RESEARCH

Inconsistencies in self-reported diabetes in a large panel study: the Survey of Health, Ageing and Retirement in Europe (SHARE)

Bernd Kowall[1*], Carolin Girschik[1] and Susanne Stolpe[1]

Abstract

Background The validity of self-reported chronic conditions has been assessed by comparing them with medical records or register data in several studies. However, the reliability of self-reports of chronic diseases has less often been examined. Our aim was to assess the proportion and determinants of inconsistent self-reports of diabetes in a long panel study.

Methods SHARE (Survey of Health, Ageing and Retirement in Europe) includes 140,000 persons aged ≥ 50 years from 28 European countries and Israel. We used data from waves 1 to 7 (except wave 3) collected between 2004 and 2017. Diabetes was assessed by self-report. An inconsistent report for diabetes was defined as reporting the condition in one wave, but denying it in at least one later wave. The analysis data set included 13,179 persons who reported diabetes, and answered the question about diabetes in at least one later wave. Log-binomial regression models were fitted to estimate crude and adjusted relative risks (RR) with 95% confidence intervals (CI) for the associations between various exposure variables and inconsistent report of diabetes.

Results The proportion of persons with inconsistent self-reports of diabetes was 33.0% (95% CI: 32.2%—33.8%). Inconsistencies occurred less often in persons taking antidiabetic drugs (RR = 0.53 (0.53—0.56)), persons with BMI ≥ 35 kg/m^2 versus BMI < 25 kg/m^2 (RR = 0.70, (0.64—0.77)), and poor versus excellent subjective health (RR = 0.87 (0.75—1.01)). Inconsistencies occurred more often in older persons (RR = 1.15 (1.12—1.18) per 10 years increase of age), and persons not reporting their age at diabetes onset (RR = 1.38 (1.31—1.45)).

Conclusion In SHARE, inconsistent self-report of diabetes is frequent. Consistent reports are more likely for persons whose characteristics make diabetes more salient, like intake of antidiabetic medication, obesity, and poor subjective health. However, lack of attention in answering the questions, and poor wording of the items may also play a role.

Keywords Diabetes mellitus, Longitudinal study, Panel study, Reliability, Self-report

*Correspondence:
Bernd Kowall
bernd.kowall@uk-essen.de
[1] Institute for Medical Informatics, Biometry and Epidemiology, University Hospital Essen, Hufelandstraße 55, 45147 Essen, Germany

Background

Multiwave panel studies are an important information source for the prevalence and incidence of chronic diseases. However, they often rely on self-reports of health conditions. If the accuracy of such self-reports is limited, this may bias estimates of disease prevalence and incidence. Effect measures of associations between various factors and disease prevalence and incidence may also be biased. The validity of self-reported diseases has

Kowall *et al. BMC Medical Research Methodology*   (2024) 24:7

Page 2 of 9

often been assessed by comparing them with medical records or register data [1–3]. However, the reliability of self-reports of chronic diseases has much less often been examined.

Diabetes mellitus as a chronic disease is in principle not curable. Remission of type 2 diabetes has been shown to be possible after intensive weight management with substantial weight loss and after bariatric surgery [4, 5]. However, for the vast majority of patients, diabetes is still a chronic condition. Thus, participants of a panel study who reported to have diabetes in one wave should give a „yes " response in all later waves. The consistency of repeated reporting of diabetes has only rarely been examined [6–11]. One such study was done in a highly specific sample of prostate cancer survivors [7], in other studies, the number of included persons with diabetes was rather small (at most 315 in [6, 8, 11]). In all these studies, substantial proportions of participants gave inconsistent reports of diabetes which ranged from 2.9% in the Health and Retirement Study (HRS) to 39.2% in the Norwegian Women and Cancer Study (NOWAC) [8, 9]. However, these earlier studies differed considerably in the number of waves, in time intervals between the waves, the age of the participants, and definitions of inconsistent reporting of diabetes, so that the proportions of respondents giving inconsistent reports are difficult to compare.

This study uses data of the longitudinal Survey of Health, Ageing and Retirement in Europe (SHARE) [12, 13]. This large study allows us to build a very large sample of people with diabetes who have participated in up to seven waves of the panel over a period of about 12 years. We aim to answer the following research questions: first, to estimate the proportions of persons giving inconsistent self-reports of diabetes; second, to examine a wide range of factors which may have an influence on inconsistent reporting of diabetes (sociodemographic factors, drug intake, age at onset of diabetes, number of visits to the doctor, self-rated health, BMI, etc.).

Methods

Study population and data analysis set

SHARE is a longitudinal study with 140,000 persons aged 50 years or older from 28 European countries and from Israel [12, 13]. We used data from waves 1, 2, 4, 5, 6 and 7 of SHARE which were collected in 2004/2005/2006 (wave 1), 2006/2007 (wave 2), 2011 (wave 4), 2013 (wave 5), 2015 (wave 6) and 2017 (wave 7) [14–19]. SHARE participants who took part in any previous wave were invited again for later waves. Moreover, refreshment samples were drawn to compensate for the loss of participants. We did not take wave 3 into account because it has a focus on the life history of the participants and, thus, differs from the other waves. Participants are interviewed every two years. The interviews cover a wide range of topics, including demographics, physical and mental health, cognitive function, health care, lifestyle, social support, housing, employment, pensions, household income, and financial transfers. The study rationale and design have been described elsewhere, and further information on SHARE is available online [12, 13]. SHARE data are available free of charge after registration.

