Institute for Fiscal Studies

How accurate are self-reported diagnoses? Comparing self-reported health events in the English Longitudinal Study of Ageing with administrative hospital records

IFS Working Paper W20/13

George Stoye
Ben Zaranko

UKRI Economic and Social Research Council

How accurate are self-reported diagnoses? Comparing self-reported health events in the English Longitudinal Study of Ageing with administrative hospital records

George Stoye and Ben Zaranko*

May 2020

Abstract

This paper uses linked survey responses and administrative hospital records to examine the accuracy of self-reported medical diagnoses. The English Longitudinal Study of Ageing (ELSA) collects self-reported information on the incidence of heart attacks, strokes and cancer in the past two years. We compare these reports with administrative hospital records to examine whether respondents are recorded as having an inpatient admission with these diagnoses during this period. We find self-reported medical diagnoses are subject to considerable response error. More than half of respondents diagnosed in hospital with a condition in the previous two years fail to report the condition when surveyed. Furthermore, one half of those who self-report a cancer or heart attack diagnosis, and two-thirds of those who self-report a stroke diagnosis, have no corresponding hospital record. A major driver of this reporting error appears to be misunderstanding or being unaware of their diagnoses, with false negative reporting rates falling significantly for heart attacks and strokes when using only primary hospital diagnoses to define objective diagnoses. Reporting error is more common among men, older respondents and those with lower cognitive function. Estimates relying on these self-reported variables are therefore potentially subject to sizeable attenuation biases. Our findings illustrate the importance of routine linkage between survey and administrative data.

*We thank Rowena Crawford, Carol Propper and seminar participants at the Institute for Fiscal Studies for helpful comments. We thank NHS Digital for providing access to the English Longitudinal Study of Ageing and the Hospital Episode Statistics under data sharing agreement CON-322640-S9V7X. This paper has been screened to ensure no confidential information is revealed. We gratefully acknowledge financial support from the Health Foundation (Efficiency Research Programme) and the Economic and Social Research Council through the Centre for the Microeconomic Analysis of Public Policy (CPP) at IFS (ES/M010147/1). Any errors are our own. Author affiliations: Stoye (University College London and Institute for Fiscal Studies, george_s@ifs.org.uk), Zaranko (Institute for Fiscal Studies, ben.zaranko@ifs.org.uk).

1 Introduction

Many researchers in medical and social sciences make extensive use of self-reported health data collected as part of surveys. However, there is concern that the accuracy of these measures may vary substantially across different population groups, based on characteristics such as age, sex, labour market status, education, income, cognitive function and institutional setting (Schrijvers et al., 1994; Bergmann et al., 1998; Baker et al., 2004; Okura et al., 2004; Corser et al., 2008; Johnston et al., 2009; Gupta & Jürges, 2012; Wolinksy et al., 2014; Wolinsky et al., 2015). Systematic errors in self-reported health measures threaten to introduce a considerable degree of measurement error into empirical analysis using such data, and pose particular challenges for studies aiming to accurately control for differences in health across different population groups.

In this paper, we examine the extent to which self-reported major health events captured by the English Longitudinal Study of Ageing (ELSA) are validated by administrative health records. ELSA contains detailed survey information on health conditions and a rich set of socioeconomic characteristics for a large, nationally representative sample of the English population aged 50 years and above.[1] For 80% of respondents, survey responses are linked to administrative hospital records containing detailed information on inpatient and outpatient hospital use. This enables us to compare self-reported medical events in the survey data with diagnoses in administrative hospital records.

Under the assumption that the hospital records contain the correct diagnoses, we study the accuracy of reporting for three major medical events: heart attacks, strokes and cancer. ELSA respondents are asked if a doctor has diagnosed them with each of these conditions in the previous two years. We focus on these cases as they are all serious medical diagnoses that should, in the vast majority of cases, require hospital treatment, and that are subject to national audit processes that give hospitals an incentive to record diagnoses accurately. As a result, hospital records should provide a good, objective measure of the incidence of these health events over the same period of time covered by the ELSA survey.

We use these data to sort responses into two categories. A 'false negative' is recorded when a respondent is recorded with having a hospital admission for a particular condition but does not report this diagnosis in the survey data. A 'false positive' is recorded if respondents report a particular diagnosis but do not have an associated hospital admission. Such under- or

[1] Similar surveys include the Health and Retirement Study (HRS) in the US and the Survey of Health, Ageing and Retirement in Europe (SHARE).

over-reporting threatens to introduce a substantial degree of measurement error into empirical analyses using self-reported health measures.

Across all respondents and all conditions, there is substantial disagreement between the diagnoses recorded in the self-reported data and in the hospital records. Of the respondents with a hospital admission for any of these three diagnoses in the previous two years, 63% did not report having been diagnosed with the condition (or conditions) during the same period when surveyed in ELSA ('false negatives'). The rate of false negative reporting was largest for heart attacks (67%), but was also substantial for cancer (59%) and stroke (54%). The false positive rate (the share of those without a hospital admission who self-report having the condition) is smaller across all conditions (1.0% for cancer and stroke, and 0.7% for heart attack). While smaller, this rate of false positive reporting means that of those who self-report having been diagnosed with cancer, 47% have no corresponding hospital record. The equivalent figures for stroke and heart attack are 69% and 47%, respectively.

The non-concordance between survey data and hospital records has a number of potential explanations, including patients misremembering or being unaware of the full extent of their past treatment, confusion between different conditions, unwillingness to truthfully discuss potentially sensitive health conditions with the interviewer, and inaccuracies in the hospital data. The final two explanations are difficult to assess with the current data.[2] To assess the importance of these first two channels we consider three alternative definitions of false positive and false negative reporting, in each case shutting off or reducing concerns about particular explanations.

First, our headline results use all hospital admissions where the patient was recorded with cancer, stroke or heart attack as a diagnosis, even where it wasn't the primary reason for the patient's admission. However, patients may be less likely to report conditions which are recorded as secondary diagnoses, either because they do not associate these conditions with the reason for attending hospital or because the doctor did not inform them of the diagnosis. We therefore repeat our analysis considering only cancer, stroke or heart attack diagnoses which were recorded as the main (primary) reason for admission.

When we carry out this exercise, rates of false negative reporting of cancer are broadly unchanged. However, they fall substantially for strokes (from 54% to 38%) and heart attacks (from 67% to 23%). This suggests that for heart attack and strokes, much of the inaccuracy in reporting is due to patients either not being fully aware of their medical history (potentially

[2]The implications of such misreporting errors are discussed below.

because they were not informed of this diagnosis by doctors at the time) or not placing as much importance on these diagnoses.

Second, we show that patients who have had multiple hospital admissions for a given condition are more likely to report the condition in ELSA. Patients with only a single admission for a particular condition may be less likely to report the relevant diagnosis, either because they forget the event or because they find it less important. Nonetheless, even on this definition, rates of false negative reporting do remain high: among patients who had four or more hospital admissions for cancer in the previous two years, close to half fail to report a cancer diagnosis when asked in ELSA. And importantly, while narrowing our analysis in this way (i.e. by using only primary diagnosis codes or by looking only at those with multiple hospital admissions) reduces the rate of false negative reporting, this comes at the cost of increasing the rate of false positive reporting.

