




OPEN ACCESS

Self-report of chronic diseases in old-aged individuals: extent of agreement with general practitioner medical records in the German AugUR study

Anna B Steinkirchner,¹ Martina E Zimmermann ,¹ Ferdinand J Donhauser,¹ Alexander Dietl,² Caroline Brandl,^{1,3} Michael Koller,⁴ Julika Loss,^{5,6} Iris M Heid,¹ Klaus J Stark ¹

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/jech-2022-219096>).

¹Department of Genetic Epidemiology, University of Regensburg, Regensburg, Germany

²Department of Internal Medicine II, University Hospital Regensburg, Regensburg, Germany

³Department of Ophthalmology, University Hospital Regensburg, Regensburg, Germany

⁴Center for Clinical Studies, University Hospital Regensburg, Regensburg, Germany

⁵Department for Epidemiology and Preventive Medicine, Medical Sociology, University of Regensburg, Regensburg, Germany

⁶Department of Epidemiology and Health Monitoring, Robert Koch Institute, Berlin, Germany

Correspondence to

Professor Klaus J Stark, Department of Genetic Epidemiology, University of Regensburg, Regensburg, Germany; klaus.stark@klinik.uni-regensburg.de

Received 31 March 2022

Accepted 5 August 2022

Published Online First 26 August 2022



© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Steinkirchner AB, Zimmermann ME, Donhauser FJ, et al. *J Epidemiol Community Health* 2022;**76**:931–938.

ABSTRACT

Background To estimate prevalence and incidence of diseases through self-reports in observational studies, it is important to understand the accuracy of participant reports. We aimed to quantify the agreement of self-reported and general practitioner-reported diseases in an old-aged population and to identify socio-demographic determinants of agreement.

Methods This analysis was conducted as part of the AugUR study (n=2449), a prospective population-based cohort study in individuals aged 70–95 years, including 2321 participants with consent to contact physicians. Self-reported chronic diseases of participants were compared with medical data provided by their respective general practitioners (n=589, response rate=25.4%). We derived overall agreement, over-reporting/under-reporting, and Cohen's kappa and used logistic regression to evaluate the dependency of agreement on participants' sociodemographic characteristics.

Results Among the 589 participants (53.1% women), 96.9% reported at least one of the evaluated chronic diseases. Overall agreement was >80% for hypertension, diabetes, myocardial infarction, stroke, cancer, asthma, bronchitis/chronic obstructive pulmonary disease and rheumatoid arthritis, but lower for heart failure, kidney disease and arthrosis. Cohen's kappa was highest for diabetes and cancer and lowest for heart failure, musculoskeletal, kidney and lung diseases. Sex was the primary determinant of agreement on stroke, kidney disease, cancer and rheumatoid arthritis. Agreement for myocardial infarction and stroke was most compromised by older age and for cancer by lower educational level.

Conclusion Self-reports may be an effective tool to assess diabetes and cancer in observational studies in the old and very old aged. In contrast, self-reports on heart failure, musculoskeletal, kidney or lung diseases may be substantially imprecise.

INTRODUCTION

As a population grows older, chronic diseases are not only burdens to the individual but also increasingly become a challenge to the healthcare system and society as a whole. Epidemiological longitudinal and panel studies have been designed to investigate older populations in terms of ageing and its impact on health.^{1–3} These studies often use questionnaire-based self-report measures of chronic disease.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Self-reports of diseases in studies are known to be of limitations. They have mostly been studied in younger cohorts and through comparison with diverse sources of medical data.

WHAT THIS STUDY ADDS

⇒ This study investigates the agreement of self-reports of the old and very old population in Germany with reports of general practitioners. The old aged are most frequently affected by the studied diseases and yet give inaccurate self-reports on heart failure, musculoskeletal, kidney or lung diseases, whereas self-reports on diabetes and cancer are more accurate. Agreement between these two sources of information is also affected by sociodemographic factors, such as sex, age and education level.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ When using data concerning self-reported diseases of study participants in a similar population group, the results this study have provided will help to interpret accuracy.

Limitations of self-reports have been the subject of studies before: sociodemographic factors, illness perceptions, severity of symptoms and resources to understand a condition may impact consistency and accuracy of self-reports.^{4–7} Several studies focus on the validity and reliability of self-reports in different study populations and with different measures of agreement;^{5–8–19} however, there has only been a small number of such studies focused on Germany's elderly population.

The German Altersbezogene Untersuchungen zur Gesundheit der Universität Regensburg (AugUR) study was established as a research platform to estimate the prevalence and incidence of chronic conditions and to understand associated risk factors for chronic health disabilities in the elderly.²⁰ In the following AugUR substudy, we covered diseases and medical events, which are dealt with in general practice: hypertension, diabetes, myocardial infarction, heart failure, stroke, kidney disease, cancer, lung diseases and musculoskeletal diseases. All of

them are either persistent or irreversible and, therefore, should become apparent in both the participants' self-reports and the medical data provided by their health professionals. Sex, age, the status of living with a partner and education level might influence the participants' awareness of the disease and, therefore, the degree of agreement with the diagnosis made by their general practitioners (GPs). Recognising sociodemographic factors can help to identify subgroups, which might benefit from better communication about diagnosed diseases.^{21 22} Limited awareness or knowledge of one's own status of disease may lead to non-compliance with therapy as well as the oversight of warning signs of deterioration.²³

Two questions were addressed in the following substudy analysis: (1) To what extent do self-reports of old-aged individuals agree with the information given by their GPs? (2) Are there sociodemographic determinants of agreement (age, sex, living status, education level)?

