ORIGINAL RESEARCH

Agreement Between Self-Reported Information and Health Register Data on Chronic Diseases in the Oldest Old

Pauliina Halonen [1–3], Esa Jämsen[4], Linda Enroth [1,2], Marja Jylhä[1–3]

[1]Faculty of Social Sciences (Health Sciences), Tampere University, Tampere, Finland; [2]Gerontology Research Center (GEREC), Tampere, Finland; [3]Tays Research Services, Wellbeing Services County of Pirkanmaa, Tampere University Hospital, Tampere, Finland; [4]Faculty of Medicine, University of Helsinki, Helsinki, Finland

Correspondence: Pauliina Halonen, Faculty of Social Sciences (Health Sciences), Tampere University, Arvo Ylpön katu 34, Tampere, 33520, Finland, Tel +358 50 4377338, Email Pauliina.Halonen@tuni.fi

Purpose: To study the agreement on disease prevalence between survey data and national health register data among people aged over 90.

Patients and Methods: The survey data were from the Vitality 90+ Study conducted among 1637 community dwellers and persons in long-term care aged 90 and over in Tampere, Finland. The survey was linked with two national health registers, including hospital discharge data and prescription information. The prevalence of 10 age-related chronic diseases was calculated for each data source and the agreement between the survey and the registers was estimated using Cohen's kappa statistics and positive and negative percent agreement.

Results: The prevalence of most diseases was higher in the survey than in the registers. The level of agreement was highest when the survey was compared with information combined from both registers. Agreement was almost perfect for Parkinson's disease (κ=0.81) and substantial for diabetes (κ=0.75) and dementia (κ=0.66). For heart disease, hypertension, stroke, cancer, osteoarthritis, depression, and hip fracture, the agreement ranged from fair to moderate.

Conclusion: Self-reported information on chronic diseases shows acceptable agreement with health register data to warrant the use of survey methods in population-based health studies among the oldest old. It is important to acknowledge the gaps in health registers when validating self-reported information against register data.

Keywords: chronic condition, reliability, survey, prevalence, health registers

Introduction

Self-report surveys are the most common and feasible method for acquiring information from large population-based samples in health research. However, electronic health registers have paved the way for the use of register information compiled by administrators or insurance providers. These registers typically cover either general practice records or inpatient care, but the variation is wide.[1] Thorough clinical and laboratory examinations, regarded as the gold standard, are seldom possible in large population studies.

In younger adults, self-report and health register data show high agreement for clearly defined diseases that require constant medication.[2,3] Yet studies have suggested that self-reports underestimate disease prevalence.[4–6] In the fast-growing group of the oldest old, multimorbidity and cognitive problems may compromise the reliability of self-reported health information.[7–9] Self-reports and register data on hypertension and heart disease have shown poor or moderate agreement in the few studies including the oldest old.[8,10] A study of Swedish people aged 80 found that the prevalence of cardiovascular diseases and diabetes was lower in a patient register than in interviews.[11] In an earlier analysis of community-dwelling Vitality 90+ Study participants in Finland, depression, dementia, and arthritis occurred more frequently in self-reports than in local medical records, while other diseases were underreported.[12]

The Nordic countries have publicly funded universal health-care systems, providing the basis for reliable and comprehensive health registers.[13] These registers are accurate for diseases that require continuous care and medication, such as diabetes, but potentially less so for risk factors (eg, high blood pressure, obesity, hyperlipidemia).[14] However, most registers are not designed and intended for research, but for administrative purposes.[1,13]

The number of the oldest old is continuing to grow worldwide[15] and therefore reliable and accurate information on morbidity is essential for informed policy-making and planning.[16] This study evaluates the reliability of self- and proxy-reported survey information on chronic diseases in the oldest old by comparing it with national health register data. We use two registers, one covering hospital treatment and the other prescription data. Our research questions are: (1) To what extent are self-reports consistent with register data on the prevalence of selected chronic diseases, and (2) How does the level of agreement vary between different diseases?

Materials and Methods

Survey Data

The Vitality 90+ Study is a multidisciplinary population-based research project focused on the health, functioning, and life circumstances of people aged 90 and over.[17] Its main component is a postal survey among both community-dwelling and long-term care residents in the city of Tampere, Finland (2019 population 238,140, of which 0.9% were over 90 years). This study uses data from a survey conducted in 2014 among people born in 1924 or earlier. Responses were received from 1637 participants, giving a response rate of 80%.

