction based on LASI Wave I (weighted).

As noted in other studies that explore the income–health gradient in India, there is a significant difference in health conditions and behaviours by income levels (Onur and Velamuri 2018; Vellakkal et al. 2013). We therefore also estimate the models separately by income levels. We use the 25th percentile of the annual earnings distribution in our sample (only among those who have any income at all), which is INR36,000 (about US$1.27 per day).

When we run the regressions separately by income threshold of INR36,000 annual earnings, the main results hold true primarily for the group above the threshold (Table 8b) but not necessarily for those below (Table 8a). Among those who earn less than INR36,000 a year, there are no significant differences in reporting error by education levels for accurate reporting of high blood pressure.

Table 8a: Probit regression results for accurate, false negative, and false positive reporting of health condition, poverty (earnings < INR36,000/year), controlling for work status and health-related variables

	Accurate reporting		False negative reporting		False positive reporting	
Variables	High blood pressure	Mental depression	High blood pressure	Mental depression	High blood pressure	Mental depression (LPM)
No education	−0.015	−0.060***	0.120***	0.000	−0.004	0.001
	(0.016)	(0.018)	(0.025)	(0.002)	(0.009)	(0.001)
Less than middle school	−0.006	−0.016	0.034	0.001	0.012	0.001
	(0.016)	(0.018)	(0.023)	(0.002)	(0.009)	(0.002)
Bachelor's	0.017	0.035	-0.038	0.003**	0.027	−0.001
	(0.027)	(0.038)	(0.042)	(0.001)	(0.023)	(0.002)
Graduate	0.012	0.118***	−0.062	0.003	0.000	−0.001
	(0.032)	(0.034)	(0.049)	(0.002)	(0.031)	(0.003)
Observations	37,451	39,890	18,000	15,690	19,460	21,926

Note: includes adults 45 and older; excluded if responses were reported by a proxy or had missing values for education variable; LPM indicates linear probability models; regressions control for number of words recalled correctly as well as indicator variables of age, gender, marital status, religion, first language, whether respondent lives in a rural area, state fixed effects, current employment status, whether respondent has insurance coverage, hospitalization in past year, and whether respondent visited a healthcare facility in past year; *** p<0.01, ** p<0.05, * p<0.10; standard errors in parentheses.

Source: authors' construction based on LASI Wave I (weighted).

It is possible that a person might be more aware of their health condition if their family members have that particular condition. This could lead some to get health checks more often. The data allow us to identify individuals who have any family members with high blood pressure. We thus estimate the main model with additional controls on income- and health-related variables (as in Tables 3 and 4) among those who report having any family members with high blood pressure. As shown in Table 9, the result on false negative reporting remains similar to our other results, i.e., the less educated are more likely to give a false negative report of high blood pressure. The results for accurate reporting, however, show a slightly weaker relationship by education levels than in our main findings.

Table 8b: Probit regression results for accurate, false negative, and false positive reporting of health condition, poverty (earnings > INR36,000/year), controlling for work status and health-related variables

	Accurate reporting		False negative reporting		False positive reporting	
Variables	High blood pressure	Mental depression	High blood pressure	Mental depression	High blood pressure	Mental depression (LPM)
No education	−0.036**	−0.055***	0.181***	0.001	−0.003	0
	(0.016)	(0.018)	(0.029)	(0.002)	(0.012)	(0.001)
Less than middle school	−0.038**	−0.060***	0.131***	0.002	0.015	0.001
	(0.015)	(0.017)	(0.025)	(0.002)	(0.015)	(0.001)
Bachelor's	0.033	−0.053	−0.100**	−0.005	−0.004	0.002
	(0.028)	(0.037)	(0.045)	(0.005)	(0.019)	(0.002)
Graduate	−0.147	0.066	0.035	−0.01	0.027	0.008
	(0.091)	(0.078)	(0.078)	(0.009)	(0.031)	(0.006)
Observations	20,118	21,751	7,784	8,475	12,342	13,301

Note: includes adults 45+; excluded if responses were reported by a proxy or had missing values for education variable; regressions control for number of words recalled correctly and indicator variables of age, gender, marital status, religion, first language, whether respondent lives in a rural area, state fixed effects, current employment status, whether respondent has insurance coverage, hospitalization in past year, and whether respondent visited a healthcare facility in past year; *** p<0.01, ** p<0.05, * p<0.10; standard errors in parentheses.