METHODS

AugUR cohort study

The German AugUR study is a prospective study of the general elderly population in and around the city of Regensburg, Bavaria. The study region comprises ~347 000 inhabitants, including 45 000 residents aged 70 or older.²⁰ From the latter, a random sample from population registries in Regensburg and selected communities of the county was identified (n=13 971) and invited for the baseline study centre visits in two consecutive, comparable surveys (AugUR-1 in 2013–2015 and AugUR-2 in 2017–2019). From a total of 13 522 contactable persons (n=449 deceased or moved outside the recruitment area before being invited), 2449 participated at the two AugUR baseline study surveys (n=1133 in AugUR-1 and n=1316 in AugUR-2, respectively), resulting in a net response rate of 18.1% (n=8171 did not respond at all, n=2902 actively refused participation). Further details on recruitment and response are presented in the online supplemental note. Consent and valid information to contact GPs was given by 2321 participants (figure 1).

For 788 AugUR-1 baseline participants, 3-year follow-up data were collected in 2016–2018. Between baseline and follow-up, 67 persons died, 37 moved away, whereas 3 did not agree to be recontacted (net response rate=788/(1133-67-37-3)=76.8%). The consecutive 6-year follow-up started in November 2019 and paused in March 2020 due to the COVID-19 pandemic after inclusion of n=123 participants (net response rate so far=72.4%).

We consider the participants physically mobile and without major cognitive impairments, since they have had to visit the study centre and actively take part at the study programme.

AugUR study programme and data management

The study programme, including a standardised in-person interview, was conducted in the study centre at the University Medical Centre Regensburg. AugUR focuses on ophthalmic diseases classified within the study²⁴ and general chronic diseases evaluated through in-person-interview.²⁰

The baseline questionnaire included an assessment of socio-demographic characteristics such as sex, age, living status and education level. Medical conditions and further medical history were evaluated at baseline and follow-up visits via self-report using the question 'Has a physician ever diagnosed one of the following conditions?'. Possible answers were 'yes', 'no' and 'I do not know'. The latter was excluded from the analysis.

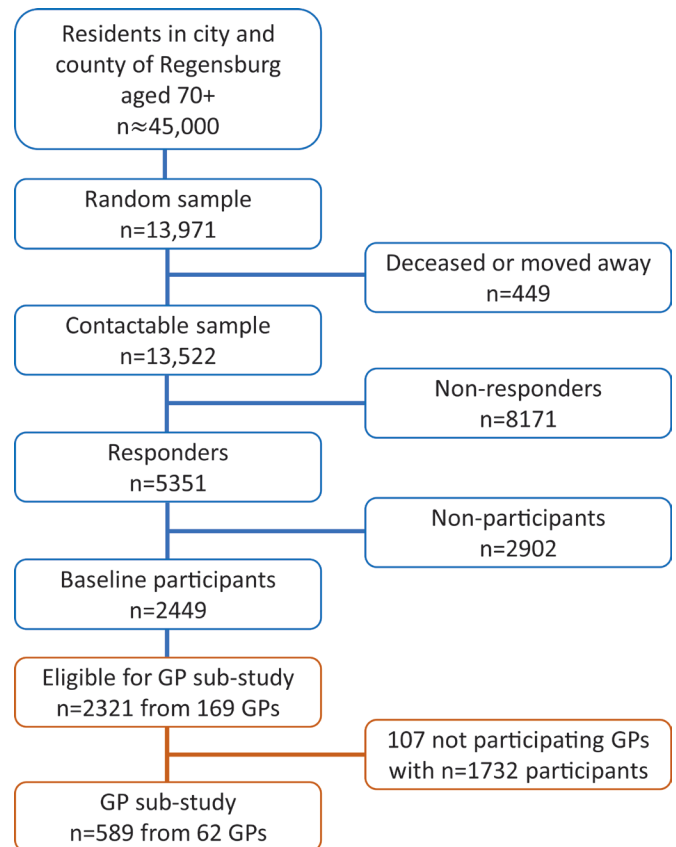


Figure 1 Overview of AugUR study recruitment with number of participants at baseline and in the general practitioner (GP) sub-study. Net response at baseline was 18.1%. For the GP sub-study, 36.7% of GPs responded, resulting in 25.4% of eligible AugUR participants. AugUR; Altersbezogene Untersuchungen zur Gesundheit der Universität Regensburg.

Patterns of missing values of both self-reports and GP reports are presented in online supplemental table 1.

Some of the terms for diagnoses had to be adapted, so participants could better understand them. For example, the patients were asked about 'heart weakness' instead of the medical term 'heart failure', which was used in the questionnaires for GPs. Questionnaire data were transferred to an electronic case report form.²⁰ For all participants, we evaluated the status of disease and the sociodemographic characteristics at their most recent interview. We included follow-up data; therefore, inconsistencies between baseline and follow-up information could arise. This is a common phenomenon among longitudinal data and needs to be dealt with.⁴ Adjustments had to be made if a participant reported a condition in an earlier interview and denied it during the following visit. Since the lifetime prevalences of diseases were asked for, and not the current prevalences of diseases, we set the disease status 'yes' for subsequent interviews and used that status in further analysis.

The study complies with the 1964 Helsinki declaration and its later amendments. All participants provided informed written consent. For this substudy, participants with given informed consent to contact their physicians were included.