The data analysis set includes all participants who said at least in one wave that they had diabetes or had been told by their doctor to have diabetes and who answered the question about diabetes in at least one later wave.

Variables

To assess diabetes at wave 1, participants were shown a card with 16 diseases, and they were asked: "Has a doctor ever told you that you had any of the conditions on this card? Please tell me the number or numbers of the conditions. … Diabetes or high blood sugar." At the subsequent waves, participants were shown the card again and asked: "Has a doctor ever told you that you had/Do you currently have any of the conditions on this card? With this we mean that a doctor has told you that you have this condition, and that you are either currently being treated for or bothered by this condition. Please tell me the number or numbers of the conditions." Diabetes duration was calculated from the answer on the following question: "About how old were you when you were first told by a doctor that you had diabetes or high blood sugar?" To assess diabetes medication, participants were shown a card with diseases, and asked: „Do you currently take drugs at least once a week for problems mentioned on this card? … Drugs for diabetes."

Heart disease and stroke were assessed with the same questions as for diabetes. Subjective health was assessed with the question „Would you say your health is (excellent / very good / good / fair / poor)? ". BMI was calculated from self-reports of weight and height. International Standard Classification of Education (ISCED) codes were provided in SHARE in the modul gv_isced. Western Europe includes Austria, Germany, the Netherlands, France, Switzerland, Belgium, and Luxembourg; Northern Europe includes Sweden and Denmark; Southern Europe includes Spain, Italy, Greece, Portugal, Croatia, and Israel; Eastern Europe includes Czech Republik, Poland, Hungary, Slovenia, and Estonia. All variables were assessed at the wave where the persons reported to have diabetes for the first time.

Definition of inconsistencies in self-report of diabetes

An inconsistency in self-report of diabetes is defined as follows: A patient states that he has diabetes or has been

told by a doctor to have diabetes in one wave, but says that he has no diabetes in at least one later wave.

Statistical analysis

Characteristics of participants in the analysis data set are given separately for each wave (means ± standard deviation for continuous variables, proportions for categorical variables). Each person in the analysis data set took part in at least two waves. For categorical variables, the proportions of persons giving inconsistent self-reports of diabetes are estimated with 95% confidence intervals for each category.

Log-binomial regression models were fitted to assess crude and adjusted relative risks with 95% confidence intervals for the association between various exposure variables and inconsistencies in self-report of diabetes (yes/no) as the outcome. An adjustment set was determined for each exposure variable. Adjustment sets can be found under Table 2.

For each wave, we additionally estimated the proportion of persons who said that they did not have diabetes among persons who said that they took diabetes medication. Moreover, we estimated the proportion of persons who said they had diabetes after an inconsistent self-report of diabetes in an earlier wave.

In a secondary analysis, we performed a multiple imputation, using the SAS procedure PROC MI with the statement FCS DISCRIM (fully conditional specification, discriminant function). The number of imputations was 10. We imputed self-report of diabetes at waves 1, 2, 4, 5, 6 and 7, and, additionally, for covariates which may have an influence on inconsistencies in self-report of diabetes (namely age, sex, marital status, country, ISCED, BMI, subjective health, times the doctor was seen, history of stroke, history of myocardial infarction, intake of diabetes drugs). The imputed value for wave 1 was subsequently reset to missing for persons who participated for the first time in wave 2. This was handled in the same way for participants with later baseline waves.

Analyses were performed with SAS (version 9.4; SAS Institute, Cary, North Carolina).

Results

All together, 139,010 persons participated with the baseline either in wave 1, 2, 4, 5, 6 or 7. From participants of wave 1, 7,116 did not participate in a later wave. From new participants in wave 2, 6,383 did not participate in a later wave (wave 4: 8,275; wave 5: 6,677; wave 6: 4,334; wave 7: 19,032). Thus, 87,193 persons remained who participated at least in one later wave after baseline. Of these, 13,179 fulfilled the condition that they reported to have diabetes in one wave and gave a self-report of diabetes (positive or negative) in at least one further wave. The

mean age of these persons increased from 64.7 years in wave 1 to 72.6 years in wave 7 (Table 1). Between 8.5% to 13.2% of the participants described their health as excellent or very good. Between 18.5% and 22.2% reported an earlier heart attack, and between 4.9% to 7.7% reported an earlier stroke.