Finally, we explore whether misreporting is driven by patients confusing similar, but distinct, medical conditions, or the timing of these events. To do so, we expand the definition of the condition in the hospital records to include any visits where they received treatment under a broadly related medical specialty: oncology (cancer), cardiology (heart attacks) and neurology (strokes). This reduces the rate of false positive reporting, but only by a modest amount. This suggests that this is not the main driving factor behind our results. Similarly, we find that confusion over the precise timings of health events (e.g. mistakenly thinking that a heart attack that occurred up to three years ago happened in the last two years) can explain only a small portion of the reporting error that we document.

We then examine variation in the accuracy of these responses across different population groups. Specifically, we study how the rates of false negatives and positives vary across age, sex, couple status, educational attainment and cognition. We find that older, cognitively impaired men are particularly prone to reporting error. The issue of cognition is of particular interest for the older population, where age-related cognitive decline is a concern for the collection of accurate survey data. After controlling for age, sex, educational attainment and couple status, we find that those in the top half of the distribution of cognitive function (as measured by orientation in time and retrospective memory score) are around 25% less likely to report a false positive, and 18% less likely to report a false negative, than those in the bottom half.

To the best of our knowledge, this is the first paper to compare self-reported health measures in ELSA with administrative hospital records, and represents the most comprehensive analysis of this type in a UK setting to date. An important contribution of this paper is therefore to extend previous work comparing 'subjective' and 'objective' measures of health to a representative sample of the older population in England and a widely used longitudinal survey. We also build on previous studies examining the individual characteristics associated with accurate self-reporting, showing in particular the importance of cognitive function (Schrijvers et al., 1994; Bergmann et al., 1998; Baker et al., 2004; Okura et al., 2004; Corser et al., 2008; Johnston et al., 2009; Gupta & Jürges, 2012; Wolinksy et al., 2014; Wolinsky et al., 2015).

The main conclusion of the analysis is that self-reported medical histories are potentially subject to a considerable degree of response error. Strokes and heart attacks do not appear to be well understood by patients when they are not the primary reason for the patient's admission to hospital, and there is some evidence that they are confused with related but distinct medical ailments. Rates of false negative reporting for cancer, especially, are worryingly and persistently high.

This builds a strong case for routine linkages between survey and administrative data to improve the accuracy and coverage of individual medical histories. This is particularly important for the groups (such as older, cognitively impaired men) that are most prone to misreporting. This would allow researchers to validate their survey responses, build a fuller picture of respondents' health, and facilitate important future research.

The rest of this paper is organised as follows. Section 2 sets out the existing related literature. Section 3 describes the data used and briefly summarises the institutional background of hospital care in England. Section 4 sets out the main results. Section 5 examines how the results vary across respondent characteristics. Section 6 concludes.

2 Related literature

Previous research has examined concordance between 'subjective' and 'objective' health measures in other settings. The study most similar to ours is Baker et al. (2004) who compare linked survey and administrative data collected during the late 1990s in Ontario, Canada. They conclude that self-reported measures for 13 separate ailments are subject to considerable reporting error. The probability of false positive reporting is found to be significantly lower among individuals in work, which the authors interpret as being consistent with unemployed respondents seeking to justify their absence from the labour market.

Gupta & Jürges (2012) use linked survey and register data in Denmark to explore a related hypothesis: that individuals in the labour market deny or seek to 'play down' health problems due to fears that employers will judge illness to be a signal of low productivity. They find widespread false negative reporting (under-reporting) of chronic health conditions, with higher rates among individuals in the labour market than individuals outside.

Age and cognitive function have been shown to matter for the accuracy of self-reports. Wolinsky et al. (2014) examine the degree of concordance between self-reports in the Survey on Assets and Health Dynamics among the Oldest Old (AHEAD) and Medicare claims in the US. They find evidence of both over-reporting (false positives) and under-reporting (false negatives) across multiple conditions, and show that reporting accuracy decreased with age. Wolinsky et al. (2015) also exploit linked AHEAD survey and Medicare claims data. They show that better mental status was associated with more concordant reporting and lower rates of under-reporting (false negatives).

Other characteristics have also been shown to be associated with reporting accuracy, though typically this has been done using smaller samples that are not necessarily representative of the wider population. Okura et al. (2004) examined the agreement between self-reported disease history and medical records in Minnesota, USA. They found substantial false negative (under) reporting, but a greater degree of agreement between data sources for diseases with acute onset (such as stroke and heart attack). Self-reporting was more accurate amongst young, female, and better-educated respondents. This is consistent with Schrijvers et al. (1994) who found that rates of false negative reporting for cancer in the Netherlands were lower among the same groups (young, female and better educated survey participants). Corser et al. (2008) found a lesser degree of concordance between survey responses and medical records for older individuals and those with higher depressive symptom levels. Bergmann et al. (1998) found higher levels of education to be associated with more accurate self-reporting, but no relationship with age or gender. Johnston et al. (2009) compare subjective and objective measures of hypertension in England. Rates of false negative reporting (self-reporting that they do not have hypertension, but being measured by a trained nurse as having hypertension) decreased with household income.

3 Data and definitions

Our analysis is based on a dataset that combines seven waves (2004-05 to 2016-17) of the English Longitudinal Study of Ageing (ELSA) with the census of public hospital records in England, the Hospital Episode Statistics (HES). This provides a rich set of survey data, including self-reported health events, which can be compared with medical diagnoses recorded in administrative hospital records. We briefly describe each of these datasets and the linked sample used for analysis below.

3.1 English Longitudinal Study of Ageing

ELSA is a panel survey of a representative sample of the English household population aged 50 and above.[3] The survey is administered every two years, and follows the same individuals over time. Respondents are interviewed on a range of core topics including demographic and economic characteristics, their health and wellbeing, and household and family structure. We use seven waves of the survey, collected between 2004-05 and 2016-17 (waves 2 - 8).[4]

All participants have a face-to-face interview, which consists of a computer-assisted interview with a trained interviewer and a self-completed questionnaire (which may or may not be completed during the interview). Each survey wave collected information on whether respondents had been told by a doctor in the previous two years that they had a number of different conditions.[5] This includes heart attack (including myocardial infarction or coronary thrombosis), stroke (cerebral vascular disease), and cancer or a malignant tumour (excluding minor skin cancers). In total, ELSA asks about 19 health conditions. We do not study the remaining conditions for four reasons. First, these conditions can be difficult to map directly into ICD-10 codes contained in hospital records (e.g. lung disease). Second, these conditions may not always require hospital treatment or could be diagnosed in a primary care setting (e.g. diabetes). As a result, these patients may not appear in hospital records. Third, a sizeable share of those with some conditions, like hypertension, may not even be aware of that fact (Johnston et al., 2009). This is considerably less likely to be the case for conditions like stroke and heart attacks, which tend to involve an abrupt onset associated with pain and/or a loss of function. Finally, for the

[3] The initial sampling frame excluded people living in institutions, such as nursing homes. Subsequent survey waves followed patients into these institutions but they remain under-represented in the data. We exclude all institutional responses from our analysis.

[4] Survey collections took place in 2002-03, 2004-05, 2006-07, 2008-09, 2010-11, 2012-13, 2014-15 and 2016-17.