Source: authors' construction based on LASI Wave I (weighted).

Table 9: Probit regression results for accurate, false negative, and false positive reporting of high blood pressure among those who report family history of high blood pressure, controlling for income- and health-related variables

	High blood pressure		
	Accurate reporting	False negative reporting	False positive reporting
No education	−0.007 (−0.009)	0.085*** (−0.017)	−0.017*** (−0.006)
Less than middle school	−0.017* (−0.01)	0.042** (−0.02)	0.017*** (−0.007)
Bachelor's	−0.002 (−0.023)	−0.026 (−0.036)	0.026 (−0.019)
Graduate	−0.009 (−0.031)	−0.064 (−0.06)	0.016 (−0.023)
Observations	44,981	18,925	25,989

Note: includes adults 45+; excluded if responses were reported by a proxy or had missing values for education variable; regressions control for number of words recalled correctly and indicator variables of age, gender, marital status, religion, first language, whether respondent lives in a rural area, state fixed effects, current employment status, whether respondent has insurance coverage, hospitalization in past year, and whether respondent visited a healthcare facility in past year; *** p<0.01, ** p<0.05, * p<0.10; standard errors in parentheses.

Source: authors' construction based on LASI Wave I (weighted).

Unlike asymptomatic conditions such as high blood pressure, whether one engages in physical activity is a conscious behaviour that does not require any diagnosis. We thus estimate the model for the group of respondents that accurately report their physical activity in the past month. There are several questions related to physical activity in the data, such as whether the respondent has engaged in vigorous or moderate physical activity in the past month and whether they walked the previous day. We use the question on how often they engaged in sports or vigorous physical activities in the past month.[9] We examine different variables for self-reports of physical activity in the analysis but mainly focus on whether the respondent reports having engaged in physical activity in the past month as a general proxy for physical activity.[10]

For objective measurement on physical activity, we use the data on timed walking, measured among older adults who do not have a health condition that limits walking ability. Respondents walk along a 4-metre path in a non-carpeted area (with a measuring tape indicating the walking space) at their usual pace of walking. The measurement is undertaken twice, with an interviewer present. We calculate the average time of walking measured in seconds and divide it by four to calculate the walking speed (seconds per metre). We proxy for objective measurement of physical activity (specifically walking) based on the walking speed from the timed walk (1 m/s or greater) (Middleton et al. 2015). The threshold of 1 m/s was chosen based on associated outcomes such as being independent and less likely to be hospitalized or have adverse events among older adults (Middleton et al. 2015).

Table 10 shows the regression results for accurate, false negative, and false positive reporting of health condition among those who accurately report their physical activity engagement in the past month. This sample includes older adults who report not engaging in physical activity in the past month and who do not have a walking speed above the threshold, and the opposite case (those who exercised in the past month and had a walking speed above the threshold). It is assumed that this will be the subgroup that correctly recalls their exercise level and accurately reports it in the data.

Among this sample, accurate reporting of high blood pressure does not seem to differ by education levels. For outcomes on depression, the main findings remain the same: those with no schooling or less than middle school education are less likely to accurately report than those with middle or high school education. False negative reporting of depression is significant for most of the groups, but the coefficient for false negative reporting is largest for those with graduate degrees and above. However, false negative reporting did not significantly differ by education levels among the group that did not accurately report physical activity.[11] This might indicate that among those respondents who accurately recall their recent exercise level and are more likely to accurately report conscious behaviour, social desirability bias[12] might be driving the results on false negative reporting of depression.

---

[9] The question asks 'We would like to know the type and amount of physical activity involved in your daily life. How often do you take part in sports or vigorous activities, such as running or jogging, swimming, going to a health centre or gym, cycling, or digging with a spade or shovel, heavy lifting, chopping, farm work, fast bicycling, cycling with loads: every day, more than once a week, once a week, one to three times a month, or hardly ever or never?'

[10] There were specific questions on walking but due to the skip pattern in the survey, the variables had lots of missing values. Results are available upon request.

[11] Results are available upon request.

[12] Social desirability bias is defined as a someone concealing their true opinions in the belief that they will be more favourably looked upon by others if they do so. Given the stigma towards mental illness (Venkatesh et al. 2015), this may well be the case in India.