The survey included a question on chronic diseases. Respondents were asked, "Has a doctor told you that you have…?", followed by a list of 10 diseases: 1) hypertension, high blood pressure, 2) heart disease (coronary artery disease, arrhythmia, myocardial infarction), 3) dementia, Alzheimer's disease or worsening of memory, 4) diabetes 5) Parkinson's disease, 6) stroke, 7) osteoarthritis 8) hip fracture, 9) depression, depressed mood, and 10) cancer. Information was also collected on place of residence (home, long-term care facility) and the use of proxy respondents. Proxy respondents (287; 16 participants had missing information on proxy status) were used for those who were unable to respond themselves. Most of the proxy responses were given by a relative/acquaintance (76.7% of proxy responses) and the rest by a member of staff in health or social care (23.2%).

Register Data

Care Register for Health Care

The Care Register for Health Care (CRHC) compiled by the National Institute for Health and Welfare is an administrative register containing nationwide information on inpatient care, specialized outpatient care, and day surgery.[13] It provides one primary diagnosis for every care episode (ie, the main reason for the care episode) and possible secondary diagnoses (ie, other relevant diagnoses with implications for prognosis, treatment, or health status) with ICD-10 code. Reporting to the register is mandatory and the main diagnosis must be specified for every care episode.[18] A review assessing the validity of the register found that the main diagnoses were complete, but information on secondary diagnoses was often incomplete.[19] In our study, data from CRHC was obtained for 1563 survey participants; 74 participants declined to give permission for the linkage.

Finnish Prescription Register

The Finnish Prescription Register (FPR) maintained by the Social Insurance Institution of Finland is a nationwide register covering all prescription drugs dispensed at pharmacies. It identifies the Anatomical Therapeutic Chemical (ATC) code for each purchase and a reimbursement code, if applicable. In Finland, most prescription drugs are partly or fully reimbursed, depending on the indication. We used the reimbursement codes to identify hypertension, heart disease (including cardiac insufficiency, chronic coronary artery disease and related fat metabolism disorders, and chronic arrhythmias), Alzheimer's disease (AD), diabetes, and Parkinson's disease from the register. To be eligible for reimbursement, patients must undergo a thorough clinical examination. Hence, the code confirms the indication of the drug and is therefore a highly reliable indicator for a disease. FPR does not cover drugs administered during hospital or other institutional care, or over-the-counter drugs.[20] The validity of FPR concerning psychotropic drug use has been shown to

be adequate.[21,22] For our study, we only had permission to use FPR data for participants who had deceased before the linkage of the registers, December 31, 2018 (1117 individuals).

Data Matching and Linkage

The matching of survey variables with ICD-10 codes from the CRHC and ATC-codes from the FPR is shown in Table 1. All events with ICD-10 diagnoses from CRHC and all medication dispenses with a reimbursement code from FPR were drawn until the beginning of survey data collection in 2014 (January 17, 2014). CRHC included information from January 1, 1996, and FPR from January 1, 2010, onwards.

The register data were linked with the survey data using the personal identification code, a unique identifier assigned to all residents in Finland.[23] The linkage was done by Statistics Finland. All data accessed were handled according to relevant data protection and privacy regulations.

The protocol for Vitality 90+ Study has been approved by the Regional Ethics Committee of Tampere University Hospital and the study has a research permit from the City of Tampere. Permission to use the register data was obtained from the register keepers. All survey participants or their representatives provided written informed consent and were asked for permission to link the survey information with the health registers.

Statistical Analysis

First, the analyses were conducted for all survey respondents with complete information on chronic diseases (n = 1548), comparing all reported diagnoses in the survey with diagnosis information from CRHC. Second, information from both CRHC and FPR was used. The comparison of survey data with both registers was carried out among the 1107 participants for whom we had information from both CRHC and FPR and complete data on chronic diseases in the survey (Figure 1).

We first calculated the frequency of all 10 diseases from the survey and CRHC for the whole sample (n = 1548) and then for the five diseases with a reimbursement code in FPR (n = 1107). Among participants in the latter sample, information from the two registers was used in combination, and a participant was considered to have the disease if it was recorded in either register.

Cohen's kappa statistic was used to test the agreement for each disease between survey and CRHC, between survey and FPR, and between survey and combined data from both registers. Level of agreement was evaluated on a scale from 0.00 to 1.00 as defined by Landis & Koch:[24] <0.00 poor, 0.00–0.20 slight, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80