[5] Data in the first wave collected information on lifetime incidence of each condition. However, we cannot be sure when this diagnosis took place and therefore cannot compare with hospital records over a limited period of time. As a result, we use data from the second wave onwards.

three conditions on which we focus, hospital records are subject to national audit processes that provide hospitals with clear incentives to record patient diagnoses accurately.[6] Consequently, hospital records for cancer, stroke and heart attack should provide a good, objective measure of the incidence of these health events for our period of interest. This does not, however, necessarily mean that patients are fully aware of the contents of those records, and in particular may be unaware of secondary diagnoses assigned but never disclosed to the patient by the doctor.

Respondents were also directly asked whether they had received any treatment for cancer during the previous two years. We use these data to create a binary indicator for each condition that takes the value of one if respondents report that a doctor diagnosed them with the condition in the past two years, and zero otherwise. We also create an indicator of whether the person reports *ever* having been diagnosed with the condition, and of whether they report receiving cancer treatment over this period.

ELSA collects information on a rich set of demographic and economic characteristics. In Section 5, we examine whether the accuracy of reported answers varies across people with different characteristics. In particular, we examine how answers vary across sex, age, educational attainment, and cognitive function.

We transform individual educational attainment into a three point scale: 'low' (no formal qualifications), 'mid' (NVQ1/NVQ2/NVQ3, O level, A level or equivalent) and 'high' (NVQ4/NVQ5, higher education below degree, or degree equivalent). We also create a dummy variable equal to one if the individual is highly educated.

Respondents are also asked to carry out a number of verbal and physical assessments to assess their mental and physical health, and their cognitive ability. We use these assessments to construct a 'cognitive score' for each individual. This score has two components. First, each respondent is asked to identify the day of the month, the month, the year, and the day of week. This gives each respondent a score out of 4 for their orientation in time. The survey also includes a word recall test, and a delayed recall test, both of which are scored out of 10, giving a total score out of 20 for retrospective memory. We combine these measures to give a combined 'cognitive score' out of 24, and split the sample into 'high cognition' (above average) and 'low cognition' (below average).

[6]These are the Sentinel Stroke National Audit Programme (SSNAP, https://www.strokeaudit.org/), the Myocardial Ischaemia National Audit Project (MINAP, https://www.nicor.org.uk/national-cardiac-audit-programme/myocardial-ischaemia-minap-heart-attack-audit/) and the National Cancer Registration and Analysis Service (NCRAS, https://www.cancerdata.nhs.uk/).

3.2 Hospital Episode Statistics

HES contains the census of visits to publicly funded hospitals in England.[7] These records include all inpatient (admitted) care between April 1997 and January 2018. This provides at least 7 years of hospital records for all respondents before each interview.[8] Inpatient records include admission and discharge dates, the specialty of the consultant assigned to each patient, and up to 20 detailed diagnoses codes. Diagnoses are recorded as ICD-10 codes, and include the primary reason for admission and up to 19 secondary diagnoses.

HES also contains information on all outpatient (non-admitted) treatment provided to NHS patients between April 2003 and January 2018. This captures all outpatient appointments with NHS consultants during this period, and includes information on the date of this appointment and the specialty of the consultant. However, exact diagnosis data are not available.

Combined, these data capture the vast majority of hospital activity in England. There is no private market for emergency hospital care. As a result, essentially all acute hospital treatment for urgent conditions - such as heart attacks and stroke - should be observed in the data. There is a small private market for elective care, which can include some cancer treatment that is privately financed. However, few people - particularly in the older age group that we study - have private medical insurance.[9] The number of self-payers is also small. As a result, the data should give us a high quality overview of hospital use among people with our three conditions of interest.

Throughout our analysis, we use diagnoses recorded in the hospital records as an objective measure of health conditions. In practice, some diagnoses in these data may be incorrect. This could be either due to errors in the data inputting process, or because hospitals purposefully enter more, or more severe, diagnoses to receive additional compensation.[10] These records also only contain diagnoses for patients who sought treatment in an NHS hospital in England. Patients who did not receive treatment, or treatment received abroad or in private hospitals, would therefore not be included in the data, and may lead to under-reporting of some conditions.

[7] This includes care in public hospitals, and publicly funded care in private hospitals.

[8] The first ELSA interviews in our sample took place in June 2004.

[9] 13.1% of our sample, and 9.8% of those aged 65 and above, are covered by private health insurance.

[10] NHS hospitals in England receive nationally set payments from the government for providing specific treatments to patients. These payments are set to cover the average cost for specific treatments or procedures. Payments also vary in some instances when patients are more complex (e.g. have certain secondary diagnoses) and therefore cost more to treat on average. Evidence from the US suggests that these financial incentives have led hospitals to enter more, and more severe, secondary diagnoses in order to attract additional compensation, in a process known as ‘upcoding’ (Geruso and Layton (2020); Cook and Averett (2020)). The extent to which this may occur in NHS hospitals is uncertain.

We discuss below how this would affect potential misreporting rates.

3.3 Constructing the linked sample

To examine the medical diagnoses of patients recorded in hospital records, we link the survey responses of all consenting ELSA respondents to their official hospital records contained in HES using a pseudonymised unique identifier for each respondent. 80% of ELSA respondents (comprising 85% of observations across the seven waves in our sample) provided consent for their hospital records to be linked to their survey responses. This provides a sample of 11,218 people across 49,459 interviews.

We use the date of each survey interview to keep all hospital records (both inpatient and outpatient) that took place in the 730 days (2 years) prior to the interview. We initially use only the diagnosis information from the inpatient records to create indicators of whether a person has an admission over this period where they have a diagnosis of a heart attack, stroke or cancer. We map each of these conditions to a set of ICD-10 codes. Appendix Table A1 shows which codes are assigned to each condition. We create two main indicators for each condition: the first takes the value of one if a relevant code is recorded in any of the primary or secondary diagnosis categories, and zero otherwise. The second repeats this process but uses only the primary diagnosis code. This addresses concerns that patients may only be aware of their primary reason for the hospital visit and may be unaware of (or fail to recall) secondary diagnoses recorded by the doctor. We also create count variables to track the number of admissions each respondent experiences where the relevant condition is recorded.

As a further examination of the underlying factors that may explain any misreporting, we also create two further indicators of having been in hospital with a heart attack, stroke or cancer. These measures are intended to capture a broader set of hospital diagnoses that patients may falsely identify as a heart attack, stroke or cancer diagnosis.

First, using inpatient hospital records, we assign an indicator a value of one if the patient has the relevant code recorded as any of the primary or secondary diagnoses *or* if the patient is recorded as an inpatient under the care of a consultant of the relevant specialty (as shown in Appendix Table A1). Second, using both inpatient and outpatient records, we construct an indicator equal to one if the patient has a relevant diagnosis code as an inpatient, *or* is recorded as either an inpatient or outpatient under the care of a consultant of the relevant specialty. Including outpatient visits is likely to be particularly important for cancer diagnoses, where treatment (such as chemotherapy) often take place in an outpatient setting. We can only construct this final measure for survey wave 3 and onwards, as we require two years of hospital records prior to the interview.[11]

In both cases, by using a broader definition based on consultant specialty, we may capture a wider set of related conditions that patients may mistakenly attribute as the specific diagnoses that they are asked about in the ELSA interview. However, we will also capture some patients who received treatment for a different condition to those we focus on here (e.g. patients who saw a neurologist for a condition other than a stroke).

