

Agreement between self-reports, proxy-reports and the National Patient Register regarding diagnoses of cardiovascular disorders and diabetes mellitus in a population-based sample of 80-year-olds

LINA RYDÉN¹, ROBERT SIGSTRÖM¹, JOHAN NILSSON¹, VALTER SUNDH¹, HANNA FALK ERHAG¹, SILKE KERN¹, MARGDA WAERN¹, SVANTE ÖSTLING¹, KATARINA WILHELMSON^{2,3}, INGMAR SKOOG¹

¹Neuropsychiatric Epidemiology Unit, Department of Psychiatry and Neurochemistry, Institute of Neuroscience and Physiology, Sahlgrenska Academy, Centre for Ageing and Health (AgeCap) at the University of Gothenburg, Sweden

²Department of Health and Rehabilitation, Institute of Neuroscience and Physiology, Sahlgrenska Academy, Centre for Ageing and Health (AgeCap) at the University of Gothenburg, Sweden

³Department of Geriatrics, Sahlgrenska University Hospital, Gothenburg, Sweden

Address correspondence to: Lina Rydén. Tel: +46 (0) 31 3438646; Email: lina.ryden@gu.se

Abstract

Background: cognitive impairment is common among older adults, necessitating the use of collateral sources in epidemiological studies involving this age group. The objective of this study was to evaluate agreement between self- and proxy-reports of cardiovascular disorders and diabetes mellitus in a population-based sample of 80-year-olds. Further, both self- and proxy-reports were compared with hospital register data.

Methods: data were obtained from the Gothenburg H70 Birth Cohort Studies in Sweden. The study had a cross-sectional design and information was collected through semi-structured interviews in 2009–2012 from participants born in 1930 ($N = 419$) and their proxy informants. The National Patient Register provided diagnoses registered during hospital stays. Agreement was measured with Kappa values (K).

Results: agreement between self- and proxy-reports was substantial for diabetes mellitus ($K = 0.79$), atrial fibrillation ($K = 0.61$), myocardial infarction ($K = 0.75$), angina pectoris ($K = 0.73$) and hypertension ($K = 0.62$), and fair for intermittent claudication ($K = 0.38$) and heart failure ($K = 0.40$). Compared to the National Patient Register, a large proportion of those with a hospital discharge diagnosis were also self- and proxy-reported.

Conclusions: proxy informants can be an important source of information, at least for well-defined conditions such as myocardial infarction, angina pectoris and diabetes mellitus.

Keywords

Dementia, Proxy-informants, Cardiovascular disease, Ageing, Agreement, Older people

Key points

- Agreement between proxy- and self-reports varied considerably between different cardiovascular disorders.
- Highest agreement was found for myocardial infarction, angina pectoris and diabetes mellitus.
- Lowest agreement was found for heart failure and intermittent claudication.
- A large proportion of those with a hospital discharge diagnosis was detected by both proxy- and self-reports.
- Proxy-reports can be an important source of information in population studies of older adults.

Introduction

The prevalence of memory problems increases with age, and dementia affects almost a fourth of the population at age 85 [1]. Yet it is important to include people with cognitive impairment in population-based studies to attain representative samples. Proxy informants, e.g. relatives, friends or medical staff can provide additional information when memory problems interfere with the individual's ability to report their medical history.

Cardiovascular disorders (CVD) and diabetes are common disorders that increase with age. Ischaemic heart diseases were the leading cause of death worldwide in 2017 and hypertension and hyperglycaemia among the top three risk factors for early death and disability, highlighting the importance to study these disorders [2]. However, few studies [3–5] have evaluated information from proxy informants regarding CVD and diabetes in population-based samples of older adults even though proxy-reports in these age groups often are needed to reduce missing information and selection-bias [6]. To our knowledge, only one study has compared self- and proxy-reports for CVD and diabetes among the very old (80+), where cognitive impairment is common. Also, degrees of agreement between self- and proxy-reports varies depending on the disorders investigated both in younger and older cohorts [3, 4, 7], necessitating studies on specific disorders.

