Comparison of socioeconomic differences in health using objective and self-reported measures of the same condition: Evidence from the Health Survey for England 2011-14

Scholes S, Morris S, Conolly A, Mindell JS

P64

Self-report data has important benefits (costs, speed) but also has limitations. **Reporting error** (a very broad term) encompasses:

- Inaccurate reporting (e.g. underestimate weight, overestimate height); and
- Undiagnosed disease (when persons *actually* have disease).

Objective data can be used with self-report data to: (1) estimate "total" prevalence (i.e. including both diagnosed and undiagnosed cases); and (2) examine systematic variations in reporting error. For example, among persons actually with disease (established using physical measurements), persons from lower socioeconomic groups may be more likely to report being undiagnosed than those in higher socioeconomic groups.[1,2] Relying only on self-report data could therefore lead to inaccurate estimates of the associations between socioeconomic status and health.

Objective (OBJ) and self-report (SR) data collected from the same participants in the Health Survey for England (HSE) was used to estimate differences in undiagnosed hypertension across income groups. We illustrate our findings using data on hypertension for men in the HSE 2011-14 (N=7,064).

Methods

Table 1. *Two definitions of hypertension*

	Self-report, diagnosed disease	Total prevalence (OBJ and SR data)
Hypertension	Doctor-diagnosed high blood pressure (BP)	SBP ≥140mmHg or DBP≥90mmHg; or currently on medication* to lower BP

** Medication use included as part of the definition of total prevalence as using the BP data alone would only identify persons whose hypertension was uncontrolled .*

The standard definition of being undiagnosed is the probability an individual self-reports not having the disease, *conditional* on actually having the disease (**Table 2**).[1]

Table 2. *Standard definition of being undiagnosed*

Self-report	Has Disease (Total Prevalence)	
	1: Yes	**0: No**
1: Yes	p11	p10
0: No	p01	p00

Probability of being undiagnosed :

$$\frac{p01}{(p01+p11)}$$

The key parameter for our study was the difference between income groups in the probability of being undiagnosed. That is, *conditional* on being hypertensive (according to BP and medication data), was there evidence to suggest that men in the highest-income group were more or less likely to be classed as undiagnosed than men in the lowest-income group. We estimated this hypothesis using a bivariate probit model (see below).