3.4 Definitions

Our main analysis focuses on the incidence of two types of reporting inaccuracies.

First, we define a 'false negative' as having occurred when the individual does not report having been diagnosed with a condition in ELSA, but is recorded as having had a hospital visit with the relevant diagnosis.

Second, a 'false positive' occurs when an individual reports having been been diagnosed with a particular condition, but does not have any hospital visits corresponding to the condition in question.

The incidence of each is defined as follows:

$$\begin{aligned} &\text{False negatives:} \quad \Pr(SelfReport = 0 \mid HospRecords = 1) \\ &\text{False positives:} \quad \Pr(SelfReport = 1 \mid HospRecords = 0) \end{aligned} \tag{1}$$

We therefore calculate the false negative rate as the proportion of individuals with a hospital record indicating the presence of condition who do not report the condition in ELSA. The false positive rate is calculated as the proportion of individuals without a hospital record who do report having the condition. In our baseline analysis, we define $SelfReport$=1 if the individual reports in ELSA the onset of the condition in the past two years, and $HospRecords$=1 if the individual is recorded in HES with an inpatient spell in the last two years, with either a primary or secondary diagnosis recorded as the relevant condition. We then employ alternative

[11]Outpatient records are available from April 2003, but the first interviews in wave 2 of ELSA took place in June 2004.

definitions to test the underlying factors that may explain these results.

4 Comparing self-reported health events with administrative diagnoses

We start in Section 4.1 by outlining our results using our baseline definition. We then employ various definitions to explore the importance of potential channels through which false negatives and false positives could arise in Sections 4.2 and 4.3.

4.1 Main analysis

The first three columns of Table 1 show the results of our baseline analysis. Column 1 shows the mean incidence of each condition in our sample, based on both primary and secondary diagnosis codes in hospital records. This varies by condition: 2.7% had a hospital admission with a (primary or secondary) cancer diagnosis in the previous two years, versus 0.9% for a stroke and 2.2% for a heart attack. Columns 2 and 3 show how self-reported (subjective) health events compare to (objective) administrative hospital records. 59% of of those who were diagnosed with cancer in hospital in the past two years fail to report having cancer when asked in ELSA (false negatives). The equivalent figures for strokes and heart attacks are 54% and 67%, respectively. Of those who have no hospital diagnosis of cancer in the past two years, around 1% self-report having cancer (false positives). The rates of false positives for stroke and heart attacks are similar in magnitude (1% and 0.7%, respectively). To put the incidence of false positives another way: of the 988 people who report a new diagnosis of cancer in the past two years, almost half (460, or 47%) do not show up in hospital inpatient records with a cancer diagnosis in that period.

These results suggest that there is substantial disagreement between self-reported diagnoses and the diagnoses recorded in the administrative health records. This could occur for a number of reasons. We therefore explore alternative definitions of false negatives and false positives to provide evidence on why misreporting might occur.

4.2 False negatives

A false negative arises when an individual is recorded in hospital with a condition ($HospRecords$=1) but does not report having it ($SelfReport$=0). This could happen for a number of reasons.

Individuals may struggle to remember and accurately report their medical history, particularly for one-off (and therefore perhaps less salient) health events. They may be unaware of the precise reasons for the treatment received in hospital (potentially confusing this with another diagnosis), particularly where this was not the main reason for their admission. This is especially likely to be the case in our sample, composed entirely of individuals aged 50 and above, for some of whom age-related cognitive decline will be a concern (an issue to which we return in section 5).

To examine this further, we use alternative definitions of false negative reporting. Our initial results used both primary and secondary hospital diagnoses. That is, we assign a respondent a value of $HospRecords$=1 if they are admitted to hospital with the condition in question at any point in the previous two years, even where that condition was not the primary reason for their admission. However, we might expect patients to be aware only of the main (primary) reason for their admission, and not to always be aware of the full extent of their ailments upon hospitalisation. For example, an individual could fall after a minor stroke, but remember the subsequent hospital admission as being only for the (more salient) fractured wrist caused by the fall, and not the stroke. In some cases the individual may not even be aware that a stroke was recorded as a secondary diagnosis by their hospital doctor. If the individual then failed to report having had a stroke when asked in ELSA, they would technically still be misreporting their medical history (as they have in fact had a stroke), but we might interpret that misreporting differently.

To examine whether this misunderstanding is driving the false negative reporting we therefore repeat the exercise using only the primary diagnosis recorded in the hospital records.[12] The results of this exercise are shown in columns 4-6 of Table 1. The mean incidence of each condition (shown in column 4) falls relative to our baseline analysis: from 2.7% to 2.4% for cancer, 0.9% to 0.6% for stroke and 2.2% to 0.6% for heart attack. This indicates that heart attacks in particular are frequently a secondary, rather than primary, diagnosis.

For cancer, the rate of false negative reporting (shown in column 5) is broadly unchanged relative to our baseline results (57% versus 59%). For strokes and heart attacks, however, using only primary diagnosis codes has a considerable effect. The rate of false negative reporting for strokes falls to 38% (from 54% when all diagnosis codes are used), and the rate for heart attacks

[12]Using primary diagnoses might also reduce concerns about the manipulation of diagnoses due to financial incentives ('up-coding') as these would be likely confined only to secondary diagnosis codes that are less well recorded and monitored.

falls to 23% (from 67%). This would suggest that when stroke or heart attack is recorded as a secondary diagnosis, patients are seemingly less aware (because they aren't told, or because only the primary diagnosis is salient), and that this plays an important role in under-reporting these specific conditions.

It is important to note that while narrowing our analysis to focus only on primary diagnoses reduces the rate of false negative reporting, this is traded off with an increase in the rate of false positive reporting for all three conditions. This is because some individuals are admitted to hospital with cancer, stroke or heart attack as only a secondary diagnosis, and then (correctly) report having been diagnosed when asked in ELSA. Such individuals are treated here as having reported a false positive, because they lack a relevant *primary* hospital diagnosis. Conversely, employing a broader mapping algorithm could bring down rates of false positive reporting, but at the cost of increased rates of false negative reporting (something we return to in Section 4.3).

If patient misunderstanding of their condition is important, we might also expect more accurate reporting when looking only at patients who make multiple hospital visits for the same reason. We explore this question in Table 2. Columns 1 to 3 show our baseline results (using primary and secondary diagnoses), as a point of comparison. In columns 4 and 5, we assign $HospRecords$=1 only if the individual had at least two inpatient admissions in the two years prior to interview, with a primary or secondary diagnosis recorded as the condition in question. In columns 6 and 7 we require a minimum of three admissions, and a minimum of four admissions in columns 8 and 9.

For cancer, the pattern is one of reduced rates of false negative reporting among those using hospitals more intensively. Rates of false negative reporting fall from 59% in our baseline results (requiring a minimum of one hospital admission) to 47% in column 8 (where we look only at those who were admitted at least four times). This means, however, that even among those who had 4 or more hospital admissions for cancer in the two previous years, close to half fail to report having been diagnosed in that time.[13]

For stroke, rates of false negative reporting do not appear to change with intensity of use. For heart attack, rates fall when we limit our analysis to those with a greater number of admissions, but remain in excess of 60%. In other words, of those individuals who had a minimum of four hospital admissions with heart attack recorded as a diagnosis code, more than 6 in 10 fail to

[13]This result also reduces concerns that false negative rates are driven by unintentional errors in the recording of hospital diagnoses, as we wouldn't expect someone with an incorrect preliminary cancer diagnosis to have the same incorrect diagnosis recorded in four separate hospital spells.

report having been diagnosed by a doctor with a heart attack in the previous two years.