Table 10: Probit regression results for accurate, false negative, and false positive reporting of health condition among those who accurately report physical activity in the past month, controlling for income- and health-related variables

	Accurate reporting		False negative reporting		False positive reporting	
Variables	High blood pressure	Mental depression	High blood pressure	Mental depression (LPM)	High blood pressure	Mental depression (LPM)
No education	−0.021	−0.073***	0.121***	−0.003	−0.002	−0.002
	(0.017)	(0.019)	(0.027)	(0.003)	(0.010)	(0.011)
Less than middle school	−0.006	−0.037**	0.037	0.006**	0.009	0.009
	(0.016)	(0.019)	(0.025)	(0.003)	(0.011)	(0.012)
Bachelor's	0.013	−0.003	−0.054	0.007**	−0.005	−0.003
	(0.028)	(0.035)	(0.044)	(0.003)	(0.021)	(0.022)
Graduate	0.007	0.061*	−0.089*	0.009**	0.034	0.026
	(0.030)	(0.036)	(0.048)	(0.004)	(0.034)	(0.032)
Observations	33,886	33,749	16,473	14,796	17,664	17,756

Note: includes adults 45 and older; excluded if responses were reported by a proxy or had missing values for education variable; LPM indicates linear probability models; sample includes only those who accurately reported physical activity (whether or not they engaged in general physical activity (sports or vigorous activities)) in the past month; regressions control for number of words recalled correctly as well as indicator variables of age, gender, marital status, religion, first language, whether respondent lives in a rural area or not, state fixed effects, current employment status, annual earnings, whether respondent has insurance coverage, hospitalization in past year, and whether respondent visited a healthcare facility in past year; *** p<0.01, ** p<0.05, * p<0.10; standard errors in parentheses';

Source: authors' construction based on LASI Wave I (weighted).

8 Conclusion

This is the first study to analyse the health–education gradient for older adults in India based on differential health reporting error. This is also the first study to analyse the accuracy of self-diagnosis for depression in a developing country. We find evidence suggesting that older adults with no schooling and less than middle school education report high blood pressure and depression less accurately than those with middle or high school education. This pattern is consistent across different controls, such as earnings, healthcare utilization, and community-based identifiers like scheduled caste status. We find false negative reporting for high blood pressure to be dominant (where someone over-reports their health, for example stating an absence of high blood pressure in the self-report but where the objective measurement indicates otherwise) for those with no education.

Sensitivity analysis for depression shows our baseline results based on the ten-item CES-D questionnaire to be robust, with the less educated less likely to accurately report depression at all thresholds than those who have middle or high school education. Further, our results show that men with no schooling and less than middle school education report high blood pressure and depression less accurately: women with no schooling are significantly more likely to over-report high blood pressure than those with middle or high school education. Surprisingly, we also find that our finding that older adults with no schooling and less than middle school education are more likely to be less accurate in their reporting for high blood pressure and depression holds true only for those whose income is above the poverty line of roughly US $1.27 per day.

References

Andresen, E.M., J.A. Malmgren, W.B. Carter, and D.L. Patrick (1994). 'Screening for Depression in Well Older Adults: Evaluation of a Short Form of the CES-D'. *American Journal of Preventive Medicine*, 10(2): 77–84. https://doi.org/10.1016/S0749-3797(18)30622-6

Bernadine, M. (2021). 'Understanding India's Mental Health Crisis'. *Ideas for India*, 6 April. Available at: www.ideasforindia.in/topics/human-development/understanding-india-s-mental-health-crisis.html (accessed 27 January 2023).

Bhan, N., L. McDougal, A. Singh, Y. Atmavilas, and A. Raj (2020). 'Access to Women Physicians and Uptake of Reproductive, Maternal and Child Health Services in India. *EClinical Medicine*, 20, 100309. https://doi.org/10.1016/j.eclinm.2020.100309

Bloom, D., T.V. Sekher, and J. Lee (2021). 'Longitudinal Aging Study in India (LASI): New Data Resources for Addressing Aging in India'. *Nature Aging*, 1: 1070–72. https://doi.org/10.1038/s43587-021-00155-y

Choi, A., and J. Cawley (2018). 'Health Disparities across Education: The Role of Differential Reporting Error'. *Health Economics*, 27(3): e1–e29. https://doi.org/10.1002/hec.3609