The proportion of persons with inconsistent self-reports of diabetes was 33.0% (95% confidence interval: 32.2%–33.8%). This proportion increases with the number of waves participants took part in: for persons participating in 2, 3, 4, 5 and 6 waves, the corresponding proportions of inconsistencies were 24.5%, 31.5%, 36.1%, 43.0%, and 42.0%, respectively. More women than men (34.4 versus 31.5%), more older (≥ the median age, which is 67 years) than younger ones (< 67 years) (35.3 versus 30.3%), and more lower educated persons (ISCED grade 0–2) than higher educated persons (ISCED grade 3–6) (35.4 versus 30.4%) gave inconsistent reports of their diabetes status (Fig. 1). The proportion of persons with inconsistent report of diabetes was lowest in Northern Europe (25.9%) and highest in Southern Europe (38.9%). The proportion of persons with inconsistent diabetes reports decreased from 39.0% in persons with normal weight to 33.8% in persons with overweight to 30.5% in persons with obesity grade 1 and to 25.6% in persons with obesity of at least grade 2. Persons taking no diabetes medication at their first self-report of diabetes were more likely to give inconsistent reports of diabetes at later waves than persons who reported that they took diabetes medication at their first self-report of diabetes (60.7 versus 25.4%). Among persons who did not report their age at diabetes diagnosis, and among persons who did not report how often they saw the doctor per year, the proportion of those giving inconsistent self-reports of diabetes was above average (43.9%, and 42.9%, respectively). The better subjective health was, the higher was the proportion of inconsistent diabetes self-reports (37.4% for excellent, and 31.7% for poor subjective health).

In the adjusted log-binomial regression analyses, women, persons older than the median age, persons from Eastern and Southern Europe, and persons not reporting their age at diabetes onset had a higher risk of giving inconsistent self-reports of diabetes (Table 2). The higher persons′ BMI was, the lower was their risk of inconsistencies in reporting diabetes: persons with a BMI ≥ 35 kg/m² had a relatively 30% lower risk of inconsistencies compared to persons with BMI < 25 kg/m² (RR = 0.70 (0.64–0.77)). Persons describing their health as poor had a relatively 13% lower risk of inconsistencies than persons with excellent subjective health (RR = 0.87 (0.75–1.01)). Persons who reported an earlier heart attack or an earlier stroke barely differed in their risk of giving inconsistent diabetes self-reports from persons without heart attack

Kowall *et al. BMC Medical Research Methodology*   (2024) 24:7

Page 4 of 9

Table 1 Characteristics of study participants in the data analysis set

		Wave 1	Wave 2	Wave 4	Wave 5	Wave 6	Wave 7
N		3793	5077	8064	9527	10,147	9130
Sex	Male	1831 (48.3%)	2406 (47.4%)	3704 (46.4%)	4625 (48.6%)	4831 (47.6%)	4310 (47.2%)
	female	1962 (51.7%)	2671 (52.6%)	4324 (53.6%)	4902 (51.4%)	5316 (52.4%)	4820 (52.8%)
Age (years)		64.7±9.3	67.0±9.3	68.4±9.3	70.0±9.5	71.2±9.3	72.6±9.0
ISCED	Grade 3–6	1431 (38.1%)	2099 (42.1%)	3761 (47.7%)	4681 (50.2%)	4867 (49.0%)	4553 (50.2%)
	Grade 0–2	2329 (61.9%)	2883 (57.9%)	4126 (52.3%)	4651 (49.8%)	5075 (51.0%)	4515 (49.8%)
Partner-ship	Not alone	2748 (72.5%)	3590 (71.5%)	5345 (67.2%)	6290 (66.9%)	6622 (66.0%)	5874 (64.5%)
	alone	1040 (27.5%)	1429 (28.5%)	2604 (32.8%)	3106 (33.1%)	3404 (34.0%)	3239 (35.5%)
Region in Europe	Northern	476 (12.6%)	582 (11.5%)	484 (6.0%)	840 (8.8%)	787 (7.8%)	646 (7.1%)
	Eastern	0	615 (12.1%)	2956 (36.7%)	2586 (27.1%)	2856 (28.2%)	2753 (30.2%)
	Southern	1788 (47.1%)	2190 (43.1%)	1760 (21.8%)	2706 (28.4%)	3633 (35.8%)	3250 (35.6%)
	Western	1529 (40.3%)	1690 (33.3%)	2864 (35.5%)	3395 (35.6%)	2871 (28.3%)	2481 (27.2%)
BMI (kg/m^2)		28.7±4.9	28.9±5.0	29.5±5.3	29.2±5.3	29.2±5.2	29.0±5.2
Age at diabetes onset	<30 years	81 (2.1%)	114 (2.3%)	232 (2.9%)	320 (3.4%)	358 (3.5%)	305 (3.3%)
	≥30 years	2172 (57.3%)	2872 (56.6%)	5026 (62.3%)	5998 (63.0%)	6458 (63.6%)	5809 (63.6%)
	Not reported	1540 (40.6%)	2091 (41.2%)	2806 (34.8%)	3209 (33.7%)	3331 (32.8%)	3016 (33.0%)
Times of visits to the doctor per year		10.0±13.0	10.1±12.1	9.5±11.1	9.9±11.7	9.7±12.3	10.3±23.1
Report of earlier diabetes	Yes	2253 (59.6%)	3203 (63.3%)	5947 (73.9%)	7349 (77.3%)	8352 (82.4%)	6629 (72.9%)
	No	1529 (40.4%)	1861 (36.7%)	2104 (26.1%)	2158 (22.7%)	1788 (17.6%)	2462 (27.1%)
Intake of diabetes medication	Yes	1853 (49.0%)	2739 (54.0%)	5162 (64.1%)	6588 (69.3%)	7514 (74.1%)	6435 (70.6%)
	No	1929 (51.0%)	2330 (46.0%)	2883 (35.8%)	2907 (30.6%)	2607 (25.7%)	2634 (28.9%)
Subjective health	Excellent	99 (2.6%)	176 (3.5%)	143 (1.8%)	174 (1.8%)	195 (1.9%)	150 (1.7%)
	Very good	400 (10.6%)	463 (9.1%)	575 (7.1%)	726 (7.6%)	704 (6.9%)	623 (6.8%)
	Good	1275 (33.7%)	1696 (33.5%)	2303 (28.6%)	2987 (31.4%)	2948 (29.1%)	2675 (29.3%)
	Fair	1401 (37.0%)	1814 (35.8%)	3168 (39.3%)	3616 (38.0%)	4045 (39.9%)	3574 (39.2%)
	Poor	609 (16.1%)	918 (18.1%)	1864 (23.1%)	2006 (21.1%)	2251 (22.2%)	2071 (22.7%)
Report of earlier heart disease	Yes	699 (18.5%)	986 (19.5%)	1786 (22.2%)	1892 (19.9%)	1989 (19.6%)	1854 (20.4%)
	No	3083 (81.5%)	4078 (80.5%)	6265 (77.8%)	7615 (80.1%)	8151 (80.4%)	7237 (79.6%)
Report of earlier stroke	Yes	184 (4.9%)	278 (5.5%)	578 (7.2%)	710 (7.5%)	715 (7.1%)	705 (7.7%)
	No	3598 (95.1%)	4786 (94.5%)	7473 (92.7%)	8797 (92.5%)	9425 (92.9%)	8386 (92.0%)