Descriptive analysis

Figure 1. *Hypertension by income tertiles*

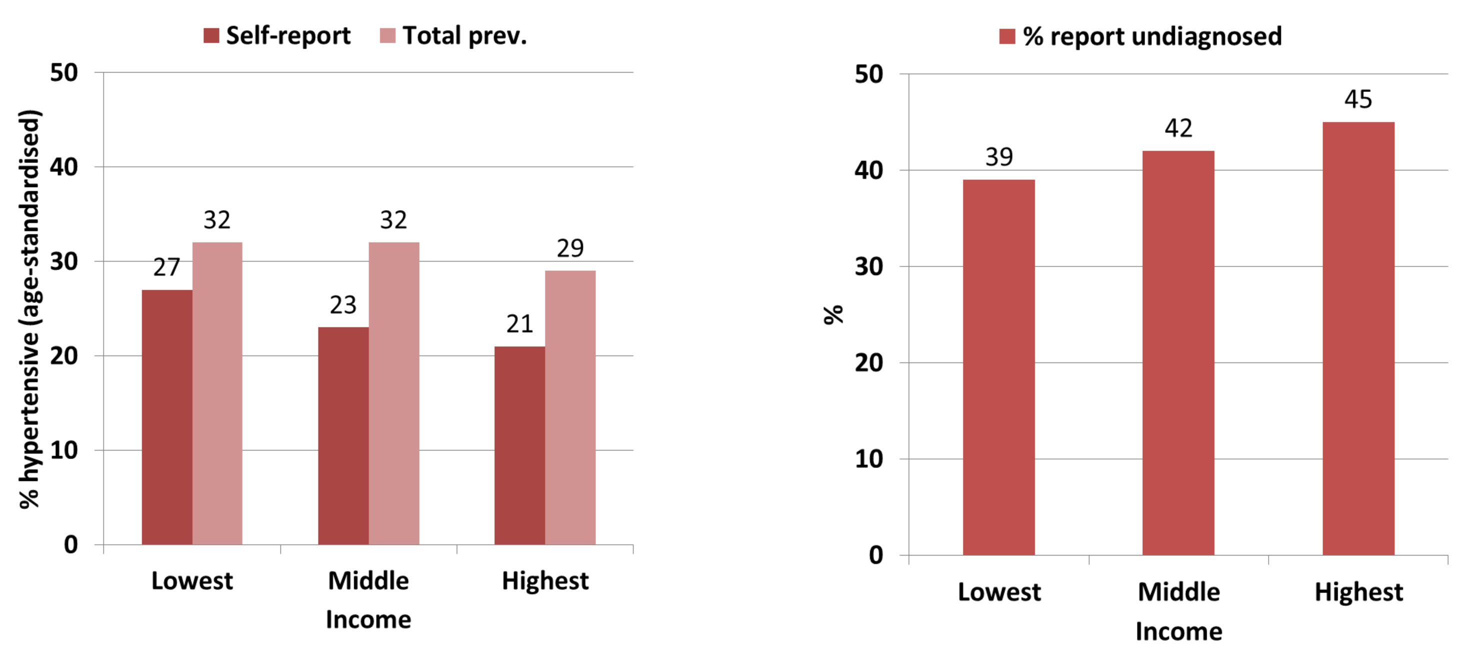

Figure 1: % hypertension varied by income tertile using self-report data (27% lowest; 21% highest); but less so using both objective and self-report data to estimate total prevalence (32%; 29%) (**left panel**). Among persons classed as hypertensive (according to BP and medication data), % who reported being undiagnosed was lowest in the lowest-income group (39%; 45%) (**right panel**).

Statistical modelling: bivariate probit model

We used a **joint** model for these two interrelated binary outcomes.[1,3] We examined the associations between income and hypertension adjusting for: age; smoking status; marital status; and BMI category.

Model coefficients can be interpreted as **A**verage **M**arginal **E**ffects (AMEs) (**Figure 2**). Results showed a significant difference between the highest and lowest income groups in the probability of reporting diagnosed hypertension (*P*=0.003) but not for being classed as hypertensive according to BP and medication data (*P*=0.776) (**Figure 2**).