Previous studies on older adults (65+) have found an almost perfect agreement for diabetes [3, 4] and substantial agreement for myocardial infarction, angina, and hypertension [3, 4]. The Hispanic Established Population for the Epidemiological Study of the Elderly (HEPESE) compared self- and proxy-reports among the very old (80+), and found an almost perfect agreement for diabetes, substantial agreement for myocardial infarction, and slight agreement for hypertension [5]. When comparing self-reports with medical records and registers in community or population-based samples, agreement also varies between disorders [8–15]. While agreement is reported to be substantial to almost perfect for diabetes in both younger and older age groups [8–13], it tends to be lower (moderate to substantial) for hypertension [9, 10, 12–15]. For claudication, agreement has been fair [9]. A wide range of agreement [8, 10, 11] is reported when treating heart disease as a single entity. Studies examining specific cardiac diseases found moderate to substantial agreement for myocardial infarction [9, 12, 14], slight to moderate for heart failure [12–14], and substantial for angina [9].

The objective of this study was to evaluate agreement between self- and proxy-reports of cardiovascular disorders and diabetes mellitus in a population-based sample of 80-year-olds. In addition, we tested validity of proxy-reports using self-reports as gold standard. Further, both self- and proxy-reports were compared with hospital register data.

Methods

Sample

Data from self- and proxy-reports were obtained from the Gothenburg H70 Birth Cohort Studies in Sweden, which has been described previously [16, 17]. In short, the H70-study started in 1971, where 70-year-olds living in Gothenburg were selected from the Swedish population register based on birth date and followed-up over time. Since then, new cohorts of older adults have been included. The present study has a cross-sectional design and participants were born in 1930 on dates 2, 3, 5, 6, 11, 12, 16, 18, 20, 21, 24, 27 or 30 of each month and examined at age 79–81. Data from participants and proxy informants were collected in 2009–2012. The study population (see flow chart in Appendix 1), comprised 662 participants (62% response rate), and 475 of these (72%) had data from both self- and proxy-reports on at least one of the seven disorders investigated (i.e. myocardial infarction, angina, claudication, heart failure, atrial fibrillation, hypertension, and diabetes). We excluded participants with dementia ($n = 42$). In some cases, the participant interview was augmented with additional information from medical records, relatives or other sources ($n = 14$). These were also excluded, leaving 419 participants with proxy informant reports.

The study was approved by the Regional Ethics Committee for Medical Research at the University of Gothenburg. Informed consent was obtained from all participants, their relatives or both.

Procedures

Participants were interviewed by research nurses or physicians at the memory clinic at Sahlgrenska University Hospital in Gothenburg. If declining examination at the outpatient clinic, home visits were offered. Information on diabetes and CVD was collected through semi-structured interviews. The questions were phrased 'Has a doctor or nurse ever told that you have or have had ...?' Tests of mental functioning (e.g. memory, language, visuospatial and executive abilities, and apraxia), the Mini Mental State Examination [18], and the Alzheimer's Disease Assessment Scale [19] were included.

Participants provided contact information for proxy informants, who were interviewed by telephone by health professionals. The semi-structured interviews included questions about changes in behaviour, cognitive function, psychiatric symptoms, and activities of daily living. Also questions about diabetes and CVD were included. These questions were phrased 'Has he/she or has he/she had any of the following disorders...?'

The diagnosis of dementia, based on information from cognitive tests and proxy-reports, according to the Diagnostic and Statistical Manual of Mental Disorders,

third edition revised (DSM-III-R), was only used as exclusion criteria [20].

Hospital discharge diagnoses were obtained from the National Patient Register (NPR), using codes from ICD-8, ICD-9, and ICD-10. All Swedish inhabitants have access to health care in Sweden and therefore equal chances to be in the NPR. The NPR was initiated in 1964 and attained full national coverage in 1978 [21].