Table 1 Matching of Survey Items, ICD-10, ATC, and Reimbursement Codes

Survey item	ICD-10	ATC code	Reimbursement code
Cancer	C00-C97	n.a.	n.a.
Diabetes	E10-E14	A10	Diabetes, insulin treatment (103), diabetes, other than insulin treatment (215)
Dementia or Alzheimer's disease, loss of memory	F00-F03, G30	N06D	Donepezil, Galantamine, memantine, rivastigmine (307)
Depression, depressed mood	F31-F34, F38-F39	n.a.	n.a.
Parkinson's disease	G20	N04	Parkinson's disease (110)
Hypertension, high blood pressure	I10-I15	C02C, C03A, C03B, C03C, C03D, C03E, C07A, C07B, C07F, C08C, C08D, C09A, C09B, C09C, C09D	Chronic hypertension (205)
Heart disease (coronary artery disease, arrhythmia, myocardial infarction)	I20-I25, I47-I50	C01A, C01D, C03A, C03C, C03D, C03E, C07A, C07B, C08C, C08D, C09A, C09B, C09C, C09D, C10A	Heart failure (201), chronic coronary artery disease and related fat metabolism disorders (206), chronic arrhythmias (207)
Hip fracture	S72.0, S72.1, S72.2	n.a.	n.a.
Stroke	I60-I64, I69	n.a.	n.a.
Osteoarthritis	M15-M19	n.a.	n.a.

Abbreviations: ICD-10, International Classification of Diseases, Tenth Revision; ATC code, Anatomical Therapeutic Chemical code; n.a, not available.
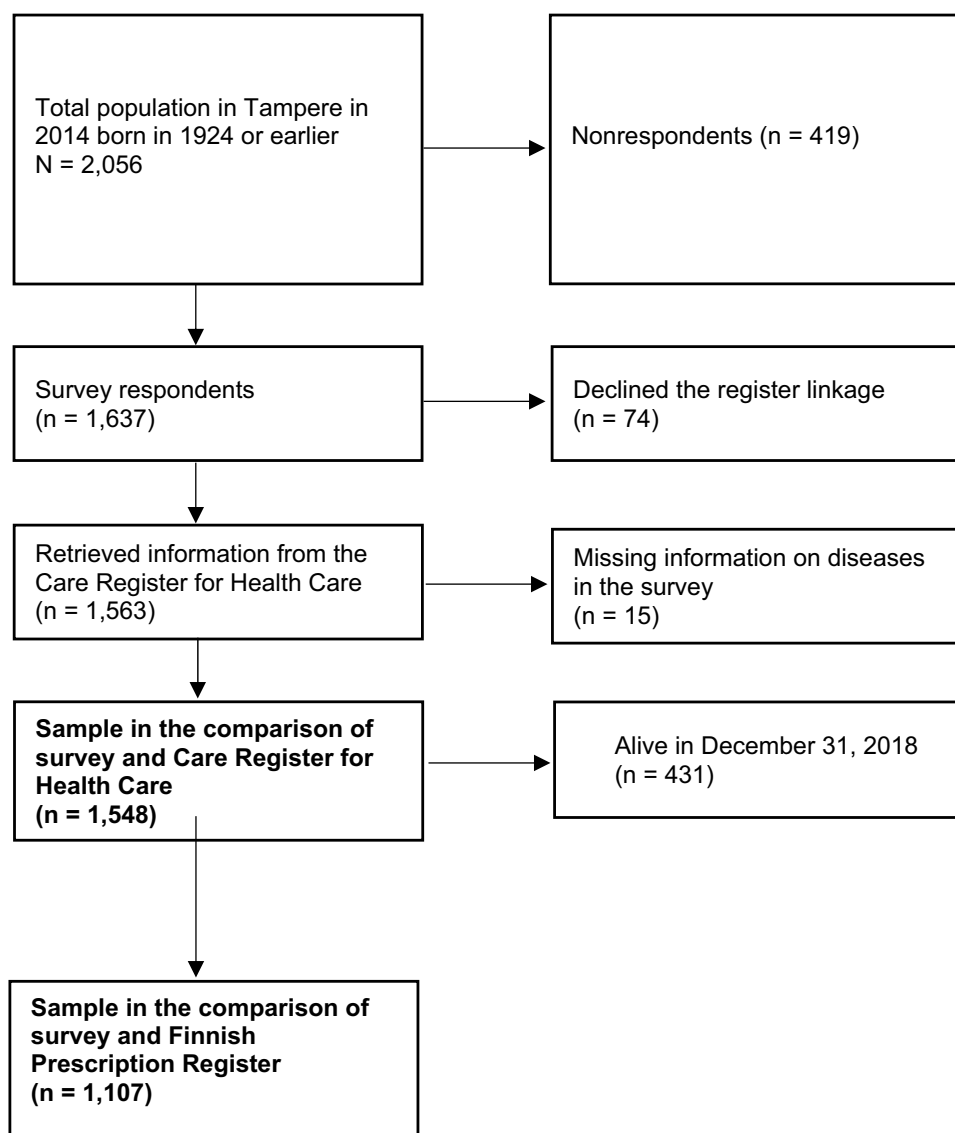
Figure 1 Flowchart of the study participants.

substantial, and 0.81–1.00 almost perfect. Positive and negative percent agreement were calculated to assess whether the discrepancy between the data sources originated from mismatch in positive or negative ratings.[25] Using the register as a reference, PPA denotes the proportion of matching positive ratings in the register and the survey out of all positive ratings in the register, and NPA denotes the proportion of matching negative ratings in the register and the survey out of all negative ratings in the register.

