

Original Article

The Agreement Between Diagnoses as Stated by Patients and Those Contained in Routine Health Insurance Data

Results of a Data Linkage Study

Felicitas Vogelgesang*, Roma Thamm*, Timm Frerk, Thomas G. Grobe, Joachim Saam, Catharina Schumacher, Julia Thom

Department 2: Epidemiology and Health Monitoring, Mental Health, Robert Koch Institute, Berlin, Germany: Felicitas Vogelgesang

Department 2: Epidemiology and Health Monitoring, Physical Health, Robert Koch Institute, Berlin, Germany: Dr. oec. troph. Roma Thamm

Department of Health Monitoring & Biometrics, aQua-Institute for Applied Quality Improvement and Research in Health Care GmbH, Göttingen, Germany: Timm Frerk

Department of Health Monitoring & Biometrics, aQua-Institute for Applied Quality Improvement and Research in Health Care GmbH, Göttingen, Germany: Dr. med. Thomas G. Grobe

BARMER Health Insurance, Berlin, Germany: Joachim Saam

BARMER Health Insurance, Berlin, Germany: Catharina Schumacher

Department 2: Epidemiology and Health Monitoring, Mental Health, Robert Koch Institute, Berlin, Germany: Dr. rer. medic. Julia Thom

*These authors share first authorship.

Summary

Background: The frequency of medical diagnoses is a figure of central importance in epidemiology and health services research. Prevalence estimates vary depending on the underlying data. For a better understanding of such discrepancies, we compared patients' diagnoses as reported by themselves in response to our questioning with their diagnoses as stated in the routine data of their health insurance carrier.

Methods: For 6558 adults insured by BARMER, one of the statutory health insurance carriers in Germany, we compared the diagnoses of various illnesses over a twelve-month period, as reported by the patients themselves in response to our questioning (October to December 2021), with their ICD-10-based diagnosis codes (Q4/2020–Q3/2021). The degree of agreement was assessed with two kappa values, sensitivity, and specificity.

Results: The patients' stated diagnoses of diabetes and hypertension agreed well or very well with their diagnosis codes, with kappa and PABAK values near 0.8, as well as very high sensitivity and specificity. Moderately good agreement with respect to kappa was seen for the diagnoses of heart failure (0.4), obesity, anxiety disorder, depression, and coronary heart disease (0.5 each). The

poorest agreement ($\kappa \leq 0.3$) was seen for post-traumatic stress disorder, alcohol-related disorder, and mental and somatoform disorder. Agreement was worse with increasing age.

Conclusion: Diagnoses as stated by patients often differ from those found in routine health insurance data. Discrepancies that can be considered negligible were found for only two of the 11 diseases that we studied. Our investigation confirms that these two sources of data yield different estimates of prevalence. Age is a key factor; further reasons for the discrepancies should be investigated, and avoidable causes should be addressed.

Cite this as

Vogelgesang F, Thamm R, Frerk T, Grobe TG, Saam J, Schumacher C, Thom J: The agreement between diagnoses as stated by patients and those contained in routine health insurance data—results of a data linkage study. *Dtsch Arztebl Int* 2024; 121: 141–7. DOI: 10.3238/arztebl.m2023.0250

The frequency of medical diagnoses (diagnostic prevalence) is a parameter of central importance in epidemiology and health services research. To estimate this frequency in the population, epidemiological survey studies as well as statutory health insurance claims data (SHI routine data) are used. Whereas medical diagnoses in survey studies are reported by the participants themselves, diagnostic prevalence in SHI data is based on diagnoses according to ICD-10 coding, the official classification for coding diagnoses in outpatient and inpatient care in Germany (1). Comparisons of these two data sources indicate differing estimates of nationwide diagnostic prevalence rates, which are particularly grave in the case of mental disorders (2–4). The heterogeneity of the evidence makes it difficult to provide a summarizing evaluation and to infer recommendations for policy and practice.