Coelho, K., C. Belden, C. Belden, and C. Belden (2016). 'A Systematic Review of the Literature on the Relationship between Caste Membership and Health-Related Risk Factors in India'. *International Journal of Medicine and Public Health*, 6(2): 61–68. https://doi.org/10.5530/ijmedph.2016.2.3

Cramm, J.M., L. Bornscheuer, A. Selivanova, and J. Lee (2015). 'The Health of India's Elderly Population: A Comparative Assessment using Subjective and Objective Health Outcomes'. *Journal of Population Ageing*, 8(4): 245–59. https://doi.org/10.1007/s12062-015-9122-2

Cutler, D.M., and A. Lleras-Muney (2010). 'Understanding Differences in Health Behaviors by Education'. *Journal of Health Economics*, 29(1): 1–28. https://doi.org/10.1016/J.JHEALECO.2009.10.003

Dandona, L., et. al. (2020). 'The Burden of Mental Disorders across the States of India: The Global Burden of Disease Study 1990–2017'. *Lancet Psychiatry*, 7(2): 148–61. https://doi.org/10.1016/S2215-0366(19)30475-4

Dey, S., D. Nambiar, J.K. Lakshmi, K. Sheikh, and K. Srinath Reddy (2012). 'Health of the Elderly in India: Challenges of Access and Affordability'. In J.P. Smith and M. Majmundar (eds), *Aging in Asia: Findings From New and Emerging Data Initiatives*. Washington, DC: National Academies Press.

Geldsetzer P, J. Manne-Goehler, M. Theilmann, J.I. Davies, A. Awasthi, S. Vollmer, L.M. Jaacks, T. Bärnighausen, and R. Atun (2018). 'Diabetes and Hypertension in India: A Nationally Representative Study of 1.3 Million Adults'. *JAMA Internal Medicine*, 178(3): 363–72. DOI: 10.1001/jamainternmed.2017.8094

Gellis, Z.D. (2010). 'Assessment of a Brief CES-D Measure for Depression in Homebound Medically Ill Older Adults'. *Journal of Gerontological Social Work*, 53(4): 289. https://doi.org/10.1080/01634371003741417

Grossman, M. (2008). 'The Relationship between Health and Schooling on JSTOR'. *Eastern Economic Journal*, 34(3), 281–292. https://www.jstor.org/stable/20642426?seq=1

Huffman, M.D., D. Prabhakaran, C. Osmond, C.H.D. Fall, N. Tandon, R. Lakshmy, S. Ramji, A. Khalil, T. Gera, P. Prabhakaran, Dey Biswas, K. Srinath Reddy, S.K. Bhargava, and H.S. Sachdev (2011). 'Incidence of Cardiovascular Risk Factors in an Indian Urban Cohort: Results from the New Delhi

Birth Cohort'. Journal of the American College of Cardiology, 57(17): 1765–74. https://doi.org/10.1016/j.jacc.2010.09.083

International Institute for Population Sciences (IIPS) (2023). 'Longitudinal Ageing Study in India (LASI): Datasets)'. Available at: www.iipsindia.ac.in/content/LASI-data (accessed 27 January 2023).

IIPS, National Programme for Health Care of Elderly (NPHCE), Indian Ministry of Health and Family Welfare (MoHFW), Harvard T.H. Chan School of Public Health (HSPH), and the University of Southern California (USC) (2020). *Longitudinal Aging Study in India (LASI) Wave 1, 2017–18, Executive Summary*. Mumbai: International Institute for Population Sciences.

Iqbal, S., R. Barolia, P. Petrucka, L. Ladak, R. Rehmani, and A. Kabir (2022). 'Smoking Cessation Interventions in South Asian Region: A Systematic Scoping Review'. BMC Public Health, 22: 1096. https://doi.org/10.1186/s12889-022-13443-y

Johnston, D.W., C. Propper, and M.A. Shields (2009). 'Comparing Subjective and Objective Measures of Health: Evidence from Hypertension for the Income/Health Gradient'. *Journal of Health Economics*, *28*(3): 540–52. https://doi.org/10.1016/j.jhealeco.2009.02.010