Means ± standard deviation, Proportions (N (%)), *ISCED* International Standard Classification of Education

and stroke. Persons taking diabetes medication at first self-report of diabetes had a lower risk of inconsistencies in diabetes self-report than persons not taking diabetes medication (RR = 0.53 (0.51–0.56)).

In the waves 1, 2, and 4 to 7, between 5.1 to 7.3% of those participants who said that they took diabetes medication said at the same time that they did not have diabetes (data not shown). Among 204 persons who said they had diabetes at wave 1, but not at wave 2, and who participated in at least one of the waves 4 to 7, 108 (52.9% (45.9–60.0)) said at wave 4 or later that they had diabetes. Accordingly, among 643 persons who said they had diabetes at wave 4, but not at wave 5, and who participated in at least one of the waves 6 and 7, 322 (50.1% (46.1–54.0)) said at wave 6 or 7 they had diabetes.

For persons starting in wave 1, the proportions of missing values for self-reported diabetes were 0.3% (wave 1), 10.3% (wave 2), 44.9% (wave 4), 36.6% (wave 5), 40.3% (wave 6), and 47.2% (wave 7) (proportions of missing values for persons starting in later waves not shown). After multiple imputation, the proportion of persons giving at least one inconsistent self-report increased to 39.6% (38.4%—40.8%).

Discussion

In wave 1 to 7 of SHARE, the proportion of participants who reported diabetes but denied it in at least one later wave was 33%. Participants who were obese, took anti-diabetic medication, or reported poor health less often gave inconsistent self-reports of diabetes. In contrast, persons who saw the doctor less often or who did not report their age at diabetes onset more often gave inconsistent reports of their condition. Moreover, inconsistencies appeared more often in women and older persons.

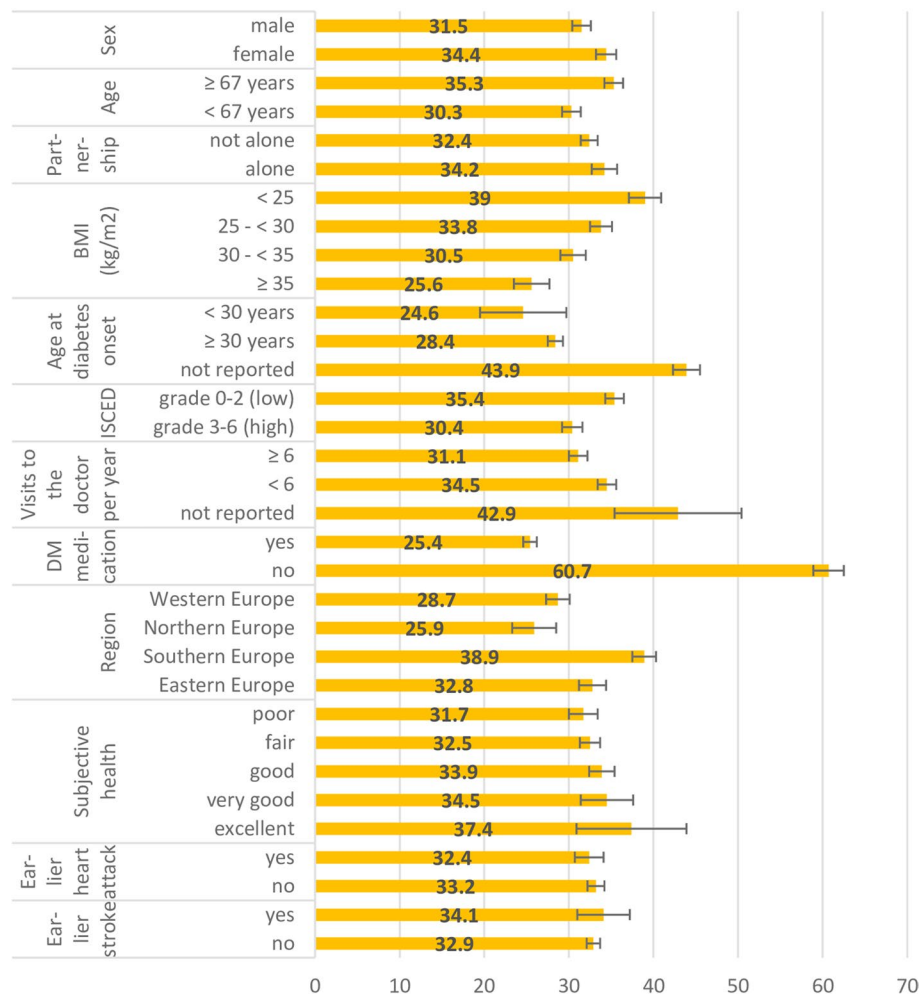## Proportion of inconsistencies (%, 95% CI) in report of diabetes status