There is a lack of conclusive information for Germany on the extent to which the diagnoses of various diseases differ between the two data sources. Discrepancies here relate to people not reporting the diagnoses that have been documented for them or reporting medical diagnoses that have not been documented. Person-specific linkage of self-reported survey data and routine data enables a valid quantification and investigation of these discrepancies (5–8). By means of this type of linkage, the present article describes the extent to which a person's self-reporting of a medical diagnosis in a survey corresponds to the health insurance documentation for that person and the extent

Box

Measures of agreement: Cohen's kappa and PABAK

To assess the agreement between two categorical variables, Cohen's kappa (κ) and the prevalence- and bias-adjusted kappa (PABAK) are used.

$$\kappa = \frac{p_0 - p_1}{1 - p_1} \quad \text{PABAK} = 2 p_0 - 1$$

p_0 = observed agreement, p_1 = expected agreement

Sample calculation for diabetes (unweighted):

	Routine data (BARMER)		
	Diagnosis	No diagnosis	Marginal frequency
Data from the written survey	n (%)	n (%)	n (%)
Diagnosis	614 (9.6)	19 (0.3)	633 (9.9)
No diagnosis	251 (3.9)	5530 (86.2)	5781 (90.1)
Marginal frequency	865 (13.5)	5549 (86.5)	6414 (100)

The data from the two sources agree in $614 + 5530 = 6144$ cases. The observed agreement is $p_0 = 6144/6414 = 0.958 = 95.8\%$.

The expected agreement describes the proportion of data in agreement from the two data sources that would be obtained if one were to make a completely random allocation to the "diagnosis" and "no diagnosis" groups for a given marginal distribution (mathematically referred to as independence). Even then, one would expect a certain proportion of cases to agree. For the agreeing judgment "diagnosis" in both data sources, it is, for example: $9.9\% \times 13.5\% = 1.34\%$. Overall, the expected agreement is calculated as $p_1 = 9.9\% \times 13.5\% + 90.1\% \times 86.5\% = 79.3\%$.

Thus, the observed agreement is approximately 16.5% higher than would have been expected with random allocation (given the marginal distribution).

For diabetes, this yields: $\kappa = (0.958 - 0.793)/(1 - 0.793) = 0.165/0.207 = 0.797$.

If the data from the two data sources are in full agreement, the value of Cohen's κ is 1. A value of 0 means that the observed agreement does not differ from the expected agreement. Negative values indicate that the observed agreement is even lower than would have been expected in the case of random allocation.

To interpret κ , the agreement rating developed by Altman et al. is often used: < 0.20 , 'insufficient'; $0.21-0.40$, 'sufficient'; $0.41-0.60$, 'moderate'; $0.61-0.80$, 'good'; and $0.81-1.0$ 'very good' (21). Thus, the κ calculated for diabetes (unweighted) is on the border between good and very good.

When using κ values, one must bear in mind that they are related to the prevalence of the entity under consideration (the lower the prevalence, the lower κ tends to be), which is why PABAK values are recommended as a measure of agreement in the recent scientific literature (10, 22).

In the case of two forms, PABAK depends only on the observed agreement and, in the example for diabetes, is: $2 \times 0.958 - 1 = 0.916$. Its interpretation is often carried out in the same way as for Cohen's κ . PABAK values tend to be very high at very low prevalence rates.

to which various non-communicable diseases differ in this respect.

Methods

This investigation was based on the study "Optimierte Datenbasis für Public Mental Health: Datenlinkage-Studie zur Aufklärung von Diskrepanzen zwischen Befragungs- und Routinedaten" (Optimized database for public mental health: data linkage study to investigate discrepancies between survey and routine data) (OptDatPMH; funded by the Innovation Committee at the German Federal Joint Committee [Gemeinsamer Bundesausschuss]: 01VSF19015). The Ethics Committee of the Lower Saxony Medical Association, Germany, granted its approval for the study.

Sample

The study population was based on a sample of people aged over 18 years who were resident in Germany and had been insured by BARMER, one of the SHI carriers in Germany, for at least 12 months. The sample was stratified by 5-year age groups, sex, and German federal state. A total of 26,000 insured persons were randomly drawn from the strata, proportionate to the population distribution. In October 2021, the selected group was asked to complete a health questionnaire and, in a separate response, to consent to the linkage of pseudonymized survey data and BARMER data. Following two reminder letters, 7110 individuals returned their completed questionnaire (response rate: 27.3%); of these, data linkage was additionally possible for 6558 individuals (92.2%). In order to address a possible sample bias due to

selective participation, an adjustment weighting to the Federal German population distribution in relation to sex, age, region (federal state, East–West, Nielsen regions, district size) (as of 31.12.2020), and education was conducted (Microcensus 2018 [9]). The results of a comprehensive analysis of (non-)response will be published separately.