In Table 3 we repeat this analysis for patients with more than one hospital admission with the condition in question recorded as the primary diagnosis. Columns 1 to 3 again show the results when we only require a minimum of one admission with a relevant primary diagnosis (as in columns 4-6 of Table 1). In columns 4 and 5, for *HospRecords*=1 we require at least two hospital admissions with a primary diagnosis code for the condition in question. For all three conditions, the rate of false negatives falls, to around one-half, one-third and one-fifth for cancer, stroke and heart attack, respectively. Again, this reduction comes at the cost of a substantial increase in the rate of false positive reporting relative to our baseline results: from 1.0% to 1.6% for cancer, 1.0% to 1.4% for stroke, and 0.7% to 1.3% for heart attacks. Nonetheless, these results are consistent with less under-reporting of specific conditions when treatment for these conditions was more salient.

Finally, we consider an alternative definition that addresses concerns over the timing of diagnoses. For strokes and heart attacks, which are typically characterised by an acute and clearly defined health event, we might worry that individuals will misremember exactly when a diagnosis occurred. In addition, for chronic conditions (such as cancer) where recurrent diagnoses are possible, the onset and diagnoses of the illness may have been more than two years ago but treatment may be ongoing. In such a case, the respondent would not report having had a new diagnoses in the last two years but would be recorded receiving hospital treatment for this diagnosis. They would be classified as reporting a false negative as a result.

To address this, in columns 6 and 7 of Table 3 we assign $SelfReport$=1 to all individuals who report *ever* having been diagnosed with the condition. The results show that the rates of false negative reporting for each condition is considerably lower than in our baseline analysis: false negative reporting falls to 31.4% for cancer, to 23.4% for heart attacks, and to 33.9% for strokes. This suggests that confusion over the timing is likely to be a factor in explaining why patients under-report these conditions.[14]

In addition to being asked whether they have been diagnosed in the past two years, respondents in ELSA are asked if they have *received treatment* for cancer in the past two years. Using this as our measure of self-reporting partly addresses concerns over timing, for as long as the patient records having received treatment in the past two years, the precise timing of

[14]The false positive rates are considerably higher when using this definition as we would expect given that respondents who (correctly) identify past diagnoses will not have a corresponding hospital admission during the same period.

the diagnosis does not matter. Column 8 of Table 3 shows that of those recorded in hospital as a cancer patient, 48% report that they have *not* had cancer treatment in the past two years. Clearly, important differences between the subjective and objective measures remain.

Taken together, these results suggest that limited awareness or recall of all relevant diagnoses, and confusion of the precise timing of conditions, are likely to be important factors in explaining the high rate of false negative reporting. In particular, the differences in these rates for heart attacks and stroke when using only primary hospital diagnoses is striking. However, even after these adjustments, relatively large differences remain in the incidence of these health events as recorded by the survey and administrative data.

4.3 False positives

We now turn to potential drivers of false positives. A false positive is recorded if the individual has no relevant hospital admission ($HospRecords$=0) but self-reports having been diagnosed with the ailment in question ($SelfReport$=1).

This could arise for a number of reasons, including confusion between precise diagnoses, receiving hospital treatment outside of an NHS inpatient setting, and misremembering the exact timing of a health event. In Table 4, we employ a number of definitions to explore the importance of some of these. Columns 1 to 3 report our headline results, as in columns 1 to 3 of Table 1. This serves as a comparison point for our other measures.

One reason why a false positive may occur is that individuals do not fully understand their medical history and misreport a similar (but distinct) ailment as being cancer, a heart attack or stroke.[15] In columns 4 and 5 of Table 4, we use a broader mapping of hospital records to conditions, by defining $HospRecords$=1 if the patient has a relevant diagnosis code *or* is admitted under the care of a consultant of the relevant medical specialty. This means, for example, that someone who self-reports having had a heart attack, and is recorded as having seen a cardiologist (with a non-heart attack diagnosis code), is no longer recorded as a false positive (as they have both $SelfReport$=1 and $HospRecords$=1). This should reduce the number of false positives due to differences in the knowledge that patients have over their exact diagnosis.[16] If this was driving our results, we would expect to see a sharp reduction in the

[15] Bergmann et al. (1998) find evidence of patients self-reporting a stroke diagnosis that in fact reflected a diagnosis of other conditions of anatomic proximity.

[16] This could additionally address two other explanations for a high rate of false positives. First, our mapping of diagnosis codes may be too narrow and we may not capture all of the respondents who have suffered a specific condition. Second, hospitals may on occasion mis-record precise diagnoses. Implicitly expanding the diagnoses

rate of false positives under this new definition. However, the results in column 5 show that the reduction in the incidence of false positives for cancer and stroke is very small. This reduction is larger for heart attacks (from 0.7% to 0.5%).

This suggests that patient misunderstanding of diagnoses does not explain the majority of misreporting. In addition, employing this broader definition also comes with a trade-off: column 4 shows that the incidence of false negatives remains broadly unchanged for cancer, but rises for stroke and heart attack, relative to our baseline definition.

An alternative reason for a high false positive rate is that patients do not receive inpatient treatment in an NHS hospital in England for the specific condition. This is unlikely to be a major concern for heart attacks or stroke, where due to the nature of the illness an emergency admission is typically required quickly. However, it may be more applicable to cancer, where outpatient treatment is common. To address this, in columns 6 and 7 of Table 4 we further widen our mapping of hospital records to conditions, using outpatient as well as inpatient records. Outpatient records do not contain detailed information on diagnoses, so we again make use of the consultant's specialty. An individual is now recorded with $HospRecords$=1 if they have an inpatient admission with a relevant diagnosis code, or if they are recorded under the care of a relevant consultant as either an inpatient or outpatient. If many false positives were because patients received only outpatient care, we would expect a marked decrease in the rate of false positives relative to the results in column 5. Column 7 shows that the incidence of false positive reporting of cancer falls by a moderate amount (from 1% to 0.7%), falls slightly for strokes (from 1% to 0.9%), and remains unchanged for heart attacks (at around 0.5%). The rate remains positive and of a considerable magnitude in each case. These sources of error therefore explain some, but by no means all, of our results. Again, using a broader mapping comes at the cost of marked increase in the number of false negatives (as shown column 6).

Another potential reason for false positive reporting is patients misremembering the precise timing of their health event. For example, when asked, a person who was admitted to hospital with a heart attack 25 months ago may mistakenly (but understandably) respond in the affirmative when asked if they have had a heart attack in the past two years. Using our baseline definition would result in the recording of a false positive. To address this, in columns 8 and 9, we use hospital records from the three years prior to interview, rather than two. Individuals such as those in the example above would no longer be recorded as false positives.

used in our definitions by using specialty code would reduce both concerns.

Using this definition reduces the rate of false positives for all three conditions relative to our baseline results, but only by a very modest amount, indicating that some patients do indeed misremember the dates of health events, but that this channel is not the driving factor behind our false positive results.