Analysis was performed with IBM SPSS Statistics version 28.0.

Results

Participant Characteristics

Mean age of the participants with complete disease information from the survey and CRHC was 92.7 (range 90–106), while in the sub-sample with information available from FPR it was 93.0 (range 90–106). The proportion of women was about the same in the whole study sample (76.5%) and in the sub-sample (76.6%) but the share of proxy respondents (18.6% and 24.6%, respectively) and residents in long-term care (36.4% vs 44.4%) was higher in the sub-sample.

Disease Prevalence in the Survey and the Registers

Among all respondents (n = 1548), the survey showed a higher prevalence than CRHC for all diseases, except for cancer and stroke, which had a higher prevalence in the register. Hypertension and heart disease were the two most prevalent diseases in both the survey and CRHC. The prevalence of osteoarthritis (43.9%) and dementia (42.4%) were also high in the survey, but much lower (24.9% and 26.0%, respectively) in CRHC (Table 2).

In the sub-sample (n = 1107), the survey showed a higher prevalence for all five diseases than either CRHC or FPR. The lowest prevalence rates for all diseases were obtained from FPR. When information from both registers was combined, the prevalence of all five diseases was closer to the figure from the survey (Table 3).

Agreement Between Survey Data and the Registers
Cohen's Kappa

Hip fracture (κ=0.65), diabetes (κ=0.63), and Parkinson's disease (κ=0.61) showed substantial agreement between the survey and CRHC among all survey participants. Stroke (κ=0.59), cancer (κ=0.57), dementia (κ=0.57), and heart disease (κ=0.52) showed moderate agreement and hypertension (κ=0.33), depression (κ=0.23), and osteoarthritis (κ=0.21) showed fair agreement (Table 2).

In the sub-sample, hypertension and Parkinson's disease showed higher agreement with FPR than CRHC. For heart disease, dementia, and diabetes, the agreement between survey and CRHC was higher than between survey and FRP. For

Table 2 Disease Prevalence in Survey and CRHC, Kappa Coefficients, Positive and Negative Percent Agreement Between Survey and CRHC and Two-by-Two Table of Frequencies (n = 1548)

	Survey % (n)	CRHC % (n)	κ	PPA %	NPA %	CHRC - Survey - n (%)	CRHC - Survey + n (%)	CRHC + Survey - n (%)	CRHC + Survey + n (%)
Osteoarthritis	43.9 (680)	24.9 (386)	0.21	63.5	62.6	727 (47)	435 (28)	141 (9)	245 (15.8)
Depression	17.1 (264)	5.7 (88)	0.23	59.1	85.5	1248 (80.6)	212 (13.7)	36 (2.3)	52 (3.4)
Hypertension	60.8 (941)	44.3 (685)	0.33	80.0	54.5	470 (30.4)	393 (25.4)	137 (8.9)	548 (35.4)
Heart disease	53.8 (833)	51.4 (796)	0.52	78.9	72.7	547 (35.3)	205 (13.2)	168 (10.9)	628 (40.6)
Cancer	16.5 (255)	24.9 (386)	0.57	54.4	96.1	1117 (72.2)	45 (2.9)	176 (11.4)	210 (13.6)
Dementia	42.4 (656)	26.0 (403)	0.57	93.3	75.5	865 (55.9)	280 (18.1)	27 (1.7)	376 (24.3)
Stroke	9.0 (140)	14.6 (226)	0.59	51.3	98.2	1298 (83.9)	24 (1.6)	110 (7.1)	116 (7.5)
Parkinson's disease	1.6 (24)	0.8 (12)	0.61	91.7	99.2	1523 (98.4)	13 (0.8)	1 (0.1)	11 (0.7)
Diabetes	15.4 (238)	10.4 (161)	0.63	83.9	92.6	1284 (82.9)	103 (6.7)	26 (1.7)	238 (15.4)
Hip fracture	17.2 (267)	13.1 (203)	0.65	81.3	92.4	1243 (80.3)	102 (6.6)	38 (2.5)	165 (10.7)

Notes: κ = Cohen's kappa. PPA is the percentage of positive ratings in CRHC and survey out of all positive ratings in CRHC. NPA is the percentage of negative ratings in CRHC and survey out of all negative ratings in CRHC. Percentages in the last four columns are proportions out of total observations.
Abbreviations: CRHC, Care Register for Health Care; PPA, positive percent agreement; NPA, negative percent agreement.