AugUR study participants' data collected from GPs

A total of 2321 AugUR participants gave informed consent to contact their physicians and provided valid contact information of their respective GPs (n=169). A standardised survey of the

Cross-table as reference for agreement terms					Cohen's kappa		
		GP-report					
		Yes	No	Total			
Self-report	Yes	a	b	a+b	$\frac{\frac{a+d}{n} - \frac{a+b}{n} * \frac{a+c}{n} + \frac{c+d}{n} * \frac{b+d}{n}}{1 - \frac{a+b}{n} * \frac{a+c}{n} + \frac{c+d}{n} * \frac{b+d}{n}}$		
	No	c	d	c+d			
	Total	a+c	b+d	n			
Overall agreement		$\frac{a+d}{n}$		Positive agreement	$\frac{2a}{(2a+d+c)}$	Sensitivity	$\frac{a}{a+c}$
Underreporters		$\frac{c}{a+c}$		Negative agreement	$\frac{2d}{(2d+b+c)}$	Specificity	$\frac{d}{b+d}$
Overreporters		$\frac{b}{b+d}$		Positive predictive value	$\frac{a}{a+b}$	Negative predictive value	$\frac{d}{c+d}$

Figure 2 Cross-table and terms used for analysing agreement. A template cross-table is exemplified (top left). The rows show the positive and negative self-reports, the columns show the positive and negative GP reports. Also shown are terms used for analysing agreement, including the formula used to calculate Cohen's kappa. Overall agreement, positive agreement, negative agreement, underreporters and overreporters are further used as percentage by multiplication of the depicted terms by factor 100. GP, general practitioner.

participants' GPs was conducted between October 2020 and March 2021. 62/169 contacted GPs responded and provided data on 589/2321 study participants (25.4%), forming the AugUR GP substudy population. Multicentric data from GPs were collected using the web-based Magana Trial Manager (MaganaMed GmbH, Regensburg, Germany). The GPs recorded the data directly online or provided the data on paper-based questionnaires, which were transferred to Magana Trial Manager by trained AugUR staff. The questions read: 'According to your records, does this patient have a diagnosis of disease X?'. Additionally, due to the study design, the date of first diagnosis was requested for diabetes, myocardial infarction, heart failure, stroke, kidney disease and cancer. It is possible that a disease was diagnosed by a GP after the participants' last chance of reporting it in an interview, as a proportion of participant interviews had been conducted 2018 or earlier. If that was the case, the disease was considered absent. Raw data and further analysis on absent cases are found in online supplemental tables 2 and 3.

Statistical analysis

Self-reports and GP reports of up to 589 participants were compared (online supplemental table 4). Figure 2 gives an overview of the agreement parameters used in the following analyses and how they were determined. Concordance in absence or presence of disease between self-reports and GP reports was analysed by calculating overall agreement (figure 2). In addition, we identified over-reporters and under-reporters. An over-reporter is defined as a person reporting an illness, which is not confirmed by the GP, while an under-reporter does not report an illness, which is reported by the GP. Under-reporters/over-reporters and sensitivity/specificity can easily be translated into one another (figure 2). However, sensitivity and specificity are terms used for gold standards,²⁵ and, therefore, unsuitable in our analyses. To describe agreement between self-reports and GP reports in an omnibus index and control for agreement by chance, Cohen's kappa was calculated. According to Landis and Koch's classification for agreement adjusted by chance, we refer to kappa values between 0.81 and 1 as 'almost perfect', 0.61 to 0.80 as 'substantial', 0.41 to 0.60 as 'moderate' and 0 to 0.40 as 'poor to fair'.^{26 27} As kappa can be influenced towards both higher and lower numbers by the distributions in the cross-table, especially in the marginal totals, we also calculated specific agreement in form of positive and negative agreement (figure 2),

as suggested by Hansen *et al* and used by Cicchetti and Feinstein.^{9 28} The proportions of specific agreement (ie, positive and negative agreement) estimate the conditional probability, in cases where one of the raters (GP or participant)—randomly selected—makes a positive/negative rating, the other rater will do so as well.

Disease frequencies and total numbers of overall agreement stratified by the independent variables sex, age ('old vs 'very old', stratified at median age of 79.03 years), the status of living with a partner and the education level were assessed. Because, in Germany, 8 years of education are often followed by vocational training and >8 years qualify for higher education, we set 8 years of schooling education as the threshold to form the two groups.

In logistic regression analyses, we used overall agreement as the dependent variable (participant overall agrees with GP: coded as 1; participant does not overall agree with GP, that is, is under-reporter or over-reporter: coded as 0), with sex, age (in years, linear), status of living with a partner and education level as independent variables. Due to limited sample size and partially low prevalence of diseases, separate analyses of true-positives and true-negatives or under-reporters and over-reporters would have led to unreliable results. A p value of <0.05 was used as criterion for statistical significance.

Data management and statistical analyses were performed using SAS 9.4 software (SAS Institute, Cary, North Carolina), IBM SPSS Statistics for Windows, V.26.0.0.1 (IBM, Armonk, New York) and Microsoft Excel V.2019.

RESULTS

Characteristics of GP substudy individuals and AugUR study population

The characteristics of all AugUR participants and the substudy sample with available GP data showed comparable distributions (table 1).

The median time gap between the self-report and the GP report was 2.66 years (range: 0.78–7.73 years) for all 589 participants (table 1).

Frequency of diseases in self-reports and GP reports

96.9% of the 589 participants reported at least one of the depicted chronic diseases. For most diseases, the frequency