5 How does reporting vary across respondent characteristics?

All of our results so far have been for our entire sample. We now consider how the likelihood of false negative or false positive varies across individuals to gain a better sense of the characteristics associated with non-concordance between subjective and objective measures of health events. Of particular interest is individuals' cognitive function. Our previous results suggested that one key explanation for the lack of concordance between objective and subjective measures of health events is that patients either do not fully understand or remember their diagnoses. Age-related cognitive decline is one reason why this might occur.

We first consider false negatives, estimating a logistic regression with the reporting of a false negative as the dependent variable.[17] The results are shown in Table 5. In the first two columns we look at the probability of reporting a false negative for any of the three conditions. Column 1 includes only a dummy for having high cognitive function.[18] The odds ratio of 0.68 indicates that individuals with high cognitive function are around a third less likely to report a false negative as those with low cognitive function. In column 2, we also include age-band[19], sex, couple status and education level as control variables. The results indicate that false negative reporting is considerably less prevalent among women than men. Although the results are not statistically significant, the results indicate that rates of false negative reporting are higher among older individuals, particularly those aged 80 and above. There are no statistically significant differences across education or couple status. The coefficient on high cognition remains statistically significant at the 5% level and indicates that even after controlling for these other characteristics, those with high cognitive function are close to 20% less likely to report a false negative (odds ratio of 0.82).

In columns 3 to 8, we examine how these results vary across our three conditions of interest.

[17] Note that the sample here is those who are observed in hospital records with a relevant diagnosis code. Those who report the condition in ELSA have FN=0, and those who fail to report have FN=1.

[18] As noted in Section 3.1, an individual is defined as having high cognitive function if their combined score for orientation in time, word recall and delayed recall is in the top half of the distribution.

[19] We assign respondents to eight age-bands: 50-54, 55-59, 60-64, 65-69, 70-74, 75-79, 80-85 and 85+.

Women are roughly half as likely to report a false negative for cancer (odds ratio of 0.55), but there are no significant differences between men and women on the other conditions. After controlling for other individual characteristics, individuals with high cognitive function are around 25% less likely to report a false negative for cancer. The odds ratio for stroke is less than 1 (0.82) but not statistically significant; for heart attack it is close to and not statistically distinguishable from 1. This suggests that lower cognitive function is an important driver of the under-reporting of cancer diagnoses, but that false negative reporting of strokes and heart attacks is driven by other factors.

In Table 6 we repeat the analysis for false positives. Here, the sample is much larger, consisting of all individuals who were *not* observed with a relevant hospital admission.[20] Those who subsequently reported the condition in ELSA have a value of FP=1. The results are broadly similar. The first column indicates that those with high cognition are around half as likely to report a false positive for any condition (odds ratio of 0.53). After controlling for other individual characteristics, such individuals are 25% less likely (odds ratio of 0.75). This points to difficulty understanding the precise nature of ailments among the cognitively impaired as a possible driver of false positive reporting. The results indicate that older individuals are more likely to report a false positive. Compared to a respondent aged 50-54, those aged 80-84 and 85+ are 3.5 and 3.2 times more likely to report a false positive, respectively. Men are more likely to report a false positive than women. Together with the results above, this shows that older men are more likely to both over-report and under-report the 'true' extent of their medical problems.

Columns 3 to 8 then repeat the analysis separately for cancer, stroke and heart attack. After controlling for other individual characteristics, the odds ratio on cognitive function is lower than 1 for all three conditions (indicating that high cognition individuals are less likely to report a false positive), but statistically significant only for stroke. This suggests that stroke diagnoses are perhaps less salient or more difficult to understand, and thus are more likely to be forgotten by those with impaired retrospective memory and cognitive function. Older individuals are more likely to report a false positive for all three conditions, with the age gradient particularly marked for strokes: those aged 80 and above are more than 6 times more likely to falsely report a stroke diagnosis than those aged 50-64. Men are more likely than women to report a false

[20]The previous analysis (of false negatives) is limited to those with a relevant hospital admission and this makes sample sizes relatively small.

positive for cancer and heart attack, but not for stroke. Notably, women are 60% less likely to report a false positive for heart attacks (odds ratio of 0.40). Highly educated individuals and those in a couple are less likely to do so (with odds ratios of 0.77 and 0.69 respectively), but our results indicate no statistically significant differences across education or couple status for cancer and stroke.

Overall, our analysis strongly suggests that older individuals, men, and those with lower cognitive ability are more likely to misreport (in terms of both false negatives and false positives). Subjective and objective measures of health events are less concordant for these individuals as a result. The pattern across conditions is more complex: low cognitive ability and being male are strongly associated with false negative reporting of cancer diagnoses, but not of strokes or heart attacks. When it comes to over-reporting, those with low cognitive ability are twice as likely to report a false positive for a stroke, but no statistically significant differences exist for cancer and heart attacks. Being highly educated and in a couple is associated with lower rates of false positive reporting for heart attacks (but not the other two conditions). This suggests that the way in which individuals understand, remember and report their medical history varies considerably across different conditions.

6 Conclusion

Self-reported health measures recorded in survey data are frequently used by researchers. This includes specific questions about an individual's disease history. In this paper, we examine the extent to which self-reported measures of major health events in the English Longitudinal Study of Ageing align with administrative hospital records. Under the assumption that hospital records represent an 'objective' and accurate record of such health events, we assess the accuracy of the self-reported ('subjective') diagnoses of cancer, strokes and heart attacks in ELSA. To the best of our knowledge, this is the first paper to do so, and represents the most comprehensive analysis of this type in a UK setting to date.

We find a substantial degree of disagreement between the diagnoses recorded in self-reported data and administrative hospital records. Of those admitted to hospital with a diagnosis of cancer, stroke or heart attack, more than half fail to report having been diagnosed with the condition in the survey data (false negatives). Of those who are not recorded as having had a relevant hospital admission, between 0.7% (for heart attacks) and 1.0% (for cancer and strokes) of individuals report having been diagnosed (false positives). This means that close to half of those who self-report having been diagnosed with cancer, for example, have no corresponding hospital record. These results are similar in magnitude to those found in other studies for other settings, most notably by Baker et al. (2004) using data collected in Ontario, Canada in the late 1990s, and suggests that such reporting errors in medical diagnoses are a common issue across survey datasets.

We employ a number of alternative definitions to assess the importance of various potential explanations for both false negative and false positive reporting. We find that respondents are considerably less likely to fail to report a stroke or heart attack when they have had a hospitalisation for which stroke or heart attack was the main (primary) reason for admission. This indicates that for these two conditions, much of the inaccuracy in self-reported data stems from either patients not being fully aware of their medical history (perhaps because they were not informed of the full extent of their diagnoses by doctors at the time), or patients placing less importance on secondary (non-primary) diagnoses. Researchers looking specifically at these conditions should therefore be aware that self-reported measures may only capture major strokes and heart attacks. We also find some evidence that patients confuse strokes and heart attacks with similar, but distinct, conditions, and some evidence of confusion over the precise timings of health events. Both of these channels play only a minor role in explaining our results, however.