Surveying medical diagnoses in the questionnaire

The survey of medical diagnoses in the area of mental health (depression; anxiety disorder; post-traumatic stress disorder [PTSD]; somatoform disorder; dependence on, or harmful use of, alcohol [alcohol-related disorder]; any mental disorder) was carried out in the same way as the health monitoring conducted by the Robert Koch Institute (RKI) with the following question: “Have you ever been diagnosed with X by a doctor or psychotherapist?” If the answer was “yes,” the following question was asked: “Has X also occurred in the last 12 months?” For medical diagnoses of physical diseases (diabetes; cardiovascular disease or coronary heart disease [CHD]; heart attack; heart failure; hypertension; obesity), participants were asked the following: “Have you ever been diagnosed with X by a doctor?” If respondents answered yes, they were then asked the following question: “Has X also been present in the last 12 months?” In the case of hypertension, respondents were additionally asked whether they were using blood pressure-lowering drugs. The selection of diseases was made according to their public health relevance, on the basis of which they are prioritized for the surveillance of non-communicable diseases by the RKI.

Diagnostic information in routine data

The diagnoses documented in routine data are based on ICD-10 codes. The proportionally most significant portion of the diagnoses from outpatient care is available only on a quarterly basis, and the survey took place in the fourth quarter of 2021. Therefore, to compare the diagnoses in relation to the occurrence of the diseases or disorders in the preceding 12 months, the quarters 4/2020–3/2021 (n = 6558) were taken into consideration. Sensitivity analyses investigated the extent to which agreement in the diagnostic information changes when other time periods are considered (4/2020–4/2021 and 1/2021–4/2021). On the one hand, this was to investigate whether selecting a different quarter would yield better agreement for individuals who returned the survey documents later in quarter 4/2021. On the other, this also investigated whether it was merely the respondents’ recollection of the previous 12 months that was not entirely correct and, as such, whether diagnoses from somewhat longer ago were also reported. For the comparison of self-reported medical diagnoses in relation to lifetime (“ever”), the period of the preceding 10 years was used in the routine data (quarters 4/2011–3/2021, n = 5849). A disease was deemed to be documented if it had been coded in at least one quarter as an outpatient (MIQ inclusion criterion) or inpatient with the diagnostic certainty “confirmed” in the that period. Sensitivity analyses were performed to investigate the extent to which agreement in the diagnoses changed if documentation of the ICD-10 code in at least two cases of treatment (M2C) or at least two quarters (M2Q) was selected as an inclusion criterion in routine data.

Measures of agreement and statistical analyses

To measure the degree of agreement between patient-reported and SHI routine data, Cohen’s kappa (κ), the prevalence- and bias-adjusted kappa (PABAK) as well as sensitivity and specificity were calculated for each disease or disorder overall and separately for women and men and for different age groups (*Box*). For the question under investigation, sensitivity corresponds to the proportion of people with a documented diagnosis in the routine data who also reported the respective diagnosis in the survey (true positive). Specificity refers to the proportion of people not reporting a diagnosis relative to all people who do not have a diagnosis documented in the routine data (true negative). Term 1-specificity represents the proportion of people reporting a diagnosis in the survey although no diagnosis is documented (false positive) (10–12). However, it is important to bear in mind that in the present study, sensitivity and specificity are not defined in relation to a reference standard with optimal validity, as is usually the case. Instead, these two parameters relate to the extent to which a diagnosis documented in the BARMER claims data was reflected in the answers given by the insured persons surveyed. All analyses were performed using SAS version 9.4 statistical software. Further information on the methodology of the OptDatPMH study can be found in the *eMethods Section*.