Category	Stratum	Value
Sex	male	31.5
	female	34.4
Age	≥ 67 years	35.3
	< 67 years	30.3
Partnership	not alone	32.4
	alone	34.2
BMI (kg/m2)	< 25	39
	25 - < 30	33.8
	30 - < 35	30.5
	≥ 35	25.6
Age at diabetes onset	< 30 years	24.6
	≥ 30 years	28.4
	not reported	43.9
ISCED	grade 0-2 (low)	35.4
	grade 3-6 (high)	30.4
Visits to the doctor per year	≥ 6	31.1
	< 6	34.5
	not reported	42.9
DM medication	yes	25.4
	no	60.7
Region	Western Europe	28.7
	Northern Europe	25.9
	Southern Europe	38.9
	Eastern Europe	32.8
Subjective health	poor	31.7
	fair	32.5
	good	33.9
	very good	34.5
	excellent	37.4
Earlier heart attack	yes	32.4
	no	33.2
Earlier stroke	yes	34.1
	no	32.9

Fig. 1 Proportions of inconsistencies in self-reports of diabetes by strata of sociodemographic and diabetes related variables. CI: confidence interval; ISCED: International Standard Classification of Education; DM: diabetes mellitus

Earlier studies on reliability of self-reports of diabetes in panel studies mainly showed large inconsistencies, too. However, it is difficult to compare results from different studies because studies vary in the number of panel waves, time intervals between the waves, characteristics of study participants, the wording of questions about diabetes, and the definition of inconsistencies. As a matter of course, participants have more opportunities to give inconsistent reports of their condition when they take part in more panel waves – so more persons give inconsistent reports when they participate in more waves as it was also observed in the present study. The exact wording of the question whether a participant had diabetes strongly influences the proportion of inconsistencies. In the Health and Retirement Study (HRS), participants who had reported diabetes in an earlier wave were reminded of this earlier report in later waves [9]. It is sensible to assume that respondents do not want to give inconsistent reports in the presence of the interviewer, and, therefore, in the HRS, the proportion of persons giving inconsistent reports of diabetes was only 2.9%. When algorithms to estimate proportions of inconsistencies differ this also has a strong impact on the study results. For example, in a study on prostate cancer survivors, the denominator also included persons who reported that they had no diabetes in all waves they took part in [7], and inconsistent response patterns for diabetes were given by 7% of the respondents. However, recalculating this figure using the

Kowall *et al. BMC Medical Research Methodology*   (2024) 24:7

Page 6 of 9

Table 2 Relative risks for the associations between exposure variables and incidence of inconsistent self-report of diabetes