Table 3 Disease Prevalence in Survey, CRHC, FPR, and in Both Registers Combined Information, Kappa Coefficients and Positive and Negative Percent Agreement Across All Data Sources (n = 1107)

	Survey % (n)	CRHC % (n)	κ[a]	PPA %	NPA %	FPR % (n)	κ[b]	PPA	NPA	Both Registers % (n)	κ[c]	PPA	NPA
Hypertension	58.0 (642)	48.0 (531)	0.37	77.2	59.7	34.5 (382)	0.42	89.8	58.8	58.0 (642)	0.51	79.4	71.6
Heart disease	56.6 (627)	56.3 (623)	0.48	77.4	70.0	21.2 (235)	0.26	89.8	52.3	59.8 (662)	0.51	77.5	74.4
Dementia	50.3 (557)	33.5 (371)	0.59	94.1	71.7	20.2 (224)	0.38	96.9	61.5	37.4 (414)	0.66	94.7	76.2
Diabetes	15.0 (166)	11.8 (131)	0.68	81.7	94.0	7.6 (84)	0.62	97.6	92	13.3 (147)	0.75	83.7	95.5
Parkinson's disease	1.6 (18)	1.0 (11)	0.69	90.9	99.3	1.0 (11)	0.76	100	99.4	1.3 (14)	0.81	92.9	99.5

Notes: [a]Kappa coefficient between survey and CRHC. [b]Kappa coefficient between survey and FPR. [c]Kappa coefficient between survey and both registers' information combined. PPA is the percentage of positive ratings in a register and survey out of all positive ratings in the register. NPA is the percentage of negative ratings in a register and survey out of all negative ratings in the register.
Abbreviations: CRHC, Care Register for Health Care; FPR, Finnish Prescription Register; PPA, positive percent agreement; NPA, negative percent agreement.
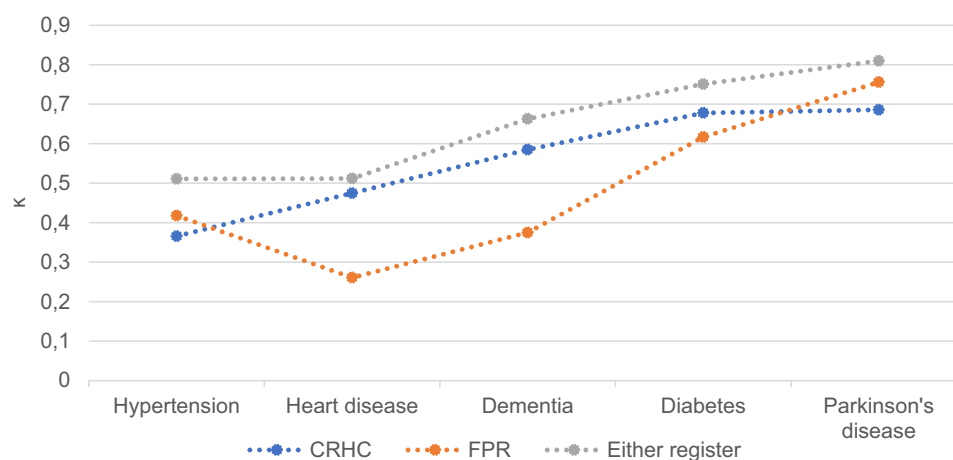
Figure 2 Kappa coefficients (κ) between survey and Care Register for Health Care (CRHC), survey and Finnish Prescription Register (FPR), and survey and either register in the sub-sample of 1107 respondents.

all five diseases, the agreement was highest when the two registers were combined. Parkinson's disease showed almost perfect agreement (κ=0.81); diabetes (κ=0.75) and dementia (κ=0.66) substantial agreement; and heart disease (κ=0.51) and hypertension (κ=0.51) moderate agreement (Table 3 and Figure 2).

For most diseases, agreement was slightly higher among proxy respondents than self-respondents between survey and CRHC. Agreement was substantial for Parkinson's disease, diabetes, and hip fracture. However, the agreement for dementia ranged from slight to moderate among proxy respondents, and agreement between proxy respondents and FPR was lower than between self-respondents and FPR (Supplementary Tables 1 and 2).

Positive and Negative Percent Agreement

Both PPA and NPA were high for diseases showing high agreement: Parkinson's disease, diabetes, and hip fracture. For hypertension and dementia, PPA was markedly higher than NPA, as in these diseases the discrepancy stemmed from a mismatch in negative ratings. A large number of hypertension and dementia cases reported in the survey were lacking from the register. Stroke and cancer showed high NPA and low PPA. Thus, some cases recorded in CRHC were lacking from the survey. Depression showed a very low prevalence in CRHC, and the mismatch was due to most survey-reported cases lacking from CRHC (Table 2).

Among participants with information available from both registers, all diseases showed higher PPA than NPA between the survey and FPR. PPA for FPR was higher than for CRHC, since there were very few cases that were registered in FPR but not reported in the survey. PPA for Parkinson's disease between FPR and the survey was 100%, indicating that all participants with a register entry of Parkinson's disease reported it in the survey (Table 4).