Table 1 Characteristics of GP substudy participants and AugUR participants

	GP sub-study participants n=589	AugUR participants n=2321
Sex n (%)		
Women	313 (53.1)	1219 (52.5)
Men	276 (46.9)	1102 (47.5)
Age median (IQR)	79.0y (75.5–82.6y)	78.9y (75.7–82.6y)
Living with partner n (%)		
No	265 (45.0)	954 (41.1)
Yes	324 (55.0)	1367 (58.9)
Education n (%)		
≤8 years	309 (53.1)	1173 (50.9)
>8 years	273 (46.9)	1130 (49.1)
Most recent interview time point n (%)		
2013–2015 (AugUR-1-BL)	72 (12.2)	322 (13.9)
2016–2018 (AugUR-1-F1)	146 (24.8)	641 (27.6)
2017–2019 (AugUR-2-BL)	340 (57.7)	1244 (53.6)
2019–2020 (AugUR-1-F2)	31 (5.3)	114 (4.9)
Median time difference to GP substudy		
All	2.66y	
AugUR-1-BL	5.86y	
AugUR-1-F1	2.81y	
AugUR-2-BL	2.09y	
AugUR-1-F2	0.86y	
Shown are the numbers for 589 GP substudy participants and all 2321 participants of AugUR, which gave consent and valid information to contact their GPs. Absolute numbers and percentages are shown. y=years; AugUR-1-BL=baseline visit AugUR-1 cohort; AugUR-1-F1=follow-up-1 visit AugUR-1 cohort; AugUR-2-BL=baseline visit AugUR-2 cohort; AugUR-1-F2=follow-up-2 visit AugUR-1 cohort. Comparisons between GP-substudy participants and non-participants are presented in online supplemental table 5.		
AugUR, Altersbezogene Untersuchungen zur Gesundheit der Universität Regensburg; GP, general practitioner.		

was higher in self-reports than in GP reports, for example, for myocardial infarction, kidney disease, lung diseases and musculoskeletal diseases. For diabetes and hypertension, GP reports showed higher frequencies than self-reports (table 2).

Self-reported disease frequencies stratified by sex, old versus very old age, status of living with a partner and education level show that men more often report diabetes, myocardial infarction, stroke, kidney diseases, cancer and COPD/chronic bronchitis. Women more often claim to suffer from hypertension, heart failure, asthma, rheumatoid arthritis and arthrosis (online supplemental figure 1). Very old participants more often report all chronic diseases except kidney diseases and asthma. Participants who lived with a partner more often self-reported cancer and asthma. Participants who had been educated for >8 years more often reported heart failure, stroke and rheumatoid arthritis (online supplemental figure 1).

Agreement on diseases between self-reports and GP reports

As one of our goals was to determine, how well self-reports depict the actual disease status of our participants, we analysed, to what extent the participants self-reports align with their GP's records. Overall agreement, over-reporting, under-reporting, kappa, positive and negative-specific agreement were evaluated.

High overall agreement with >80% was found for hypertension, diabetes, myocardial infarction, stroke, cancer, asthma, chronic bronchitis/COPD and rheumatoid arthritis. Overall agreement <80% was found for heart failure, kidney disease and arthrosis (table 2). High numbers of over-reporting were seen for hypertension (35.5%), kidney disease (25.6%) and arthrosis (59.5%). For under-reporting, high numbers are seen for heart failure (57.9%), kidney disease (57.1%), chronic bronchitis/COPD (54.8%) and rheumatoid arthritis (53.8%) (table 2).

Positive agreement of >80% was discovered for hypertension and diabetes (table 2). For most conditions, negative agreement of >80% was found, except for hypertension, kidney disease and arthrosis (table 2).

Cohen's kappa was substantial for diabetes and cancer. Moderate agreement in observed data was found for myocardial infarction, stroke, asthma and hypertension. Poor to fair agreement was observed for chronic bronchitis/COPD, heart failure, rheumatoid arthritis, arthrosis and kidney disease (table 2).

Table 2 Frequencies of chronic diseases in self-reports and GP reports and parameters for agreement

Disease (n)	Frequencies % (n)		Parameters of agreement					
	Self-reports	GP reports	Overall agreement % (n)	Overreporters% (n)	Underreporters % (n)	Positive agreement %	Negative agreement %	Cohen's kappa
Hypertension (n=570)	74.2 (423)	76.7 (437)	81.1 (462)	35.5 (47)	14.0 (61)	87.4	61.4	0.50
Diabetes (n=586)	24.1 (141)	28.3 (166)	90.6 (531)	3.6 (15)	24.1 (40)	82.1	93.6	0.76
Myocardial infarction (n=577)	9.7 (56)	6.6 (38)	93.8 (541)	5.0 (27)	23.7 (9)	61.7	96.6	0.58
Heart failure (n=577)	19.1 (110)	16.5 (95)	78.3 (452)	14.5 (70)	57.9 (55)	39.0	86.8	0.26
Stroke (n=581)	9.1 (53)	7.4 (43)	93.8 (545)	4.3 (23)	30.2 (13)	62.5	96.6	0.59
Kidney disease (n=581)	28.9 (168)	19.3 (112)	68.3 (397)	25.6 (120)	57.1 (64)	34.3	79.1	0.15
Cancer (n=584)	26.9 (157)	23.6 (138)	87.2 (509)	10.5 (47)	20.3 (28)	74.6	91.4	0.66
Asthma (n=519)	10.0 (52)	6.7 (35)	92.9 (482)	5.6 (27)	28.6 (10)	57.7	96.1	0.53
Bronchitis/COPD (n=514)	8.9 (46)	6.0 (31)	90.5 (465)	6.6 (32)	54.8 (17)	36.4	94.8	0.31
Rheumatoid arthritis (n=511)	13.5 (69)	5.1 (26)	86.1 (440)	11.8 (57)	53.8 (14)	25.3	92.3	0.19
Arthrosis (n=537)	66.9 (359)	44.9 (241)	56.4 (303)	59.5 (176)	24.1 (58)	61.0	50.6	0.16

Hypertension, diabetes, myocardial infarction, heart weakness, stroke, kidney disease, cancer (excluding white skin cancer), lung diseases and musculoskeletal diseases were addressed. Self-reports and GP reports, overall agreement, over-reporters and under-reporters are shown as totals and in percentages. Specific positive and negative agreement is shown in percentages. All parameters were calculated as shown in figure 2. Total n varies due to missing values. COPD, chronic obstructive pulmonary disease; GP, general practitioner.