Lee, J., M.E. McGovern, D.E. Bloom, P. Arokiasamy, A. Risbud, J. O'Brien, V. Kale, and P. Hu (2015). 'Education, Gender, and State-Level Disparities in the Health of Older Indians: Evidence from Biomarker Data'. *Economics & Human Biology*, 19: 145–56. https://doi.org/10.1016/J.EHB.2015.09.003

Lee, J., J. Wilkens, E. Meijer, T.V. Sekher, D. Bloom, and P. Hu (2022). 'Hypertension Awareness, Treatment, and Control and Their Association with Healthcare Access in the Middle-Aged and Older Indian Population: A Nationwide Cohort Study'. *PLOS Medicine*, 4 January. https://doi.org/10.1371/journal.pmed.1003855

Lowes, S., and E. Montero (2021). 'The Legacy of Colonial Medicine in Central Africa'. *American Economic Review*, 111(4): 1284–314. https://doi.org/10.1257/aer.20180284

Middleton, A., S.L. Fritz, and M. Lusardi (2015). 'Walking Speed: The Functional Vital Sign'. *Journal of Aging and Physical Activity*, 23(2): 314. https://doi.org/10.1123/JAPA.2013-0236

Najnin, N., K. Leder, A. Forbes, L. Unicomb, P. Winch, P. Ram, F. Nizame, S. Arman, F. Begum, S. Biswas, A. Cravioto. and S. Luby (2019). 'Impact of a Large-Scale Handwashing Intervention on Reported Respiratory Illness: Findings from a Cluster-Randomized Controlled Trial'. *American Journal of Tropical Medicine and Hygiene*, 100(3): 742–49. DOI: 10.4269/ajtmh.18-0644

Onur, I., and M. Velamuri (2018). 'The Gap between Self-Reported and Objective Measures of Disease Status in India'. *PLOS ONE*, 13(8): e0202786. https://doi.org/10.1371/JOURNAL.PONE.0202786

Pednekar, M.S., R. Gupta, and P.C. Gupta (2011). 'Illiteracy, Low Educational Status, and Cardiovascular Mortality in India'. *BMC Public Health*, 11: article 567. http://www.biomedcentral.com/1471-2458/11/567

Perianayagam, A., D. Bloom, J. Lee, S. Parasuraman, T.V. Sekher, S. Mohanty, A. Chattopadhyay, D. Govil, S. Pedgaonkar, S. Gupta, A. Agarwal, A. Posture, A. Weerman, and S. Pramanik (2022). 'Cohort Profile: The Longitudinal Aging Study in India (LASI)'. *International Journal of Epidemiology*, 51(4): e167–e176. https://doi.org/10.1093/ije/dyab266

Radloff, L.S. (1977). 'The CES-D Scale: A Self-Report Depression Scale for Research in the General Population'. *Applied Psychological* Measurement, 1: 385–401. https://doi.org/10.1177/014662167700100306

Raghupathi, V., and W. Raghupathi (2020). 'The Influence of Education on Health: An Empirical Assessment of OECD Countries for the Period 1995–2015'. *Archives of Public Health*, 78: article 20. https://doi.org/10.1186/s13690-020-00402-5

Van der Heide, I., J. Wang, M. Droomers, P. Spreeuwenberg, J. Rademakers, and E. Uiters (2013). 'The Relationship between Health, Education, and Health Literacy: Results from the Dutch Adult Literacy and Life Skills Survey'. Journal of Health Communication, 18 (Sup 1): 172–84. https://doi.org/10.1080/10810730.2013.825668

Vellakkal, S., S.V. Subramanian, C. Millett, S. Basu, S., D. Stuckler, and S. Ebrahim (2013). 'Socioeconomic Inequalities in Non-Communicable Diseases Prevalence in India: Disparities between Self-Reported Diagnoses and Standardized Measures'. *PLoS ONE*, *8(7*). https://doi.org/10.1371/journal.pone.0068219

Venkatesh, B., T. Andrews, S. Mayya, M. Singh and S. Parsekar (2015). 'Perception of Stigma towards Mental Illness in South India'. *Journal of Family Medicine and Primary Care*, 4(3): 449–53. https://Doi.Org/10.4103/2249-4863.161352

World Bank (2019). World Health Organization Global Health Expenditure database, World Bank, Washington, DC. https://data.worldbank.org/indicator/SH.XPD.CHEX.GD.ZS