In proxy responses, NPA for dementia was low. Proxy respondents reported dementia far more frequently than either one of the registers (Supplementary Tables 1 and 2).

Discussion

This study assessed the agreement of disease information between a population-based survey and two national administrative health registers in a population aged over 90. Prevalence rates varied quite widely between the data sources. Agreement was high for clear, well-defined physical diseases: hip fracture, Parkinson's disease, and diabetes, and higher when information from the two registers was combined.

We used two Finnish national registers that are designed for administrative rather than research purposes yet nonetheless commonly used in research.[13] They cover the whole population and are therefore valuable sources of information. Our aim was to capture both inpatient and outpatient care by using one register with information originating from specialized care and hospitalization data, and the other with information on medication dispended from pharmacies, representing outpatient care.

Table 4 Two-by-Two Tables of Frequencies for Each Disease in the Comparison Between Survey and CRHC, Survey and FPR, and Survey and both Registers Information Combined. Percentages are Proportions Out of Total Observations for Each Disease (n = 1107)

	CRHC - Survey -	CRHC - Survey +	CRHC + Survey -	CRHC + Survey +	FPR - Survey -	FPR - Survey +	FPR + Survey -	FPR + Survey +	Registers - Survey -	Registers - Survey +	Registers + Survey -	Registers + Survey +
Hypertension n (%)	344 (31.1)	232 (21.0)	121 (10.9)	410 (37.0)	426 (38.5)	299 (27.0)	39 (3.5)	343 (31.0)	333 (30.1)	132 (11.9)	132 (11.9)	510 (46.1)
Heart disease n (%)	339 (30.6)	145 (13.1)	141 (12.7)	482 (43.5)	456 (41.2)	416 (37.6)	24 (2.2)	211 (19.1)	331 (29.9)	114 (10.3)	149 (13.5)	513 (46.3)
Dementia n (%)	528 (47.7)	208 (18.8)	22 (2.0)	349 (31.5)	543 (49.1)	340 (30.7)	7 (0.6)	217 (19.6)	528 (47.7)	165 (14.9)	22 (2.0)	392 (35.4)
Diabetes n (%)	917 (82.8)	59 (5.3)	24 (2.2)	107 (9.7)	939 (84.8)	84 (7.6)	2 (0.2)	82 (7.4)	917 (82.8)	43 (3.9)	24 (2.2)	123 (11.1)
Parkinson's disease n (%)	1088 (98.3)	8 (0.7)	1 (0.1)	10 (0.9)	1089 (98.4)	7 (0.6)	0 (0)	11 (1.0)	1088 (98.3)	5 (0.5)	1 (0.1)	13 (1.2)

Abbreviations: CRHC, Care Register for Health Care; FPR, Finnish Prescription Register.

For most of the diseases studied, self-reports showed a higher prevalence than the registers, which is consistent with an earlier study among octogenarians.[11] The only exceptions were cancer and stroke, which showed a higher register prevalence, probably because our data dated back several years and at the time of the survey, the diseases may not have been active or shown any symptoms. It is also possible that cancer and stroke are more often recorded in CRHC since they can lead to hospitalization more often than other diseases included. Some studies have found higher prevalence rates in registers than in self-reports.[9,10] These studies used general practice records[9] and clinical examination[10] as a reference for self-reported information, whereas we used a patient register in which the diagnoses are based on the need for inpatient treatment or specialized care, possibly explaining the differing findings. Furthermore, as multimorbidity is high among the oldest old,[26] it is reasonable to assume that not all diseases are recorded during every care episode. To cover outpatient care, we used FPR data and indeed, the two registers seemed to complement each other well: the prevalence rates were closer to those reported in the survey and the agreement was better when the information from the two registers was combined. Similar findings have been presented concerning the identification of dementia from CRHC and FPR.[27]

In line with earlier studies with the oldest old,[8,10] we found that agreement between self-reports and health registers is fair to moderate for hypertension and heart disease. Likewise, our results are consistent with studies among younger old people, which have found that registers and survey data show high agreement for clearly defined diseases that require specific medication.[2,3,28] In our study, these diseases included Parkinson's disease, diabetes, and hip fracture, which showed an agreement ranging from substantial to almost perfect.

The level of agreement was lower for osteoarthritis and depression, which rarely lead to hospitalization and are not always treated with prescription drugs. Again, this is consistent with earlier studies.[5,9] Reimbursement for antidepressants is limited to severe depression, whereas in the survey we included any depression and depressed mood.