Table 3 Multiple regression models analysing the association of overall agreement between self-reports and GP reports with independent variables

Overall agreement for	Variable	OR	CI (95%)	P value
Hypertension	Sex men	0.702	0.445 to 1.107	0.128
	Age per 1 year	0.999	0.957 to 1.043	0.960
	Living w. partner yes	0.926	0.581 to 1.477	0.748
	Education >8 y	0.840	0.549 to 1.284	0.420
Diabetes	Sex men	1.021	0.550 to 1.895	0.947
	Age per 1 year	0.968	0.915 to 1.024	0.261
	Living w. partner yes	1.736	0.931 to 3.235	0.083
	Education >8 y	1.173	0.661 to 2.081	0.586
Myocardial infarction	Sex men	0.518	0.246 to 1.092	0.084
	Age per 1 year	0.931	0.873 to 0.994	0.031
	Living w. partner yes	1.460	0.694 to 3.070	0.319
	Education >8 y	1.033	0.522 to 2.047	0.925
Heart failure	Sex men	0.808	0.522 to 1.249	0.338
	Age per 1 year	0.965	0.927 to 1.005	0.083
	Living w. partner yes	1.492	0.961 to 2.316	0.075
	Education >8 y	0.866	0.579 to 1.295	0.484
Stroke	Sex men	0.391	0.181 to 0.844	0.017
	Age per 1 year	0.927	0.868 to 0.989	0.022
	Living w. partner yes	1.078	0.509 to 2.286	0.844
	Education >8 y	1.048	0.528 to 1.078	0.893
Kidney disease	Sex men	0.606	0.412 to 0.891	0.011
	Age per 1 year	0.965	0.931 to 1.000	0.051
	Living w. partner yes	1.276	0.863 to 1.886	0.222
	Education >8 y	1.157	0.810 to 1.653	0.424
Cancer	Sex men	0.528	0.309 to 0.901	0.019
	Age per 1 year	1.011	0.961 to 1.065	0.663
	Living w. partner yes	0.907	0.526 to 1.566	0.727
	Education >8 y	1.927	1.149 to 3.231	0.013
Asthma	Sex men	1.294	0.628 to 2.668	0.484
	Age per 1 year	1.017	0.948 to 1.090	0.642
	Living w. partner yes	0.688	0.326 to 1.453	0.327
	Education >8 y	1.864	0.925 to 3.755	0.082
Bronchitis/COPD	Sex men	0.869	0.454 to 1.662	0.671
	Age per 1 year	0.963	0.909 to 1.021	0.204
	Living w. partner yes	0.788	0.403 to 1.539	0.485
	Education >8 y	0.839	0.461 to 1.527	0.565
Rheumatoid arthritis	Sex men	2.276	1.280 to 4.047	0.005
	Age per 1 year	0.977	0.929 to 1.028	0.369
	Living w. partner yes	0.693	0.395 to 1.214	0.200
	Education >8 y	0.764	0.458 to 1.276	0.304
Arthrosis	Sex men	1.158	0.797 to 1.682	0.443
	Age per 1 year	0.993	0.959 to 1.027	0.675
	Living w. partner yes	0.866	0.593 to 1.263	0.454
	Education >8 y	1.155	0.818 to 1.632	0.413

Each model for the respective disease (hypertension, diabetes, myocardial infarction, heart failure, stroke, kidney disease, cancer, lung diseases and musculoskeletal diseases) was adjusted for all four independent variables (sex, age, living status and education level). ORs with 95% CIs and p-value are shown. Significantly associated characteristics are shown in bold letters. Independent variables were coded 0/1, so numbers show the disease's OR for men versus women, being 1 year older, living with a partner versus without and having >8 y of education versus less.
COPD, chronic obstructive pulmonary disease; GP, general practitioner.

Analysis of age, sex, living status and education level

We depicted overall agreement, stratified by sex, age, status of living with a partner and education level (online supplemental figure 1).

Regression analyses showed that men were significantly less likely than women to report the same as their GP regarding stroke (OR=0.391, p value=0.017), kidney disease (OR=0.606, p value=0.011) and cancer (OR=0.528, p value=0.019) (table 3).

For rheumatoid arthritis, the opposite effect was found, as women were significantly less likely to agree with their GPs (OR=2.276, p value=0.005) (table 3).

Older participants were significantly less likely to agree with their GPs on myocardial infarction (OR=0.931, p value=0.031) and stroke (OR=0.927, p value=0.022) (table 3). The regression analysis also showed that participants with low education of ≤ 8 years of schooling were significantly less likely to agree with their GPs on their cancer status (OR=1.927, p value=0.013) (table 3).

The status of living with a partner did not show significant association with overall agreement.

DISCUSSION

Our study highlights chances and challenges of interview-based self-reports on diseases and medical events in the population aged 70+. We found high agreement of self-reports and GP reports for diabetes and cancer. Therefore, self-reports are an effective tool to assess these diseases in observational studies in the old and very old population. Low agreement not only for heart failure but also for musculoskeletal, kidney or lung diseases indicate substantial imprecision when relying on self-reports. Our association analyses showed that being male, of very old age or having received less than 8 years of schooling education, was associated with lower agreement and, thus, with more inaccurate information concerning the self-reported disease status, which can induce differential misclassifications.

The data here presented showed highest overall agreement and kappa values for *diabetes and cancer* between self-reports and GP reports. This can be explained by both diseases requiring intensive intervention and being very present in a person's daily life: to lower the risk of organ damage, participants with diabetes need to adjust their lifestyles in terms of diet and exercise or require medication.²⁹ Furthermore, in Germany, a high proportion of patients suffering from diabetes is enrolled in disease management programmes, which include educational events and trainings, continuous monitoring by GPs and repeated screening for organ damage by ophthalmologists and by physicians specialised in internal medicine. A cancer diagnosis often causes physical and mental strains,³⁰ leading to higher awareness of the disease. Our results are in line with other studies (table 4).