		N	Number of persons with inconsistent diabetes self-report	RR_crude (95% CI)	RR_adj (95% CI)
Sex	Female	6835	2349	1.09 (1.04 – 1.15)	1.08 (1.03 – 1.13) [a]
	Male	6344	1999	1	1
Age (per 10 years)		13,179	4348	1.14 (1.11 – 1.16)	1.15 (1.12 – 1.18) [a]
ISCED	Grade 3–6 (high)	6211	1885	0.86 (0.82 – 0.90)	0.96 (0.91 – 1.02) [b]
	Grade 0–2 (low)	6828	2416	1	1
Partnership	Not alone	9090	2944	0.95 (0.90 – 1.00)	0.98 (0.93 – 1.03) [c]
	Alone	3921	1340	1	1
Region	Northern Europe	1010	262	0.91 (0.81 – 1.01)	0.87 (0.77 – 0.97) [d]
	Eastern Europe	3548	1162	1.14 (1.07 – 1.22)	1.17 (1.10 – 1.25)
	Southern Europe	4431	1723	1.36 (1.28 – 1.44)	1.32 (1.24 – 1.41)
	Western Europe	4190	1201	1	1
BMI (kg/m^2)	< 25	2432	949	1	1
	25—< 30	5277	1781	0.86 (0.81 – 0.92)	0.88 (0.83 – 0.93) [e]
	30—< 35	3333	1017	0.78 (0.73 – 0.84)	0.81 (0.75 – 0.87)
	≥ 35	1627	417	0.66 (0.60 – 0.72)	0.70 (0.64 – 0.77)
Age at diabetes onset	< 30 years	260	64	0.87 (0.70 – 1.07)	0.88 (0.71 – 1.08) [f]
	≥ 30 years	8972	2552	1	1
	Not reported	3947	1732	1.54 (1.47 – 1.63)	1.38 (1.31 – 1.45)
Times of visit to the doctor (per 10 visits)		13,179	4348	0.97 (0.95 – 0.99)	0.97 (0.95 – 0.99) [g]
Intake of diabetes medication	Yes	10,344	2626	0.42 (0.40 – 0.44)	0.53 (0.51 – 0.56) [h]
	No	2827	1717	1	1
Subjective health	Excellent	214	80	1	1
	Very good	869	300	0.92 (0.76 – 1.12)	0.93 (0.80 – 1.10) [i]
	Good	3882	1317	0.91 (0.76 – 1.08)	0.91 (0.79 – 1.06)
	Fair	5271	1715	0.87 (0.73 – 1.04)	0.89 (0.77 – 1.03)
	Poor	2939	932	0.85 (0.71 – 1.02)	0.87 (0.75 – 1.01)
Report of earlier heart attack	Yes	2796	906	0.98 (0.92 – 1.04)	0.97 (0.91 – 1.03) [j]
	No	10,383	3442	1	1
Report of earlier stroke	Yes	918	313	1.04 (0.94 – 1.14)	1.05 (0.96 – 1.15) [j]
	No	12,261	4035	1	1

RR_crude: Crude relative risk; RR_adj Adjusted relative risk, *ISCED* International Standard Classification of Education

[a] Adjusted for the number of waves a participant took part in

[b] Adjusted for age, sex, region, number of waves

[c] Adjusted for age, sex, ISCED, region, number of waves

[d] Adjusted for age, sex, ISCED, partnership, number of waves

[e] Adjusted for age, sex, ISCED, region, partnership, number of waves

[f] Adjusted for region, number of waves

[g] Adjusted for age, sex, ISCED, region, partnership, earlier stroke, earlier heart attack, subjective health, number of waves

[h] Adjusted for age, sex, ISCED, region, number of waves, partnership, BMI, subjective health

[i] Adjusted for age, sex, ISCED, region, partnership, earlier stroke, earlier heart attack, diabetes medication, number of waves

[j] Adjusted for age, sex, ISCED, region, partnership, number of waves

definition of inconsistence applied in the present study led to 33.8% of persons giving inconsistent responses for that study.

In the present study, inconsistent responses were given less often when respondents had characteristics which make their condition more salient. This is the case when they take antidiabetic drugs, are obese, see the doctor more often, or report poor health. Moreover, in this study, inconsistencies were more frequent in persons who did not report the age of diabetes onset. It seems plausible that persons who do not know their age of diabetes onset have less involvement with their disease,

and are more likely to give inconsistent self-reports. Inconsistent self-report of diabetes was more frequent in participants with age at diabetes onset ≥ 30 years than in participants with age at onset < 30 years. This seems plausible because older patients with early onset of diabetes are confronted with their health condition for several decades, and should therefore be more aware of it. Persons with age at diabetes onset < 30 years may more often have type 1 diabetes, but age at onset is only a poor proxy to distinguish type 1 and type 2 diabetes. Associations between type of diabetes and inconsistent reporting of diabetes have so far only been investigated in the NOWAC Study showing only small differences between persons with type 1 and type 2, respectively [8]. However, in the NOWAC Study, the distinction between type 1 and type 2 diabetes was based solely on the age at diabetes onset [20]. In the present study, women, and older persons gave more inconsistent reports of diabetes. For age, this result is in line with results from other studies [6, 7, 9]. A decrease in consistent reports with older age may be explained by worsening of memory in older age in general, and, moreover, by an increased risk of cognitive decline in persons with diabetes [21–24]. For sex, results were mixed: in the HRS, and in the Danish Health and Morbidity Study, hardly any association between sex and inconsistent reporting was found, whereas women gave less inconsistent reports in NHANES I [6, 9, 11]. The result in the present study may be made more plausible by the possibility that women who had gestational diabetes during pregnancy may be confused about how to take this into account when answering the question whether a doctor had ever told them they had diabetes. To avoid this, 'outside pregnancy' should be added to the item. Like in the present study, associations with education were generally weak in other studies, and only in NOWAC and in the Taiwanese Survey of Health and Living Status of the Elderly, inconsistent reports of diabetes were given less often by the higher educated [6, 8].