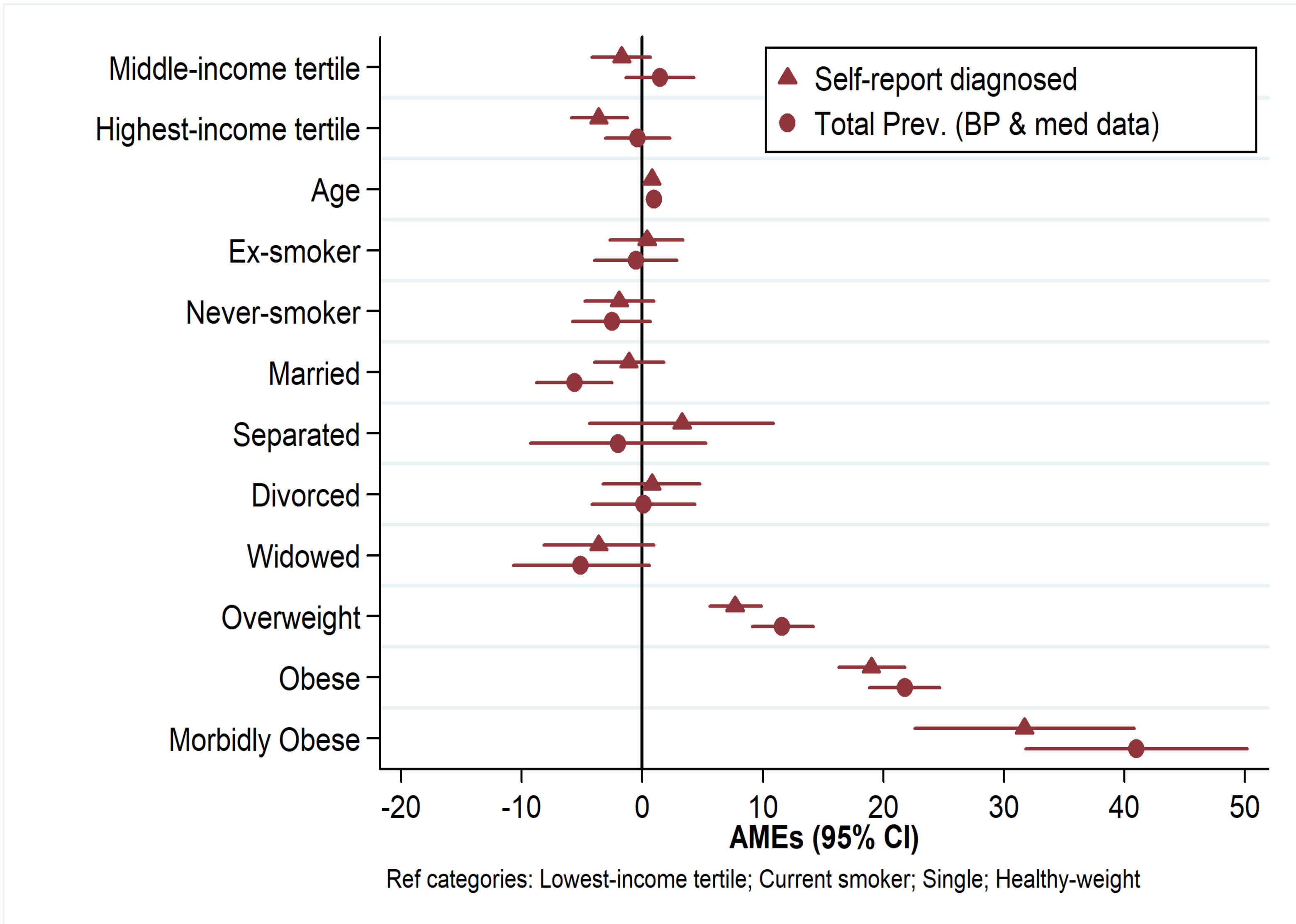

Figure 2: *Associations between income and hypertension from bivariate probit model. Lines represent AMEs and its 95% CI. Outcomes in the model were: (1) self-report, diagnosed hypertension (▲); and (2) total prevalence (●) based on objective and self-report data (BP & medication use). Negative value for the AME (▲) indicates that the estimated probability for men in the highest-income tertile reporting diagnosed hypertension was 3.6 percentage points lower (AME = -3.6; 95%CI: -5.9 to -1.2; p=0.003) than for men in the lowest-income tertile, holding all other covariates at their observed values.*

Differences in probability of being undiagnosed

Using the results from the bivariate probit model, **Figure 3** shows the estimated difference in the probability of being undiagnosed (defined in **Table 2**) between the highest and lowest income groups.