However, the agreement on cancer was not regularly addressed and varied among studies with kappa values from 0.33 to 0.67 with our kappa value of 0.66 on the top end. In comparably old study populations, similar values were found (table 4).

In AugUR, overall agreement was high, but Cohen's kappa was moderate for *myocardial infarction and stroke*, which were both mainly under-reported. Myocardial infarction and stroke can be life-threatening events. However, not all events may have been explicitly explained by the physician, leading to a lack of awareness. Comparable results are seen in other studies with kappa values ranging from 0.33 to 0.80 and 0.36 to 0.71 for myocardial infarction and stroke, respectively (table 4). The wide range of kappa values may be caused by the variety of terms used to ask for both diseases, for example, while we asked participants for a 'brain attack', other studies also differentiated transient ischaemic attacks. Variability in the severity of the events between study populations may also explain some of the differences in reported agreement, since more threatening events might lead to higher awareness and, thus, higher agreement.

In our study, agreement for *asthma* was moderate, while agreement for *chronic bronchitis/COPD* was only poor to fair. Lung diseases were asked about in different ways across studies,

Table 4 Selection of literature: kappa values and characteristics for studies evaluating kappa statistics for agreement

Authors	Ref	Cohort characteristics	Age group	Kappa*													
				HYP	DM	MI	HF	Stroke	KD	Cancer	Asthma	CB/COPD	RA	Arthrosis			
Camplain <i>et al</i>	5	Prospective cohort recruited from US communities (n=11 846)	6083y					0.39									
Okura <i>et al</i>	8	Population-based prospective cohort to study cardiac function (n=2307)	Mean: 61y	0.75	0.76	0.80	0.46	0.71									
Hansen <i>et al</i>	9	Multimorbid participants recruited from GP practices (n=3189)	mean: 74.4y	0.56	0.80		0.24	0.55				0.50	0.61§	0.27	0.29		
van den Akker <i>et al</i>	10	Prospective, dynamic cohort recruited from healthcare centres (n=2893)	Mean: 68y	0.86	0.86	0.57		0.52				0.60	0.58§	0.17			
Kriegsman <i>et al</i>	11	Independently living participants drawn from population registries (n=2380)	5585y	0.85	0.85		0.56	0.63§				0.66	0.63§	0.35			
Merzenich <i>et al</i>	12	Breast cancer patients recruited from hospitals (n=1212)	Mean: 65.9y	0.71	0.78	0.54	0.34	0.61	0.40					0.57			
Barber <i>et al</i>	13	Population-based cohort, recruited from three GP practices (n=5889)	≥50y	0.67	0.90												
Englert <i>et al</i>	14	Patients with primary hypercholesterinaemia enrolled in primary care centres (n=7640)	Mean: 61y	0.69	0.89	0.55	0.29	0.44									
Merkin <i>et al</i>	15	End-stage renal disease patients recruited from dialysis clinics (n=965)	Mean: 58y	0.20	0.93	0.55	0.47	0.59			0.67			0.20			
Muggah <i>et al</i>	16	Sample of community dwelling general population (n=85 549)	≤20y	0.66	0.80	0.48	0.33	0.36				0.55		0.29			
Leikauf <i>et al</i>	17	Participants randomly recruited from outpatient practices (n=323)	Mean: 73.1y	0.59	0.94						0.66						
Corser <i>et al</i>	18	Hospitalised Acute Coronary Syndrome patients (n=525)	Mean: 59.7y	0.80	0.63	0.09	0.54	0.47	0.33	0.43				0.07			
Teh <i>et al</i>	19	Octogenarians of general population (n=878)	Mean: 82.3y†, 84.6y‡	0.44		0.45	0.19	0.43									
This study		Mobile elderly general population drawn from population registries (n=589)	Mean: 79.0y	0.50	0.76	0.58	0.26	0.59	0.15	0.66	0.53	0.31	0.19	0.16			

We listed the literature reference, the author, a concise summary describing the study cohort and an age reference. Mean age was used if available, otherwise we depicted the study cohort's age range. Kappa values shown for HYP, DM, MI, HF, KD, CB/COPD, RA, arthrosis.

*Highest reported kappa values for each disease were used, as they may vary due to different disease-definitions within a study.

† Age evaluated for Māori and

‡ Non-Māori participants in New Zealand;

§ Evaluated asthma and COPD as one category or as 'chronic lung disease'.

CB, chronic bronchitis; COPD, chronic obstructive pulmonary disease ; DM, diabetes; HF, heart failure; HYP, hypertension; KD, stroke, kidney disease; MI, myocardial infarction; RA, rheumatoid arthritis; REF, literature reference.

some asking for either asthma or COPD and some for chronic lung diseases in general, but reported agreement was comparable to our results for asthma (table 4). Previous reports indicated that 50% of patients with asthma believed they only had asthma when they were currently experiencing symptoms,³¹ which might explain our finding of high under-reporting.