Statistics

Prevalence figures regarding specific disorders were reported for self-reports, proxy-reports, and the NPR based on persons with both self- and proxy-reports. Differences in prevalence for self-report compared to proxy-report and the NPR were analysed with McNemar’s test. To measure agreement between self- and proxy-reports, overall agreement, $P_o = (a + d)/n$, as well as positive agreement, which measures the probability that one rater will classify the disease as present if the other rater has also done so, $P_{pos} = 2a/((a + c)+(a + b))$, and negative agreement, which measures the probability that one rater will classify the disease as absent if the other rater has also done so, $P_{neg} = 2d/((b + d)+(c + d))$, were calculated (Appendix 2). Also, Cohen’s Kappa (*K*) was calculated, which measures agreement beyond chance [22]. Kappa values were interpreted as slight between 0 and 0.20, fair between 0.21 and 0.40, moderate between 0.41 and 0.60, substantial between 0.61 and 0.80, and almost perfect between 0.81 and 1.0 [23].

Self-reports are often used in epidemiological studies [6] and to test validity for proxy-reports compared to self-reports, sensitivity, specificity, positive (PPV) and negative (NPV) predicted values were calculated, using self-reported information as gold standard. Since self-report is not a perfect information source, we compared self- and proxy-reports to a third source of information, the NPR, using Kappa statistics. Also, sensitivity was calculated to identify the proportion of diagnoses in the NPR that was identified by self- and proxy-reports. Statistical analyses were performed with R 3.4.2.

Results

Among participants (*N* = 419), mean age was 80.2 years (SD 0.2), 63% (*n* = 262) were women, half had education beyond compulsory (i.e. more than 7 years), and mean MMSE score was 28.5 (SD 1.4). Among proxy-informants, 71% (*n* = 298) were women and 35% (*n* = 136) were living together with the participant. Regarding the proxy informant’s relationship to the participant, 35% (*n* = 146) were partners, 52% (*n* = 217) were children, 6% (*n* = 25) were siblings, and 7% (*n* = 31) had another type of relationship (Appendix 3).

The prevalence of CVD and diabetes according to different information sources is given in Table 1. Self- and proxy-reports showed similar prevalence figures for all conditions, except hypertension, which was more common according to self-reports. The NPR yielded lower prevalence figures in comparison to self-reports for all conditions, except myocardial infarction and angina, where prevalence figures were similar.

Table 2 shows agreement between self- and proxy-reports. Overall and negative agreement were 90% or above for all disorders except hypertension ($P_o = 81\%$ and $P_{neg} = 78\%$). Positive agreement was lower than negative agreement for all conditions except hypertension (83%), and varied substantially between disorders (41–83%), with the lowest level of positive agreement for claudication and heart failure. Kappa values were substantial for myocardial infarction (*K* = 0.75), angina (*K* = 0.73), atrial fibrillation (*K* = 0.61), hypertension (*K* = 0.62) and diabetes (*K* = 0.79) and fair for claudication (*K* = 0.38) and heart failure (*K* = 0.40).

Table 3 shows sensitivity, specificity, PPV, and NPV for proxy-reports compared to self-reports. Proxy-reports had high specificity and NPV for all conditions, except NPV for hypertension (67%). Sensitivity and PPV varied considerably among disorders. The highest sensitivity (i.e. the proportion of self-reported conditions detected by proxy informants) was found for diabetes (78%) and myocardial infarction and hypertension (73%), and the lowest was found for heart failure (45%) and claudication (35%). The highest PPV (i.e. the proportion of proxy-reports that was

Table 1. Prevalence of specified disorders in a population-based sample of older adults, by information source

| | Information sources | | | | |
|---|-----------------------------|------------------------------|------------------------------|--------------------|------------------------------|
| | Self-reports % (<i>n</i>) | Proxy-reports % (<i>n</i>) | <i>P</i> -value ^a | NPR % (<i>n</i>) | <i>P</i> -value ^b |
| Myocardial infarction (<i>N</i> = 413) | 9.0 (37) | 8.0 (33) | 0.453 | 7.5 (31) | 0.211 |
| Angina pectoris (<i>N</i> = 398) | 13.8 (55) | 11.3 (45) | 0.066 | 12.6 (50) | 0.499 |
| Intermittent claudication (<i>N</i> = 408) | 4.2 (17) | 3.2 (13) | 0.480 | 0.2 (1) | <0.001 |
| Heart failure (<i>N</i> = 394) | 7.9 (31) | 8.1 (32) | 1.000 | 3.0 (12) | <0.001 |
| Atrial fibrillation (<i>N</i> = 400) | 15.5 (62) | 13.8 (55) | 0.337 | 12.0 (48) | 0.006 |
| Hypertension (<i>N</i> = 386) | 63.5 (245) | 49.0 (189) | <0.001 | 31.3 (121) | <0.001 |
| Diabetes mellitus (<i>N</i> = 407) | 11.3 (46) | 10.6 (43) | 0.628 | 7.1 (29) | 0.003 |

^aDifferences in prevalence for proxy-reports compared to self-reports, analysed with McNemar’s test.