Results

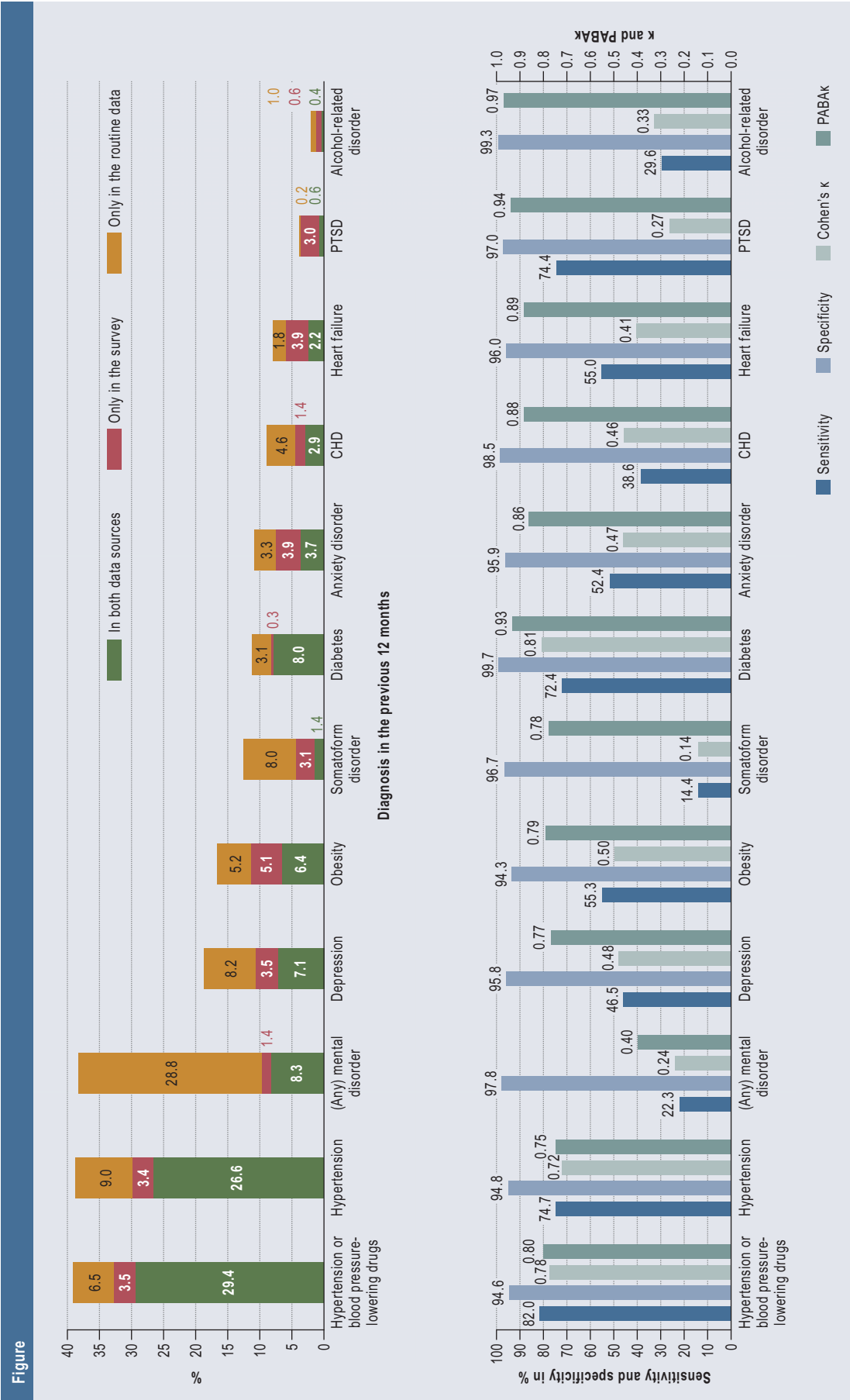
The *Figure* as well as *Tables 1–3* show the proportions of people for each of the following groups, in addition to the agreement measures:

- People for whom there is a diagnosis in both data sources
- People who report a diagnosis that is not documented in the routine data (only in the survey)
- People who have a diagnosis according to routine data but who do not report this (only in the routine data).

Agreement between diagnoses as stated by patients and those contained in routine data relating to the previous 12 months

For the majority of the 11 diseases or disorders, diagnoses are more frequently documented in routine data than reported by patients in the survey. However, the diagnoses for heart failure and PTSD are more frequently reported in the survey than they are documented (*Figure, eTable 1*).