For cancer, rates of false negative reporting are worryingly and persistently high. Of those observed in hospital with a cancer diagnosis in the previous two years, 59% fail to report a diagnosis in that time when surveyed in ELSA, and 48% report that they have *not* had cancer treatment over that period. Whereas strokes and heart attacks are acute conditions typically characterised with an abrupt onset, cancer is more chronic in nature and recurrent diagnoses are possible. We show that expanding our analysis to those who report *ever* having been diagnosed with cancer (as opposed to those who report having diagnosed in the past two years) reduces rates of false negative reporting, but that there remains substantial non-concordance between self-reported measures and administrative hospital records.

Some groups are more subject to errors in self-reporting than others. In line with previous work, we show that older, cognitively impaired men are more prone to both under-reporting and over-reporting. We find low cognitive ability to be associated with greater likelihood of under-reporting cancer, and over-reporting stroke diagnoses, in particular. The very oldest individuals are particularly prone to false positive reporting.

The disagreement between self-reported health measures and hospital records is likely to be important for studies that use these variables to control for individual health, or are used as ‘shocks’ to study impacts on other outcomes: for example, declines in health based on these measurements may also reflect broader changes in health or cognitive decline that affect patterns of self-reporting. Concerns over these data quality could be mitigated through routine linkage between surveys and more objective measures, such as biomarkers or administrative hospital records. These measures could then be used to validate self-reported diagnoses, used in combination with, or as an alternative control. When self-reported health events cannot be verified in this way, researchers should exert caution and, where possible, control for objective measures of cognitive and mental function.

References

1. Baker, M., Stabile, M. & Deri, C. (2004), 'What Do Self-Reported, Objective, Measures of Health Measure?', *Journal of Human Resources*; 39(4):1067-1093.

2. Bergmann, M.M., Byers, T., Freedman, D.S. & Mokdad, A. (1998), 'Validity of Self-reported Diagnoses Leading to Hospitalization: A Comparison of Self-reports with Hospital Records in a Prospective Study of American Adults', *American Journal of Epidemiology*; 147(10): 969-977.

3. Cook, A. & Averett, S. (2020), 'Do hospitals respond to changing incentive structures: Evidence from MEdicare's 2007 DRG restructuring', forthcoming in *Journal of Health Economics.*

4. Corser, W., Sikorskii, A., Olomu, A., Stommel, M., Proden, C. & Holmes-Rovner, M. (2008), 'Concordance between comorbidity data from patient self-report interviews and medical record documentation', *BMC Health Services Research*; 8:85.

5. Geruso, M. & Layton, T. (2020), 'Upcoding: Evidence from Medicare on squishy risk adjustment', *Journal of Political Economy*; 128(3):984-1026 .

6. Gupta, D. & Jürges, H. (2012), ''Do workers underreport morbidity? The accuracy of self-reports of chronic conditions', *Social Science & Medicine*; 75(9): 1589-1594.

7. Johnston, D.W., Propper, C., Shields, M.A. (2009), 'Comparing subjective and objective measures of health: Evidence from hypertension for the income/health gradient', *Journal of Health Economics*; 28(3): 540-552.

8. Okura, Y., Urban, L.H., Mahoney, D.W., Jacobsen, S.J. & Rodeheffer, R.J. (2004), 'Agreement between self-report questionnaires and medical record data was substantial for diabetes, hypertension, myocardial infarction and stroke but not for heart failure', *Journal of Clinical Epidemiology*; 57(10): 1096-1103.

9. Schrijvers, C.T.M., Stronks, K., van de Mheen, D.H., Coebergh, J.W. & Mackenbach, J.P. (1994), 'Validation of Cancer Prevalence Data from a Postal Survey by Comparison with Cancer Registry Records', *American Journal of Epidemiology*; 139(4): 408-414.

10. Wolinsky, F.D., Jones, M.P., Ullrich, F., Lou, Y. & Wehby, G.L. (2014), 'The Concordance of Survey Reports and Medicare Claims in a Nationally Representative Longitudinal Cohort of Older Adults', *Medical Care*; 52(5): 462-468.

11. Wolinsky, F.D., Jones, M.P., Ullrich, F., Lou, Y. & Wehby, G.L. (2015), 'Cognitive Function and the Concordance Between Survey Reports and Medicare Claims in a Nationally Representative Cohort of Older Adults', *Medical Care*; 53(5): 455-462.

Table 1: Reporting of false negatives and false negatives, by condition

Self-reporting:	Onset of condition in the last two years			Onset of condition in the last two years		
Hospital data:	Min. one hospital visit in the last two years			Min. one hospital visit in the last two years		
Mapping to conditions:	All IP diagnosis codes			Primary IP diagnosis code only		
	Mean incidence	False negatives	False positives	Mean incidence	False negatives	False positives
	(1)	(2)	(3)	(4)	(5)	(6)
Cancer	0.027 (0.161)	58.72	1.03	0.024 (0.153)	56.59	1.09
Stroke	0.009 (0.096)	53.98	1.02	0.006 (0.075)	37.87	1.10
Heart Attack	0.022 (0.148)	67.14	0.67	0.006 (0.075)	22.99	0.97

Note: Mean incidence in column 1 shows the fraction of the population who had an inpatient admission in the previous two years with either the primary or secondary diagnosis recorded as the relevant condition; mean incidence in column 4 shows the fraction with an inpatient admission in the past two years with the relevant primary diagnosis code. Standard deviation is shown in parentheses below. A false negative is defined as when a person is recorded in hospital as having a condition, but fails to report it in ELSA. A false positive is defined as when a person is not recorded in hospital as having a condition, but does report it in ELSA. IP denotes inpatient. In all columns, an individual is defined as self-reporting a condition if they report the onset of the condition in the last two years. Figures reported in columns 2, 3, 5 and 6 are percentages.

Table 2: Reporting under alternative definitions (2)

Self-reporting:	Onset of condition in the previous two years								
Hospital data:	Min. one hospital visit in last two years			Min. two hospital visits in last two years		Min. three hospital visits in last two years		Min. four hospital visits in last two years	
Mapping to conditions:	All IP diagnosis codes			All IP diagnosis codes		All IP diagnosis codes		All IP diagnosis codes	
	Mean incidence	False negatives	False positives	False negatives	False positives	False negatives	False positives	False negatives	False positives
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Cancer	0.027 (0.161)	58.72	1.03	50.43	1.54	48.63	1.75	46.74	1.84
Stroke	0.009 (0.096)	53.98	1.02	54.84	1.34	51.28	1.42	54.55	1.45
Heart Attack	0.022 (0.148)	67.14	0.67	60.62	1.00	62.21	1.23	61.21	1.31

Note: Mean incidence shows the fraction of the population who had an inpatient admission with either the primary or secondary diagnosis recorded as the relevant condition. Standard deviation is shown in parentheses below. A false negative is defined as when a person is recorded in hospital as having a condition, but fails to report it in ELSA. A false positive is defined as when a person is not recorded in hospital as having a condition, but does report it in ELSA. IP and OP denote inpatient and outpatient, respectively. In all columns, an individual is defined as self-reporting a condition if they report the onset of the condition in the last two years. Figures reported in columns 2-9 are percentages.