^bDifferences in prevalence for the National Patient Register (NPR) compared to self-reports, analysed with McNemar’s test.

Table 2. Measures of agreement between self- and proxy-reports

| | P _o ^a % (95% CI) | P _{pos} ^b % (95% CI) | P _{neg} ^c % (95% CI) | Kappa (95% CI) |
|-------------------------------------|--|--|--|------------------|
| Myocardial infarction (N = 413) | 96 (94–98) | 77 (66–88) | 98 (97–99) | 0.75 (0.63–0.87) |
| Angina pectoris (N = 398) | 94 (91–96) | 76 (66–85) | 97 (95–98) | 0.73 (0.62–0.83) |
| Intermittent claudication (N = 408) | 96 (93–97) | 41 (19–64) | 98 (97–99) | 0.38 (0.15–0.60) |
| Heart failure (N = 394) | 91 (88–94) | 45 (30–61) | 95 (94–97) | 0.40 (0.23–0.56) |
| Atrial fibrillation (N = 400) | 90 (87–93) | 67 (57–77) | 94 (92–96) | 0.61 (0.50–0.72) |
| Hypertension (N = 386) | 81 (77–85) | 83 (79–87) | 78 (73–83) | 0.62 (0.54–0.69) |
| Diabetes mellitus (N = 407) | 96 (93–98) | 80 (71–89) | 98 (97–99) | 0.79 (0.69–0.88) |

^aOverall agreement.^bPositive agreement.^cNegative agreement.**Table 3.** Sensitivity, specificity, positive predicted value (PPV), and negative predicted value (NPV) for proxy informants, using self-reports as gold standard

| | Sensitivity % (95% CI) | Specificity % (95% CI) | PPV % (95% CI) | NPV % (95% CI) |
|-------------------------------------|------------------------|------------------------|----------------|----------------|
| Myocardial infarction (N = 413) | 73 (56–86) | 98 (97–99) | 82 (65–93) | 97 (95–99) |
| Angina pectoris (N = 398) | 69 (55–81) | 98 (96–99) | 84 (71–94) | 95 (92–97) |
| Intermittent claudication (N = 408) | 35 (14–62) | 98 (96–99) | 46 (19–75) | 97 (95–99) |
| Heart failure (N = 394) | 45 (27–64) | 95 (92–97) | 44 (26–62) | 95 (93–97) |
| Atrial fibrillation (N = 400) | 63 (50–75) | 95 (92–97) | 71 (57–82) | 93 (90–96) |
| Hypertension (N = 386) | 73 (67–79) | 94 (88–97) | 95 (91–98) | 67 (60–74) |
| Diabetes mellitus (N = 407) | 78 (64–89) | 98 (96–99) | 84 (69–93) | 97 (95–99) |

confirmed by self-reports) was found for hypertension (95%), angina and diabetes (84%), and myocardial infarction (82%) and the lowest was found for claudication (46%) and heart failure (44%).

To calculate the proportion of hospital register diagnoses that were captured by self- and proxy-reports, sensitivity for self- and proxy-reports was calculated using the NPR as gold standard. Kappa values were calculated to compare self- and proxy-reports with the NPR (Appendix 4). Claudication was excluded due to few cases in the NPR. Sensitivity was 70% or more for all disorders regarding self-reports except for heart failure (67%). Regarding proxy-reports, sensitivity was over 70% for all disorders, except angina (64%) and atrial fibrillation (69%). Thus, a large proportion of the participants and proxy informants reported the condition if present in the NPR. Atrial fibrillation showed the largest difference regarding sensitivity between self- and proxy-reports compared to the NPR, with higher sensitivity for self-reports than proxy-reports. Regarding agreement for self- and proxy-reports compared to the NPR, the only disorder in this study where all patients are expected to receive inpatient treatment is myocardial infarction. The Kappa value for myocardial infarction was substantial for both self-reports ($K = 0.74$) and proxy-reports ($K = 0.73$) compared to the NPR.