The highest sensitivity is seen for the self-reporting of the diagnosis of high blood pressure. Taking the question on current use of antihypertensive drugs into account, 82 % of persons with a documented hypertension diagnoses in routine data reported this diagnosis in the survey. With sensitivities of between 74.4% and 38.6%, this was followed by patient-reported diagnoses of PTSD, diabetes, obesity, heart failure, anxiety disorder, depression, and CHD. Sensitivities of under 30% were seen for patients’ stated diagnoses of alcohol-related disorder, any type of mental disorder, and somatoform disorder. The latter was reported in the survey by only 14.4% of those with a documented diagnosis in the routine data. Specificity varied between 99.7% for diabetes and 94.3% for obesity. This means that out of all the people who do not have a



Proportion of medical diagnoses as stated by patients in the survey and/or contained in routine data as well as the agreement in diagnoses between the two data sources using Cohen's κ , PABAK, sensitivity, and specificity. The data relate to the 12 months preceding the time of the survey or to the 4th quarter of 2020–3rd quarter of 2021 (see data table in eTable 1 for illustration). The proportion of individuals identified in at least one of the two data sources is calculated as the sum of the three proportions listed. Sensitivity corresponds to the percentage of people with a documented diagnosis in the routine data who also state the respective diagnosis in the survey, in relation to all people who have the diagnosis documented in the routine data. Specificity is the percentage of people without a documented diagnosis in the routine data who also do not state a respective diagnosis in the survey, in relation to all people who do not have a corresponding diagnosis documented in the routine data.

κ , Cohen's kappa measure of agreement; PABAK, prevalence- and bias-adjusted kappa; PTSD, post-traumatic stress disorder

documented diagnosis of diabetes, only a very small proportion (0.3%) stated this diagnosis in the survey. For obesity, this percentage of people is 5.7%, closely followed by hypertension (5.4%).

The degrees of agreement between diagnoses in the two data sources also vary according to Cohen's κ . The highest κ values of around 0.8 were achieved for patient-reporting of diabetes and hypertension. This shows good to very good agreement between the two data sources. Moderate agreement was seen for self-reported obesity, depression, anxiety disorder, CHD, and heart failure. Sufficient agreement was found for patients' reporting of an alcohol-related disorder, PTSD, or any mental disorder. The lowest κ (0.14) was for somatoform disorder, for which there was insufficient agreement in the stated diagnoses.

Good to very good agreement based on the calculation of PABAK can be observed for patients' stated diagnoses for virtually all diseases and disorders considered. Only for the diagnosis of any mental disorder is there moderate agreement with a PABAK = 0.4 (Figure, eTable 1).

Differentiated results for measures of agreement relating to diseases or disorders in the preceding 12 months according to sex and age group are reported in eTable 2. While no relevant gender differences can be seen for sensitivity, specificity, κ , or PABAK, the degrees of agreement decline with increasing age (eTable 2).

Agreement between diagnoses as stated by patients and those contained in routine data relating to the previous 10 years

The measures of agreement as well as the frequencies of patients' stated diagnoses taken from the survey on diagnoses ever made by a physician and the documented diagnoses of people who had been continuously insured with BARMER over the preceding 10 years are shown in eTable 3 for 12 diseases or disorders (including heart attack). The measures and frequencies are given for the overall group as well as stratified by gender and age groups. Again, there is good to very good agreement for patient-reported diagnoses of diabetes and hypertension, as well as of heart attack. Patients' stated diagnoses of CHD, obesity, heart failure, depression, and anxiety disorder continue to show moderate agreement. Better agreement compared to the 12-month reference period was found for diagnoses as stated by patients for alcohol-related disorder. There is poorer agreement for patient reporting of diagnoses of any mental disorder. Agreement for patient-stated diagnoses of PTSD and somatoform disorder remains the lowest (eTable 3).

Sensitivity analyses

Varying the comparison periods in the routine data from Q4/2020–Q3/2021 to Q4/2020–Q4/2021 or to Q1/2021–Q4/2021 did not result in any relevant changes in Cohen's κ or PABAK. Sensitivities increased if the inclusion criteria M2C and M2Q were taken into consideration. However, since specificities simultaneously declined, Cohen's κ and PABAK changed only marginally (data not shown).