For hypertension, our kappa value of 0.50 was well within the range of other reports (table 4). We found high over-reporting and low negative agreement, indicating that the participants were more likely to invent a diagnosis than denying it. In the literature, however, an under-reporting of hypertension is predominantly documented.⁶

We observed low kappa values for *heart failure, kidney disease and musculoskeletal diseases* (table 4). Throughout literature, heart failure is an example for rather poor agreement between self-reports and GP reports or medical records. Low agreement may be due to the complexity of the heart failure diagnosis and communication with patients.⁵ In contrast, published data on the agreement for kidney disease are limited. Two studies showed kappa values of 0.40 and 0.47 (table 4), which are high compared with our kappa of 0.15, possibly due to the younger age of their study participants with lower kidney disease frequency and the setting of interviewing hospitalised participants. For musculoskeletal diseases such as rheumatoid arthritis and arthrosis, we found poor to fair agreement in line with the literature (table 4). Fluctuating symptoms of joint diseases and the tendency to treat them without consulting a physician³² may explain low agreement. As we see especially high numbers of over-reporting for arthrosis, while rheumatoid arthritis is under-reported more often, there is the possibility that patients confuse the diseases for one another.

Agreement of self-reported disease with GP reports in the general population provides insights into awareness of the disease or of awareness of not having a disease. Awareness of disease is a prerequisite for compliance with treatment plans and lifestyle changes^{33–35}; awareness of not having a disease documents a general understanding of one's own health status.

In stratified analyses, overall agreement might be higher for the group in which disease frequency is lower compared with the other group (eg, men vs women), so subgroup disease frequencies must be considered in the interpretation of the results. According to our data, men are more likely to overall agree with their GPs on rheumatoid arthritis; however, disease frequencies are much lower in that group (9.6% in men vs 17.0% in women), so significant association may be influenced by that difference, as it might be easier to correctly report the absence of a disease than the presence of a disease. While we found that men were less likely to agree with their GPs on stroke, kidney disease and cancer, there is no consensus on the influence of sex on agreement for these diseases in the literature.^{9–11 14 15}

Our findings of older age being associated with lower agreement for myocardial infarction and stroke were in accordance with a general decline of agreement by age reported in other studies.^{8–12 18}

Living with a partner was found to improve health awareness³⁶ and, thus, to potentially increase agreement. We did not find this status to be significantly associated with agreement and none of the discussed studies evaluated the influence of living with a partner.

Our finding that participants with higher education of >8 years were more likely to agree with their GPs regarding cancer indicates a higher ability to comprehend this diagnosis. Other studies also found higher education to be positively associated with higher agreement.^{8 12 14} The fully adjusted model with sex,

age and living with a partner as covariables showed that the education effect on overall agreement for cancer self-report is not confounded by the other aspects.

Strengths and limitations

We need to acknowledge a limitation caused by the low response rate of GPs (26.0%), which is partly because we refrained from reminding GPs in the middle of the Corona pandemic. To further address the limitation by selection bias, we compared the characteristics and disease count of our study subgroup with the group of AugUR participants not being represented in our GP substudy. As a random selection of practitioners gave information on the disease status, we consider the subgroup sample to not to underlie selection bias. Yet, our sample from Regensburg and selected communities of the county may not be representative for other regions. The median time gap of 2.66 years between self-reports and GP reports raises limitations as well, however, we tried to minimise them by including all available follow-up information. Furthermore, our usage of lay language in participants' questionnaires for diseases may be a reason to potentially increase disagreement.^{9 14 37} Regarding the exclusion of missing values, for most diseases, small numbers in both self-reports and GP reports were found. However, in GP reports, we found more missing values in diseases where treatment by a specialist rather than the GP is plausible, that is, lung diseases and musculoskeletal diseases.

This study's major strength is its focus on old and very old individuals. The old and very old population is difficult to address in observational studies of wider age range and, therefore, often under-represented. This is in sharp contrast to the fact that the chronic diseases studied here are more frequent in old age. We tailored the study programme to the needs of the elderly by employing a leaner questionnaire and allocating additional time for walking around the study premises as well as for answering questions. The standardised face-to-face interviews with participants, rather than the use of unsupervised questionnaires, are also strength in our study; as is the wide variety of diseases evaluated here.

Acknowledgements The authors greatly appreciate the outstanding and committed study assistance of Lydia Mayerhofer, Magdalena Scharl, Sabine Schelter and Josef Simon. We thank Miriam Stoffregen and Caroline Kästner for critically reading the manuscript. Moreover, we thank all study participants for contributing to the AugUR study. We would like to express our special thanks to the general practitioners who made this work possible by providing data on diagnoses.

Contributors All authors have contributed to interpreting results and manuscript writing. All authors have read and approved the manuscript. Further contributions are: ABS: manuscript design, statistical analyses. MEZ: data management. FJD: data curation. AD: study physician. CB: study physician. MK: study support. JL: study support. IMH: study PI, project supervision, manuscript design. KJS: study coordination, project initiation and supervision, data management, statistical analysis, manuscript design. KJS is the guarantor of the paper.

Funding The AugUR study was supported by grants from the German Federal Ministry of Education and Research (BMBF 01ER1206 and BMBF 01ER1507) to IMH, by the German Research Foundation (DFG HE 3690/7-1 and BR 6028/2-1) to IMH and CB and by institutional budget (University of Regensburg).

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by Ethics Committee of the University of Regensburg, Germany, vote 12-101-0258. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have

been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Martina E Zimmermann <http://orcid.org/0000-0002-6916-4404>