Discussion

In a population sample of 80-year-olds, we found substantial agreement between self- and proxy-reports for diabetes,

atrial fibrillation, myocardial infarction, angina, and hypertension, but only fair agreement for claudication and heart failure. Our findings suggest that proxy informants can provide reliable information, at least for well-defined disorders.

The Kappa value for hypertension was 0.62, which is in line with most studies on older adults ($K = 0.64$ – 0.67) [3, 4], except the HEPSE study in the very old where Kappa for hypertension was 0.07 [5]. For diabetes, agreement was higher ($K = 0.79$), which is consistent with previous studies ($K = 0.86$ – 0.94) [3, 4]. Also, agreement for myocardial infarction ($K = 0.75$) and angina ($K = 0.73$) were consistent with previous studies ($K = 0.67$ and $K = 0.72$). Heart diseases are sometimes treated as one group, [4, 24] which could be a problem according to our results, since agreement differed substantially between cardiac diagnoses. However, only studying Kappa for agreement has been criticised since it is hard to interpret [25–27], e.g. the prevalence of disorders impacts Kappa, resulting in a paradox where overall agreement is high but Kappa low at high and low prevalence figures [27, 28]. Our prevalence figures for myocardial infarction, diabetes, and angina were similar to other studies in older adults, except HEPSE reporting higher prevalence of diabetes. However, our prevalence figure for hypertension (63.5%) differed from other studies that reported prevalence figures of 41.7% [4], 32.7% [3], and 98% [5]. The low Kappa for hypertension in the HEPSE study might be explained by the high prevalence (98%), since they found high overall agreement (94%) despite low Kappa ($K = 0.07$) [5].

In our study, myocardial infarction, angina, and diabetes had high overall (94–96%), positive (76–80%), and negative

agreement (97–98%), as well as substantial Kappa values (0.73–0.79). Atrial fibrillation follows the same pattern, with somewhat lower number for the different indices. Claudication and heart failure on the other hand, had fair Kappa values despite high overall agreement. The disagreement for claudication and heart failure was found among the positive answers. To further analyse the disagreement among the positive answers, we calculated the proportion of self-reports that was identified by proxy informants, i.e. sensitivity, and the proportion of proxy-reports that were confirmed by self-reports, i.e. PPV. Both sensitivity and PPV were lower than 50% for claudication and heart failure, suggesting that proxy-reports are poor predictors for self-reports regarding these disorders. One reason for this could be low knowledge about these disorders in the community. It has for example been shown that awareness in the general population is lower for heart failure than for other heart conditions [29]. Myocardial infarction, hypertension and diabetes were the self-reported disorders most often detected by proxy informants, with a sensitivity of 73–78%, meaning that one fourth of the cases were missed. PPV was high for these disorders, meaning that 82–95% of the cases reported by proxy informants were confirmed by self-reports. The finding that negative agreement was high is expected in population-based studies where prevalence figures are relatively low. Hypertension differed from the other disorders with its higher prevalence and lower negative agreement.

In our study, it was possible to compare both self- and proxy-reports with a third source, the NPR, which has high numbers of correct diagnoses, with PPV between 85 and 95% for most diagnoses compared to external sources, mainly medical records [21]. On the other hand, sensitivity for the NPR compared to external sources varied substantially between disorders, e.g. high sensitivity for myocardial infarction and low sensitivity for hypertension [21]. Myocardial infarction was the only disorder in our study where all patients were expected to be treated as inpatients. Agreement for myocardial infarction was substantial for both self- and proxy-reports compared to the NPR. Regarding the other disorders studied, we take note of the fact that a large proportion of those with hospital diagnoses were detected by self- and proxy-reports.

Other researchers report both under- and over-reporting by proxy informants in relation to self-reports [3, 4, 30]. In our study, only the prevalence of hypertension was lower for proxy-reports compared to self-reports. The NPR had lower prevalence for all conditions, except myocardial infarction and angina, compared to self-reports, illustrating that this source often underestimates the occurrence of disorders.