Discussion

To our knowledge, this study is the first in Germany to quantify the agreement between diagnoses as stated by patients and those contained in routine health insurance data for a variety of diseases. The results vary between the investigated diseases. For diabetes and hypertension, there is good to very good agreement, while agreement is moderate for obesity, heart failure, anxiety disorder, depression, and CHD. The lowest level of agreement was seen for patients' stated diagnoses of PTSD, alcohol-related disorder, and any mental or somatoform disorder. Thus, discrepancies are common between diagnoses as stated by the patients themselves and those contained in SHI routine data. Discrepancies that can be considered negligible were found for only two of the 11 diseases studied.

In concordance with our results, studies from Canada (6, 8, 13) and Korea (7) reported good agreement according to Cohen's κ between the two data sources for patient-stated diagnoses of diabetes (0.7–0.9) and hypertension (0.6–0.8), as well as moderate agreement for patient-reported CHD (0.5) (7, 13) and depression (0.5) (8).

The determination of four different parameters (Cohen's κ , PABAK, sensitivity, specificity) shows that they should be viewed together when comparing different diseases, since each parameter reflects specific aspects of the agreement between diagnoses as stated by patients and those contained in routine data and is therefore limited. For example, disorders with a very low prevalence, such as PTSD or alcohol-related disorders, showed very low agreement using Cohen's κ (0.27 and 0.33, respectively), while the calculation using PABAK resulted in very high values (0.94 and 0.97, respectively). The differences between Cohen's κ and PABAK tended to increase with decreasing disease prevalence. This indicates that the meaningfulness of these parameters is only limited for rare diseases.

Implications of the results relate not only to the possible causes of the discrepancies but also to the validity of the two data sources for epidemiology and health-care

research. If one assumes that patients are predominantly aware of their diagnoses and should be able to state these in a survey, avoidable causes of the observed discrepancies should be addressed to the extent possible. These causes include problems or shortcomings in the provision of medical information about a documented diagnosis, in patients' understanding and recollection of this information, and in their willingness to state the diagnosis in the survey. What also needs to be investigated is the extent to which problems with the validity of documented diagnoses in terms of the actual presence of a disease affect agreement. For example, coding quality of medical and psychotherapeutic diagnoses are the subject of controversy (14–16). For mental disorders, for example, a comparison of primary care diagnoses with the results of standardized assessments of the same individuals shows both under-reporting and over-reporting of disorders in routine data (17–20). Therefore, it is conceivable that the self-reporting of a diagnosis is also based on respondents' experience of their disease, which was not always medically diagnosed or documented. An accurate assessment of this type of misclassification could be achieved by a

survey and investigation study (in line with a gold standard) of standardized clinical examination and diagnosis in conjunction with the linkage of patient-specific routine data. On the routine data side, a comparison with medical records or even a patient follow-up examination could yield important insights.

The investigated sample of individuals insured by BARMER, one of the largest German SHI carriers, can be seen as a strength of this study. The fact that consent to personal data linkage was given in over 90% of cases, and by adjusting the weighting to the population distribution of adults living in Germany for sex, age, region, and education, means that it was possible to effectively counteract a possible systematic bias of the results due to selective participation.

A possible limiting factor is that the wording of the question regarding the occurrence of a disease does not allow a precise inference as to whether the disease has also been medically diagnosed in the preceding 12 months. The question asked whether the disease had also been present during that time period. However, one can assume that affected individuals who answered in the affirmative would also answer “yes” to a more precise formulation.

Conclusion

Data linkage enables a valid quantification of differences between diagnoses as stated by patients and those contained in routine health insurance data. When a variety of agreement measures were taken into consideration, frequent and strongly varying discrepancies with no clear pattern became apparent. For example, patients’ stated diagnoses of somatic disorders did not generally agree better than those of mental disorders. While agreement worsened with increasing age, there were no general differences according to sex. Changes to the 12-month reference period and higher requirements in terms of the criteria for inclusion in routine data did not affect the results. Against this background, the discrepancies found here between the data sources should be reflected in a disease-specific manner when using diagnoses as stated by patients. Only further research can reveal to what extent these discrepancies reflect under- or over-recording of morbidity or disease experience in routine data and whether self-reported medical diagnoses are, as a result, informative even in the absence of agreement.