Table 3: Reporting under alternative definitions (3)

Self-reporting:	Onset of condition in the previous two years					Ever diagnosed with the condition		Treated for cancer in past two years	
Hospital data:	Min. one hospital visit in last two years			Min. two hospital visits in last two years		Min. one hospital visit in last two years		Min. one hospital visit in last two years	
Mapping to conditions:	Primary IP diagnosis code			Primary IP diagnosis code		All IP diagnosis codes		All IP diagnosis codes	
	Mean incidence	False negatives	False positives	False negatives	False positives	False negatives	False positives	False negatives	False positives
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Cancer	0.024 (0.153)	56.59	1.09	48.97	1.63	31.41	8.88	47.76	1.68
Stroke	0.006 (0.075)	37.87	1.10	31.15	1.37	33.91	3.99	-	-
Heart Attack	0.006 (0.075)	22.99	0.97	18.18	1.31	23.37	3.97	-	-

Note: Mean incidence shows the fraction of the population who had an inpatient admission with the primary diagnosis recorded as the relevant condition. Standard deviation is shown in parentheses below. A false negative is defined as when a person is recorded in hospital as having a condition, but fails to report it in ELSA. A false positive is defined as when a person is not recorded in hospital as having a condition, but does report it in ELSA. IP and OP denote inpatient and outpatient, respectively. Figures reported in columns 2-9 are percentages.

Table 4: Reporting under alternative definitions (1)

Self-reporting:	Onset of condition in the previous two years								
Hospital data:	Min. one hospital visit in last two years							Min. one hospital visits in last three years	
Mapping to conditions:	All IP diagnosis codes			All IP diagnosis codes & IP consultant specialties		All IP diagnosis codes, IP consultant specialties & OP consultant specialties		All IP diagnosis codes	
	Mean incidence	False negatives	False positives	False negatives	False positives	False negatives	False positives	False negatives	False positives
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Cancer	0.027 (0.161)	58.72	1.03	58.50	0.99	68.84	0.72	67.50	0.99
Stroke	0.009 (0.096)	53.98	1.02	59.23	1.01	84.34	0.88	60.36	0.95
Heart Attack	0.022 (0.148)	67.14	0.67	79.23	0.53	87.14	0.54	71.45	0.60

Note: Mean incidence shows the fraction of the population who had an inpatient admission with either the primary or secondary diagnosis recorded as the relevant condition. Standard deviation is shown in parentheses below. A false negative is defined as when a person is recorded in hospital as having a condition, but fails to report it in ELSA. A false positive is defined as when a person is not recorded in hospital as having a condition, but does report it in ELSA. IP and OP denote inpatient and outpatient, respectively. In all columns, an individual is defined as self-reporting a condition if they report the onset of the condition in the last two years. Figures reported in columns 2-9 are percentages.

Table 5: Relationship between individual characteristics and probability of reporting a false negative

	False negative reported for:							
	Any condition		Cancer		Stroke		Heart Attack	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High cognitive function	0.676***	0.818**	0.608***	0.735**	0.798	0.822	0.854	0.954
	(0.061)	(0.078)	(0.074)	(0.097)	(0.186)	(0.210)	(0.138)	(0.158)
Age-band (ref: 50-54):								
55-59		0.727		0.654		1.269		0.796
		(0.346)		(0.388)		(1.373)		(0.912)
60-64		1.041		0.607		1.962		2.471
		(0.487)		(0.349)		(2.020)		(2.820)
65-69		1.385		1.045		1.153		2.724
		(0.639)		(0.592)		(1.153)		(3.085)
70-74		1.357		0.890		1.164		3.207
		(0.626)		(0.503)		(1.164)		(3.633)
75-79		1.715		1.003		1.063		4.143
		(0.793)		(0.568)		(1.067)		(4.694)
80-84		2.064		1.891		1.252		3.513
		(0.966)		(1.102)		(1.268)		(3.979)
85+		1.738		1.437		1.259		3.083
		(0.841)		(0.890)		(1.317)		(3.525)
Female		0.663***		0.552***		0.843		0.978
		(0.060)		(0.070)		(0.176)		(0.153)
In a couple		0.989		0.975		0.912		1.030
		(0.099)		(0.140)		(0.206)		(0.162)
Highly educated		1.058		1.212		0.907		0.945
		(0.106)		(0.166)		(0.223)		(0.166)
Observations	2,606	2,606	1,277	1,277	448	448	1,045	1,045

Note: All results reported are odds ratios from a logistic regression model. Robust standard errors shown in parentheses are clustered at the individual level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. A false negative is defined as when a person is recorded in hospital as having a condition, but fails to report it in ELSA. High cognitive function defined as being in the top half of the distribution for orientation and retrospective memory (word recall) score. Highly educated is defined as having at least some higher education, or equivalent.

Table 6: Relationship between individual characteristics and probability of reporting a false positive

	False positive reported for:							
	Any condition		Cancer		Stroke		Heart Attack	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High cognitive function	0.533***	0.753***	0.750***	0.997	0.321***	0.484***	0.617***	0.884
	(0.038)	(0.057)	(0.080)	(0.115)	(0.039)	(0.064)	(0.086)	(0.130)
Age-band (ref: 50-54):								
55-59		0.908		1.076		1.095		0.683
		(0.228)		(0.446)		(0.639)		(0.247)
60-64		1.173		1.360		1.742		0.791
		(0.287)		(0.539)		(1.004)		(0.284)
65-69		1.612**		2.021*		2.270		1.041
		(0.391)		(0.789)		(1.307)		(0.378)
70-74		1.814**		2.251**		2.896*		0.996
		(0.441)		(0.881)		(1.668)		(0.365)
75-79		2.725***		2.739**		4.704***		1.810*
		(0.662)		(1.076)		(2.711)		(0.643)
80-84		3.472***		3.205***		6.860***		1.890*
		(0.858)		(1.291)		(3.972)		(0.715)
85+		3.222***		2.735**		6.449***		1.950
		(0.889)		(1.260)		(3.915)		(0.871)
Female		0.586***		0.484***		0.877		0.398***
		(0.044)		(0.052)		(0.111)		(0.061)
In a couple		0.999		1.062		1.146		0.773*
		(0.079)		(0.123)		(0.158)		(0.119)
Highly educated		0.858*		0.933		0.875		0.687**
		(0.069)		(0.111)		(0.121)		(0.120)
Observations	45,832	45,832	44,557	44,557	45,386	45,386	44,785	44,785

Note: All results reported are odds ratios from a logistic regression model. Robust standard errors shown in parentheses are clustered at the individual level. *** p < 0.01, ** p < 0.05, * p < 0.1. A false negative is defined as when a person is recorded in hospital as having a condition, but fails to report it in ELSA. High cognitive function defined as being in the top half of the distribution for orientation and retrospective memory (word recall) score. Highly educated is defined as having at least some higher education, or equivalent.

A Appendix

Table A1: Relationship between ELSA diagnosis and HES diagnosis codes

ELSA diagnosis	ICD-10 codes	Outpatient specialty
Heart attack	I21, I22, I23, I24, I252, I258	Cardiology
Stroke	G450, G451, G452, G454, G458, G459, G46, I6	Neuorology
Cancer	C0, C1, C2, C3, C4, C5, C6, C7, C8, C91, C92, C93, C94, C95, C96, C97	Oncology

Notes: (1) Codes are taken from 'Understanding HSMRs: A toolkit on Hospital Standardised Mortality Ratios', De. Foster Intelligence, Version 9 (July 2014); (2) The ELSA documentation specifies that heart attack also includes coronary thrombosis (ICD10 code I24).