Strengths of the study include the population-based sample, that interview questions were posed by health professionals, and the possibility to differentiate between specific cardiac disorders. The possibility to compare self- and proxy-reports with the NPR is a further strength. There are also limitations. First, only participants with proxy-reports are included, which may hamper representativeness. Second, it was not possible to compare medical severity.

Third, there are problems with the NPR such as diagnostic and coding errors and selection biases, e.g. persons not seeking medical care are not captured in the register. Many of the disorders investigated are treated in primary care and therefore not included in the NPR [21]. Fourth, the study only included 79 to 81-year-olds living in Gothenburg which may limit generalisability to other age groups and settings.

In conclusion, proxy-reports can be an important source of information in epidemiological studies where self-reports are difficult to obtain, at least for well-defined disorders such as myocardial infarction, angina and diabetes. The use of proxy informants can reduce the amount of missing information and increases representativeness in studies. However, even though agreement was high for these disorders, still 20–30% of the self-reported diagnoses were missed, indicating that caution is needed when using proxy-informants in clinical settings where adequate reporting of disorders are essential.

Supplementary data mentioned in the text are available to subscribers in *Age and Ageing* online.

Declaration of Sources of Funding: This work was supported by Swedish Research Council [grant numbers 11267, 2005-8460, 2007-7462, 2012-5041, 2015-02830, 2013-8717]; Swedish Research Council for Health, Working Life and Welfare [grant numbers 2006-0020, 2008-1229, 2012-1138, 2006-0596, 2008-1111, 2010-0870, 2013-1202, and 2013-2300 and 2013-2496 to Centre for Aging and Health]; Swedish Brain Power; Alzheimerfonden; The Swedish Brain Foundation; Konung Gustaf V:s och Drottning Victorias Frimurarestiftelse; Sahlgrenska University Hospital (ALF); The Alzheimer's Association Zenith Award [grant number ZEN-01-3151]; The Alzheimer's Association Stephanie B. Overstreet Scholars [grant number IIRG-00-2159]; Eivind och Elsa K:son Sylvans stiftelse; Stiftelsen Söderström-Königska Sjukhemmet; Stiftelsen för Gamla Tjänarinnor; Handlanden Hjalmar Svenssons Forskningsfond; and Stiftelsen Professor Bror Gadelius Minnesfond. No funding sources had any role in study design, in the collection, analysis and interpretation of data, in the writing of the report or in the decision to submit the article for publication.

Declaration of Conflict of Interest: Dr. Skoog reports grants from Swedish Research Council, grants from Swedish Research Council for Health, Working Life and Welfare, during the conduct of the study; personal fees from Takeda, outside the submitted work.

References

1. Skoog I, Borjesson-Hanson A, Kern S *et al.* Decreasing prevalence of dementia in 85-year olds examined 22 years apart: the influence of education and stroke. *Sci Rep* 2017; 7: 6136.