Dementia showed moderate to substantial agreement but was more frequently reported in the survey than in the registers. A previous Finnish study found that the register cases could be verified in clinical examinations, but all the cases diagnosed in clinical examinations were not recorded in registers. Thus, CRHC underestimated the prevalence of dementia.[27] Reimbursement in FPR only covers medication for AD, but among the oldest old cognitive impairment may be due to other conditions as well.[29] It is also possible that cases of early dementia are missed and hence not recorded in a register,[27] or that diagnostic examinations are not conducted because of the patient's overall clinical situation. Importantly, the wording used for dementia in our survey included "dementia, Alzheimer's disease or weakening of memory", and even though the question concerned diseases diagnosed by a doctor, it is possible that respondents also reported their own experiences and perceptions of a weakening of memory.

Our results confirm the importance of allowing the use of proxy respondents in studies of the oldest old.[30] Most of the participants who used a proxy in our study had dementia, and they more often lived in long-term care than those who responded independently. Consistent with earlier findings, we found high agreement between proxy-reported information and health register data on well-defined diseases.[11] Proxies may have underreported diseases not requiring active treatment or not showing prominent symptoms, such as hypertension and osteoarthritis. However, they reported a clearly higher prevalence of dementia than indicated by the registers. In addition to the wording of the question, another possible explanation is the tendency of proxies to overrate the level of disabilities, especially cognitive problems.[31]

The strengths of our study include the linking of survey data with two different nationwide registers covering inpatient and specialized care and all reimbursed drug purchases from pharmacies. The survey included all residents of Tampere aged over 90, regardless of their health or place of living, and it had a high response rate. The 90+ population in Tampere is rather homogenous in terms of ethnic background and represents well the overall 90+ population in Finland in terms of life expectancy and sociodemographic factors.[32] Limitations include the lack of register information on primary outpatient care and the fact that FPR data were only available for part of our survey respondents. Moreover, because the criteria for drug reimbursement are sometimes stricter than the diagnostic criteria, we have missed some cases with at least milder hypertension and heart failure. On the other hand, as only drug purchases eligible for reimbursement were included, we know they represent true clinically confirmed cases.

792

Clinical Epidemiology 2023:15

Conclusion

Self-reported disease data show acceptable agreement with register data among the oldest old people. Yet the use of proxy respondents is necessary in order to gain information on representative population samples, including people with cognitive decline. It is also important to acknowledge the gaps in health registers when validating self-reports against register information. In all, our study implies that also in the oldest old, survey data can be reliably used in population studies where clinical examinations are not feasible.

Funding

Disclosure

The authors report no conflicts of interest in this work.

References

1. Casey JA, Schwartz BS, Stewart WF, Adler NE. Using electronic health records for population health research: a review of methods and applications. *Annu Rev Public Health*. 2016;37(1):61–81. doi:10.1146/annurev-publhealth-032315-021353
2. Hansen H, Schäfer I, Schön G, et al. Agreement between self-reported and general practitioner-reported chronic conditions among multimorbid patients in primary care - results of the multicare cohort study. *BMC Fam Pract*. 2014;15(1):39. doi:10.1186/1471-2296-15-39
3. 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(10):1096–1103. doi:10.1016/j.jclinepi.2004.04.005
4. Griffith LE, Gruneir A, Fisher KA, et al. Measuring multimorbidity series—an overlooked complexity comparison of self-report vs. administrative data in community-living adults: paper 2. Prevalence estimates depend on the data source. *J Clin Epidemiol*. 2020;124:163–172. doi:10.1016/j.jclinepi.2020.04.019
5. 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(3):160–165. doi:10.1177/2150131913517902
6. 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(1):16. doi:10.1186/1471-2458-13-16
7. Skinner KM, Miller DR, Lincoln E, Lee A, Kazis LE. Concordance between respondent self-reports and medical records for chronic conditions: experience from the veterans health study. *J Ambul Care Manag*. 2005;28(2):102–110. doi:10.1097/00004479-200504000-00002
8. 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(10):237–245. doi:10.1007/s10433-013-0275-7
9. Hale MD, Santorelli G, Brundle C, Clegg A. A cross-sectional study assessing agreement between self-reported and general practice-recorded health conditions among community dwelling older adults. *Age Ageing*. 2019;49(1):135–140. doi:10.1093/ageing/afz124
10. Andersen-Ranberg K, Fjederholt KT, Madzak A, Nybo M, Jeune B. Cardiovascular diseases are largely underreported in Danish centenarians. *Age Ageing*. 2013;42(2):249–253. doi:10.1093/ageing/afs108
11. Rydén L, Sigström R, Nilsson J, et al. 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. *Age Ageing*. 2019;48(4):513–518. doi:10.1093/ageing/afz033
12. Goebeler S, Jylhä M, Hervonen A. Self-reported medical history and self-rated health at age 90. Agreement with medical records. *Aging Clin Exp Res*. 2007;19(3):213–219. doi:10.1007/BF03324692
13. Laugesen K, Ludvigsson JF, Schmidt M, et al. Nordic health registry-based research: a review of health care systems and key registries. *Clin Epidemiol*. 2021;13:533–554. doi:10.2147/CLEP.S314959
14. Laatikainen T, Koponen P, Reinikainen J, et al. Mitä tietoa Suomessa saadaan hoitoilmoitusrekistereistä ja mitä västötutkimuksista? [Monitoring, assessment and prediction of public health: what type of information can be obtained in Finland from care registers and what from population studies?]. *Finn Med J*. 2020;75(37):1853–1858. Finnish.
15. United Nations, Department of Economic and Social Affairs, Population Division. *World Population Prospects 2022* Online Edition. 2022.
16. Prince MJ, Wu F, Guo Y, et al. The burden of disease in older people and implications for health policy and practice. *Lancet*. 2015;385(9967):549–562. doi:10.1016/S0140-6736(14)61347-7
17. Enroth L, Halonen P, Tiainen K, Raitanen J, Jylhä M. Cohort profile: the Vitality 90+ Study. A cohort study on health and living conditions of the oldest old in Tampere, Finland. *BMJ Open*. 2023;13:e068509. doi:10.1136/bmjopen-2022-068509
18. Care Register for Health Care - THL. Finnish institute for health and welfare, THL Finland; 2021. Available from: https://thl.fi/en/web/thlfi-en/statistics-and-data/data-and-services/register-descriptions/care-register-for-health-care. Accessed May 9, 2022.
19. Sund R. Quality of the Finnish hospital discharge register: a systematic review. *Scand J Public Health*. 2012;40(6):505–515. doi:10.1177/1403494812456637
20. Furu K, Wettermark B, Andersen M, Martikainen JE, Almarsdottir AB, Sorensen HT. The Nordic countries as a cohort for pharmacoepidemiological research. *Basic Clin Pharmacol Toxicol*. 2010;106(2):86–94. doi:10.1111/j.1742-7843.2009.00494.x