Several reasons for inconsistent self-reports of diabetes are conceivable: (1) Participants with diabetes do not report the disease because they do not have any diabetes related complaints at the time of the interview. (2) Participants may be unwilling to report their health condition for fear of discrimination or out of shame. However, if these potential psychological barriers actually played a role, the question comes up why they had not prevented the same participants from reporting diabetes in an earlier interview. (3) Participants forgot the diagnosis given by their doctor which may be due to cognitive impairment in some older participants. (4) Their doctors may have given them a diagnosis of prediabetes which they misunderstood as a diagnosis of diabetes. In this case, the earlier self-report was wrong when they reported

diabetes in one wave but not any more in a later wave. However, positive predictive values of self-reported diabetes are usually assessed as high or very high, and are larger than 90% in many, albeit not in all studies [25–29]. (5) The diagnosis of diabetes given by the doctor was wrong. This is very unlikely because there are clear glucose and HbA1c based definitions of diabetes [30]. (6) Some participants may have lacked attention during the interviews as suggested by 5 – 7% of participants who did not report diabetes although they had reported taking antidiabetic drugs in the same interview.

As estimates of diabetes prevalence and incidence may be biased by inconstencies of diabetes self-report efforts must be taken to reduce the frequencies of these inconsisties. One measure is to asssess not only present diabetes or former diagnoses of diabetes but also the intake of antidiabetic medication and age at onset. Thus, corrections could be made for persons who report the latter but give no report of diabetes. Moreover, participants should be reminded of their earlier statements of diabetes („in the last interview you said that … ") – so participants with an earlier „yes " response and a later „no " response may either dispute their earlier „yes " response or think their later „no " response over. Furthermore, showing interviewees lists with 16 or more disases as done in SHARE and ask them which health conditions apply to them may be overwhelming for a part of the older participants. Instead, each disease should be addressed separately. In addition, questions about diabetes could be asked a second time. For example, in the Heinz Nixdorf Recall Study, participants are asked about diabetes once more when blood samples are taken [31]. Finally, the question about diabetes in SHARE should be worded more clearly. The additional explanation „ With this we mean that a doctor has told you that you have this condition, and that you are either currently being treated for or bothered by this condition " might bring participants who do not suffer from diabetes at the time of the interview to give a „no " response.

This study has advantages over previous studies. With an analysis data set of 13,179 participants with diabetes it is by far the largest study to examine inconsistencies of self-reports of diabetes in panel studies – earlier studies were either much smaller or focussed on a specific population [6–11]. Moreover, the data set included seven waves with time intervals of two years allowing some specific additional analyses (e.g., on correction of inconsistencies in later waves). Additionally, we took a wide range of determinants of inconsistencies into account some of which had not been adressed in earlier studies. This study also has limitations. There are no data from medical records like glucose and HbA1c values which would make it possible to look for false

Kowall *et al. BMC Medical Research Methodology*   (2024) 24:7

Page 8 of 9

positive and false negative self-reports of diabetes. Unawareness of diabetes is still frequent [32, 33], and blood parameters would help to improve and deepen the analyses of this study. As indicated above, a further limitation of this study was that type of diabetes was not asked.

Conclusions

Although diabetes is in principle not curable and should be reported again after a „yes " response in the past, inconsistencies in self-report of diabetes are frequent. Measures should be taken to make this phenomenon less likely in panel studies, in particular by wording questions about diabetes more clearly. Factors like obesity, intake of antidiabetic drugs or poor self-rated health which make diabetes more salient for a patient lead to lower frequencies of inconsistent self-reports of diabetes.

Availability of data and materials
SHARE data are available free of charge.

Declarations

Ethics approval and consent to participate
This study solely relies on publically available data. Therefore, consent to participate is not required.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