Klaus J Stark <http://orcid.org/0000-0002-7832-1942>

REFERENCES

- Dapp U, Anders J, von Renteln-Kruse W, et al. The longitudinal urban cohort ageing study (LUCAS): study protocol and participation in the first decade. *BMC Geriatr* 2012;12.
- Stephoe A, Breeze E, Banks J, et al. Cohort profile: the English longitudinal study of ageing. *Int J Epidemiol* 2013;42:1640–8.
- Schäfer I, Hansen H, Schön G, et al. The German MultiCare-study: patterns of multimorbidity in primary health care – protocol of a prospective cohort study. *BMC Health Serv Res* 2009;9:145.
- Cigolle CT, Nagel CL, Blaum CS. Inconsistency in the self-report of chronic diseases in panel surveys: developing an adjudication method for the health and retirement study. *J Gerontol B Psychol Sci Soc Sci* 2018;73:901–12.
- Camplain R, Kucharska-Newton A, Loehr L, et al. Accuracy of self-reported heart failure. The Atherosclerosis risk in communities (ARIC) study. *J Card Fail* 2017;23:802–8.
- Gonçalves VSS, Andrade KRC, Carvalho KMB. Accuracy of self-reported hypertension: a systematic review and meta-analysis. *J Hypertens* 2018;36:970–8.
- Woodfield R, Sudlow CLM. Accuracy of patient self-report of stroke: a systematic review from the UK Biobank stroke outcomes group. *PLoS One* 2015;10:e0137538.
- Okura Y, Urban LH, Mahoney DW, et al. 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.
- 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.
- van den Akker M, van Steenkiste B, Krutwagen E, et al. Disease or no disease? disagreement on diagnoses between self-reports and medical records of adult patients. *European Journal of General Practice* 2015;21:45–51.
- Kriegsman DM, Penninx BW, van Eijk JT. 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.
- Merzenich H, Blettner M, Niehoff D, et al. Cardiac late events in German breast cancer patients: a validation study on the agreement between patient self-reports and information from physicians. *BMC Cardiovasc Disord* 2018;18.
- Barber J, Muller S, Whitehurst T, et al. Measuring morbidity: self-report or health care records? *Fam Pract* 2010;27:25–30.
- Englert H, Müller-Nordhorn J, Seewald S, et al. Is patient self-report an adequate tool for monitoring cardiovascular conditions in patients with hypercholesterolemia? *J Public Health* 2010;32:387–94.
- Merkin SS, Cavanaugh K, Longenecker JC, et al. Agreement of self-reported comorbid conditions with medical and physician reports varied by disease among end-stage renal disease patients. *J Clin Epidemiol* 2007;60:634–42.
- Muggah E, Graves E, Bennett C, et al. Ascertainment of chronic diseases using population health data: a comparison of health administrative data and patient self-report. *BMC Public Health* 2013;13.
- Leikauf J, Federman AD. Comparisons of self-reported and chart-identified chronic diseases in inner-city seniors. *J Am Geriatr Soc* 2009;57:1219–25.
- Corser W, Sikorskii A, Olomu A, et al. "Concordance between comorbidity data from patient self-report interviews and medical record documentation". *BMC Health Serv Res* 2008;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;66:1135–43.
- Stark K, Olden M, Brandl C, et al. The German augur study: study protocol of a prospective study to investigate chronic diseases in the elderly. *BMC Geriatr* 2015;15.
- Hödlmoser S, Winkelmayer WC, Zee J, et al. Sex differences in chronic kidney disease awareness among US adults, 1999 to 2018. *PLoS One* 2020;15:e0243431.
- Redondo A, Benach J, Subirana I, et al. Trends in the prevalence, awareness, treatment, and control of cardiovascular risk factors across educational level in the 1995–2005 period. *Ann Epidemiol* 2011;21:555–63.
- Jin J, Sklar GE, Min Sen Oh V, et al. Factors affecting therapeutic compliance: a review from the patient's perspective. *Ther Clin Risk Manag* 2008;4:269–86.
- Brandl C, Zimmermann ME, Günther F, et al. On the impact of different approaches to classify age-related macular degeneration: results from the German augur study. *Sci Rep* 2018;8:8675.
- Trevethan R. Sensitivity, specificity, and predictive values: foundations, Pliabilities, and pitfalls in research and practice. *Front. Public Health* 2017;5.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
- McHugh ML. Interrater reliability: the kappa statistic. *Biochemia Medica* 2012;22:276–82.
- Cicchetti DV, Feinstein AR. High agreement but low kappa: II. Resolving the paradoxes. *J Clin Epidemiol* 1990;43:551–8.
- Deutsche Diabetes Gesellschaft (DDG). S2k-Leitlinie Diagnostik, Therapie und Verlaufskontrolle des Diabetes mellitus im Alter [AWMF-Registriernummer: 057-017], 2018. Available: <https://www.awmf.org/leitlinien/detail/ll/057-017.html>
- Caruso R, Breitbart W. Mental health care in oncology. contemporary perspective on the psychosocial burden of cancer and evidence-based interventions. *Epidemiol Psychiatr Sci* 2020;29:1–4.
- Halm EA, Mora P, Leventhal H. No symptoms, no asthma: the acute episodic disease belief is associated with poor self-management among inner-city adults with persistent asthma. *Chest* 2006;129:573–80.
- Peat G, McCarney R, Croft P. Knee pain and osteoarthritis in older adults: a review of community burden and current use of primary health care. *Ann Rheum Dis* 2001;60:91–7.
- Lagi A, Rossi A, Passaleva MT, et al. Compliance with therapy in hypertensive patients. *Intern Emerg Med* 2006;1:204–8.
- Fernandez-Lazaro CI, García-González JM, Adams DP, et al. Adherence to treatment and related factors among patients with chronic conditions in primary care: a cross-sectional study. *BMC Fam Pract* 2019;20.
- Heydari A, Ziaee ES, Gazrani A. Relationship between awareness of disease and adherence to therapeutic regimen among cardiac patients. *Int J Community Based Nurs Midwifery* 2015;3:23–30.
- Umberson D, Karas Montez J, Montez JK. Social relationships and health: a flashpoint for health policy. *J Health Soc Behav* 2010;51:554–66.
- Beckett M, Weinstein M, Goldman N. Do health interview surveys yield reliable data on chronic illness among older respondents? *Oxford University Press* 2000;151:315–23.