21. Haukka J, Suvisaari J, Tuulio-Henriksson A, Lönnqvist J. High concordance between self-reported medication and official prescription database information. *Eur J Clin Pharmacol*. 2007;63(11):1069–1074. doi:10.1007/s00228-007-0349-6

22. Rikala M, Hartikainen S, Sulkava R, Korhonen MJ. Validity of the Finnish prescription register for measuring psychotropic drug exposures among elderly finns. *Drugs Aging*. 2010;27(4):337–349. doi:10.2165/11315960-000000000-00000

23. Digital and Population Data Services Agency. Population information system; 2023. Available from: https://dvv.fi/en/population-information-system. Accessed January 11, 2023.

24. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–174. doi:10.2307/2529310

25. Cicchetti DV, Feinstein AR. High agreement but low Kappa: II. Resolving the paradoxes. *J Clin Epidemiol*. 1990;43(6):551–558. doi:10.1016/0895-4356(90)90158-L

26. Halonen P, Raitanen J, Jämsen E, Enroth L, Jylhä M. Chronic conditions and multimorbidity in population aged 90 years and over: associations with mortality and long-term care admission. *Age Ageing*. 2019;48(4):564–570. doi:10.1093/ageing/afz019

27. Solomon A, Ngandu T, Soininen H, Hallikainen MM, Kivipelto M, Laatikainen T. Validity of dementia and Alzheimer's disease diagnoses in Finnish national registers. *Alzheimers Dement*. 2014;10(3):303–309. doi:10.1016/j.jalz.2013.03.004

28. Guerra SG, Berbiche D, Vasiliadis H-M. Measuring multimorbidity in older adults: comparing different data sources. *BMC Geriatr*. 2019;19(1):166. doi:10.1186/s12877-019-1173-4

29. Brumback-Peltz C, Balasubramanian AB, Corrada MM, Kawas CH. Diagnosing dementia in the oldest-old. *Matur*. 2011;70(2):164–168. doi:10.1016/j.maturitas.2011.07.008

30. Kelfve S, Thorslund M, Lennartsson C. Sampling and non-response bias on health-outcomes in surveys of the oldest old. *Eur J Ageing*. 2013;10(3):237–245. doi:10.1007/s10433-013-0275-7

31. Li M, Harris I, Lu ZK. Differences in proxy-reported and patient-reported outcomes: assessing health and functional status among medicare beneficiaries. *BMC Med Res Methodol*. 2015;15(1):62. doi:10.1186/s12874-015-0053-7

32. Jylhä M, Enroth L, Luukkaala T. Trends of Functioning and Health in Nonagenarians: the Vitality 90+ Study. *Annu Rev Gerontol Geriatr*. 2013;33(1):313–332. doi:10.1891/0198-8794.33.313