References
1. Martin LM, Leff M, Calonge N, Garrett C, Nelson DE. Validation of self-reported chronic conditions and health services in a managed care population. Am J Prev Med. 2000;18:215–8.
2. Wilhelmson K, Rubenowitz Lundin E, Andersson C, Sundh V, Waern M. Interviews or medical records, which type of data yields the best information on elderly people´s health status? Aging Clin Exp Res. 2006;18:25–33.
3. Van den Akker M, Van Steenkiste B, Krutwagen E, Metsemakers JFM. Disease or no disease? Disagreement on diagnoses between self-reports and medical records of adult patients. Eur J Gen Pract. 2015;21:45–51.
4. Dambha-Miller, H et al. Behaviour change, weight loss and remission of type 2 diabetes: a community based prospective cohort study. Diabetic Medicine; DOI: https://doi.org/10.1111/dme.14122
5. Lean ME, et al. Primary careled weight management for remission of type 2 diabetes (DiRECT): an open-label, cluster-randomised trial. Lancet. 2018;391(10120):541–51.
6. Beckett M, Weinstein M, Goldman N, Yu-Hsuan L. Do health interview subjects yield reliable data on chronic illness among older respondents? Am J Epidemiol. 2000;151:315–23.
7. Klabunde CN, Reeve BB, Harlan LC, Davis WW, Potosky AL. Do patients consistently report comorbid conditions over time? Results from the prostate cancer outcome study. Med Care. 2005;43:391–400.
8. Sheik MA, Lund E, Braaten T. Test-retest reliability of self-reported diabetes diagnosis in the Norwegian Women and Cancer Study: A population-based longitudinal study (n=33,919). SAGE Open Med 2016; 8;4:2050312115622857. doi: https://doi.org/10.1177/2050312115622857. eCollection 2016.
9. Cigolle CT, Nagel CL, Blaum CS, Liang J, Quinones AR. 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.
10. Quinones AR, Melekin A, Cigolle CT, Nagel CL. Disputes of self-reported chronic disease over time: the role of race, ethnicity, nativity and language of interview. Med Care. 2019;57:625–32.
11. Rosendahl Jensen HA, Davidsen M, Illemann Christensen A, Ekholm O. Inconsistencies in self-reported health conditions: results of a nationwide panel study. Int J Public Health. 2019;64:1243–6.
12. Börsch-Supan A, Brandt M, Hunkler C, et al. Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE). Int J Epidemiol. 2013;42:992–1001.
13. SHARE - Survey of Health, Ageing and Retirement in Europe. http://www.share-project.org/home0.html (last access december 30, 2021).
14. Börsch-Supan, A. (2020). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 1. Release version: 7.1.0. SHARE-ERIC. Data set. https://doi.org/10.6103/SHARE.w1.710
15. Börsch-Supan, A. (2020). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 2. Release version: 7.1.0. SHARE-ERIC. Data set. https://doi.org/10.6103/SHARE.w2.710
16. Börsch-Supan, A. (2020). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 4. Release version: 7.1.0. SHARE-ERIC. Data set. https://doi.org/10.6103/SHARE.w4.710
17. Börsch-Supan, A. (2020). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 5. Release version: 7.1.0. SHARE-ERIC. Data set. https://doi.org/10.6103/SHARE.w5.710
18. Börsch-Supan, A. (2020). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 6. Release version: 7.1.0. SHARE-ERIC. Data set. https://doi.org/10.6103/SHARE.w6.710
19. Börsch-Supan, A. (2020). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 7. Release version: 7.1.1. SHARE-ERIC. Data set. https://doi.org/10.6103/SHARE.w7.711
20. Nooney JG, Kirkman MS, Bullard MK, et al. Identifying optimal survey-based algorithms to distinguish diabetes type among adults with diabetes. J Clin Transl Endocrinol. 2020;21:100231.
21. Cheng G, Huang C, Deng H, Wang H. Diabetes as a risk factor for dementia and mild cognitive impairment: a meta-analysis of longitudinal studies. Intern Med J. 2012;42:484–91.

Kowall *et al. BMC Medical Research Methodology*   (2024) 24:7

Page 9 of 9

22. Zhang J, Chen C, Hua S, et al. An updated meta-analysis of cohort studies: Diabetes and risk of Alzheimer´s disease. Diabetes Res Clin Pract. 2017;124:41–7.
23. You Y, Lou Z, Chen Y, et al. The prevalence of mild cognitive impairment in type 2 diabetes mellitus patients: a systematic review and meta-analysis. Acta Diabetol. 2021;58:671–85.
24. Palta P, Schneider ALC, Biessels GJ, Touradji P, Hill-Briggs F. Magnitude of cognitive dysfunction in adults with type 2 diabetes: A meta-analysis of six cognitive domains and the most frequently reported neuropsychological tests within domains. J Int Neuropsychol Soc. 2014;20:278–91.
25. Huerta JM, Tormo MJ, Egea-Caparrós JM, Ortolá-Devesa JB, Navarro C. Accuracy of self-reported diabetes, hypertension, and hyperlipidemia in the adult Spanish population. DINO Study findings Rev Esp Cardiol. 2009;62:143–52.
26. Goto A, Morita A, Goto M, et al. Validity of diabetes self-reports in the Saku Diabetes Study. J Epidemiol. 2013;23:295–300.
27. Jackson JM, DeFor TA, Crain AL, et al. Validity of diabetes self-reports in the Women´s Health Initiative. Menopause. 2014;21:861–8.
28. Pastorino S, Richards M, Hardy R, et al. Validation of self-reported diagnosis of diabetes in the 1946 British birth cohort. Prim Care Diabetes. 2015;9:397–400.
29. Li HL, Fang J, Zhao LG, et al. Personal characteristics effects on validation of self-reported type 2 diabetes from a cross-sectional survey among Chinese adults. J Epidemiol. 2020;30:516–21.
30. American Diabetes Association. 2. Classification and diagnosis of diabetes: standards of Medical Care in Diabetes-2020. Diabetes Care. 2020;43(Suppl 1):S14–31.
31. Schmermund A, Möhlenkamp S, Stang A, et al. Assessment of clinically silent atherosclerotic disease and established and novel risk factors for predicting myocardial infarction and cardiac death in healthy middle-aged subjects: Rationale and design of the Heinz Nixdorf RECALL Study. Risk factors, evaluation of coronary calcium and lifestyle. Am Heart J. 2002;144:212–8.
32. Rathmann W, Haastert B, Icks A, Löwel H, Meisinger C, Holle R, et al. High prevalence of undiagnosed diabetes mellitus in Southern Germany: target populations for efficient screening. The KORA survey 2000. Diabetologia. 2003;46:182–9.
33. Kowall B, Rathmann W, Stang A, Bongaerts B, Kuss O, Herder C, Roden M, Quante A, Holle R, Huth C, Peters A, Meisinger C. Perceived risk of diabetes seriously underestimates actual diabetes risk: The KORA FF4 Study. PLoS ONE. 2017;12(1):e0171152.

Publisher's Note