2. Findings from the Global Burden of Disease Study 2017. Seattle, WA: IHME, 2018.
3. The Medical Research Council Cognitive Function and Ageing Study. Survey into health problems of elderly people: a comparison of self-report with proxy information. *Int J Epidemiol* 2000; 29: 684–97.
4. Magaziner J, Bassett SS, Hebel JR, Gruber-Baldini A. Use of proxies to measure health and functional status in epidemiologic studies of community-dwelling women aged 65 years and older. *Am J Epidemiol* 1996; 143: 283–92.
5. Siordia C. Reliability of self-report on basic health conditions in Mexican Americans Age 80 and older. *Maturitas* 2013; 74: 95–8.
6. Gruber-Baldini AL, Shardell M, Lloyd KD, Magaziner J. Use of proxies and informants. In: Newman A, Cauley J, eds. *The Epidemiology of Aging*. Dordrecht: Springer, 2012; 81–90.
7. Halabi S, Zurayk H, Awaida R, Darwish M, Saab B. Reliability and validity of self and proxy reporting of morbidity data: a case study from Beirut, Lebanon. *Int J Epidemiol* 1992; 21: 607–12.
8. Galenkamp H, Huisman M, Braam AW, Schellevis FG, Deeg DJ. Disease prevalence based on older people's self-reports increased, but patient-general practitioner agreement remained stable, 1992-2009. *J Clin Epidemiol* 2014; 67: 773–80.
9. Haapanen N, Miilunpalo S, Pasanen M, Oja P, Vuori I. Agreement between questionnaire data and medical records of chronic diseases in middle-aged and elderly Finnish men and women. *Am J Epidemiol* 1997; 145: 762–9.
10. Koller KR, Wilson AS, Asay ED, Metzger JS, Neal DE. Agreement between self-report and medical record prevalence of 16 chronic conditions in the Alaska EARTH Study. *J Prim Care Community Health* 2014; 5: 160–5.
11. Kriegsman DM, Penninx BW, van Eijk JT, Boeke AJ, Deeg DJ. Self-reports and general practitioner information on the presence of chronic diseases in community dwelling elderly. A study on the accuracy of patients' self-reports and on determinants of inaccuracy. *J Clin Epidemiol* 1996; 49: 1407–17.
12. Muggah E, Graves E, Bennett C, Manuel DG. Ascertainment of chronic diseases using population health data: a comparison of health administrative data and patient self-report. *BMC Public Health* 2013; 13: 16.
13. Okura Y, Urban LH, Mahoney DW, Jacobsen SJ, Rodeheffer RJ. Agreement between self-report questionnaires and medical record data was substantial for diabetes, hypertension, myocardial infarction and stroke but not for heart failure. *J Clin Epidemiol* 2004; 57: 1096–103.
14. Teh R, Doughty R, Connolly M *et al.* Agreement between self-reports and medical records of cardiovascular disease in octogenarians. *J Clin Epidemiol* 2013; 66: 1135–43.
15. Tormo MJ, Navarro C, Chirlaque MD, Barber X. Validation of self diagnosis of high blood pressure in a sample of the Spanish EPIC cohort: overall agreement and predictive values. EPIC Group of Spain. *J Epidemiol Community Health* 2000; 54: 221–6.
16. Joas E, Guo X, Kern S, Ostling S, Skoog I. Sex differences in time trends of blood pressure among Swedish septuagenarians examined three decades apart: a longitudinal population study. *J Hypertens* 2017; 35: 1424–31.
17. Rydberg Sterner T, Ahlner F, Blennow K *et al.* The Gothenburg H70 Birth cohort study 2014-16: design, methods and study population. *Eur J Epidemiol* 2019; 34: 191–209.
18. Folstein MF, Folstein SE, McHugh PR. 'Mini-mental state'. A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975; 12: 189–98.
19. Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. *Am J Psychiatry* 1984; 141: 1356–64.
20. Skoog I, Nilsson L, Palmertz B, Andreasson LA, Svanborg A. A population-based study of dementia in 85-year-olds. *N Engl J Med* 1993; 328: 153–8.
21. Ludvigsson JF, Andersson E, Ekblom A *et al.* External review and validation of the Swedish national inpatient register. *BMC Public Health* 2011; 11: 450.
22. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960; 20: 37–46.
23. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159–74.
24. Long K, Sudha S, Mutran EJ. Elder-proxy agreement concerning the functional status and medical history of the older person: the impact of caregiver burden and depressive symptomatology. *J Am Geriatr Soc* 1998; 46: 1103–11.
25. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *J Clin Epidemiol* 1993; 46: 423–9.
26. Cicchetti DV, Feinstein AR. High agreement but low kappa: II. Resolving the paradoxes. *J Clin Epidemiol* 1990; 43: 551–8.
27. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther* 2005; 85: 257–68.
28. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol* 1990; 43: 543–9.
29. Remme WJ, McMurray JJ, Rauch B *et al.* Public awareness of heart failure in Europe: first results from SHAPE. *Eur Heart J* 2005; 26: 2413–21.
30. Shields M. Proxy reporting in the National Population Health Survey. *Health Rep* 2000; 12: 21–39. (Eng); 23-44 (Fre).

Received 29 June 2018; editorial decision 14 February 2019