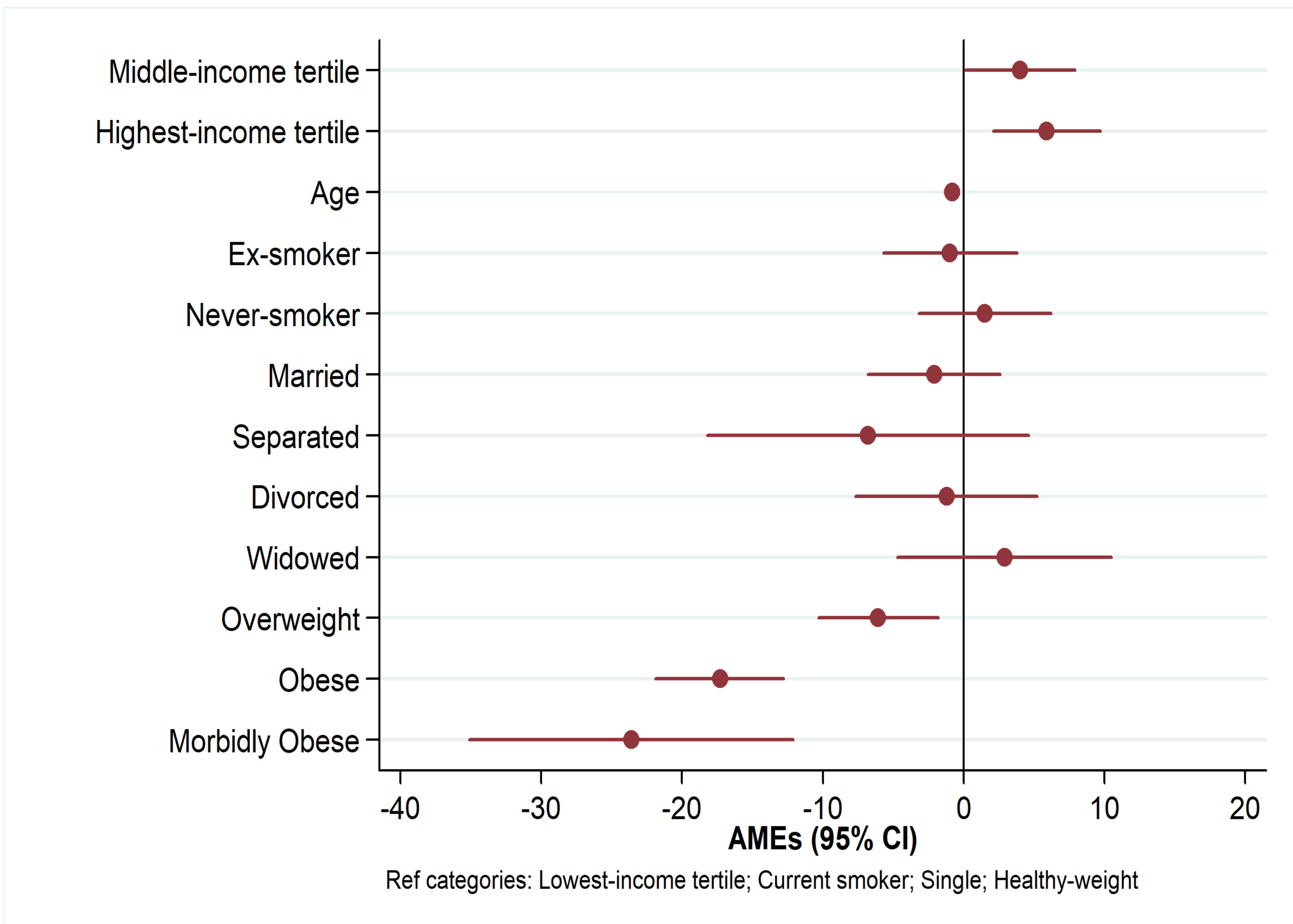

Figure 3: *Difference in probability of being undiagnosed from the bivariate probit model. Lines represent AMEs and its 95% CI. Positive value for the AME indicates that the estimated probability of being undiagnosed for men in the highest-income tertile was 5.9 percentage points higher (AME = 5.9; 95% CI: 2.1 to 9.7; p=0.002) than for men in the lowest-income tertile, holding all other covariates at their observed values.*

Conclusions

Hypertension was higher in the lowest income group, especially using self-report data. Men in the highest-income group were more (not less) likely to negatively report. Likely explanations for this finding include higher treatment rates (among hypertensive cases) in low-income groups and disease severity.

References

[1] Chatterji P, Joo H, Lahiri K. *Examining the education gradient in chronic illness*. Educ Econ 2015;**23**:735-750.

[2] Johnston DW, Propper C, Shields MA**.** *Comparing subjective and objective measures of health: Evidence from hypertension for the income/health gradient.* J Health Econ *2009;**28**:540-552.*

[3] Yin M, Augustin B, Shu C, Qin T, Yin P. *Probit Models to Investigate Prevalence of Total Diagnosed and Undiagnosed Diabetes among Aged 45 Years or Older Adults in China*. PLoS One 2016;**11(10)**: e0164481.