Funding

The present study was undertaken as part of the project titled “Optimierte Datenbasis für Public Mental Health: Datenlinkage-Studie zur Aufklärung von Diskrepanzen zwischen Befragungs- und Routinedaten” (OptDatPMH), funded by the Innovation Fund of the Joint Federal Committee (Grant No.: 01VSF19015).

Conflict of interest statement

The authors declare that no conflict of interests exists.

Manuscript received on 3 July 2023, revised version accepted on 14 November 2023.

Translated from the original German by Christine Rye.

References

1. Bundesinstitut für Arzneimittel und Medizinprodukte: Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme, German Modification. www.bfarm.de/DE/Kodier-systeme/Klassifikationen/ICD/ICD-10-GM/_node.html (last accessed on 09 June 2023).

2. Frank J: Comparing nationwide prevalences of hypertension and depression based on claims data and survey data: an example from Germany. *Health Policy* 2016; 120: 1061–9.
3. Grobe TG, Kleine-Budde K, Bramesfeld A, Thom J, Bretschneider J, Hapke U: Prävalenzen von Depressionen bei Erwachsenen—eine vergleichende Analyse bundesweiter Survey- und Routinedaten. *Gesundheitswesen* 2019; 81: 1011–17.
4. Jacobi F, Bretschneider J, Müllender S: Veränderungen und Variationen der Häufigkeit psychischer Störungen in Deutschland—Krankenkassenstatistiken und epidemiologische Befunde. In: Kliner K, Rennert D, Richter M (eds.): *Gesundheit in Regionen—Blickpunkt Psyche*. BKK Gesundheitsatlas 2015. Berlin: Medizinisch wissenschaftliche Verlagsgesellschaft und BKK Dachverband 2015; S. 63–71.
5. March S, Andrich S, Drepper J, et al.: Gute Praxis Datenlinkage (GPD). *Gesundheitswesen* 2019; 81: 636–50.
6. Fortin M, Haggerty J, Sanche S, Almirall J: Self-reported versus health administrative data: implications for assessing chronic illness burden in populations. A cross-sectional study. *CMAJ Open* 2017; 5: e729–33.
7. Kim YY, Park JH, Kang HJ, Lee EJ, Ha S, Shin SA: Level of agreement and factors associated with discrepancies between nationwide medical history questionnaires and hospital claims data. *J Prev Med Public Health* 2017; 50: 294–302.
8. Payette Y, Moura CSd, Boileau C, Bernatsky S, Noisel N: Is there an agreement between self-reported medical diagnosis in the CARTaGENE cohort and the Québec administrative health databases? *Int J Pop Data Sci* 2020; 5(1): 1155.
9. Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder: Mikrozensus 2018, eigene Berechnungen. www.forschungsdatenzentrum.de/bestand/mikrozensus (last accessed on 09 June 2023).
10. Byrt T, Bishop J, Carlin JB: Bias, prevalence and kappa. *J Clin Epidemiol* 1993; 46: 423–9.
11. Grouven U, Bender R, Ziegler A, Lange S: Der Kappa-Koeffizient. *Dtsch Med Wochenschr* 2007; 132: e65–8.
12. Bland JM, Altman DG: Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999; 8: 135–60.
13. Lix LM, Yogendran MS, Shaw SY, Burchill C, Metge C, Bond R: Population-based data sources for chronic disease surveillance. *Chronic Dis Can* 2008; 29: 31–8.
14. Slagman A, Hoffmann F, Horenkamp-Sonntag D, Swart E, Vogt V, Herrmann WJ: Analyse von Routinedaten in der Gesundheitsforschung: Validität, Generalisierbarkeit und Herausforderungen. *Z Allg Med* 2023; 99: 86–92.
15. Drösl SE, Neukirch B: Evaluation der Kodierqualität von vertragsärztlichen Diagnosen. Gutachten im Auftrag der Kassenärztlichen Bundesvereinigung. www.kbv.de/media/sp/2014-11-18_Gutachten_Kodierqualitaet.pdf (last accessed on 07 November 2023).
16. IGES: Bewertung der Kodierqualität von vertragsärztlichen Diagnosen – Eine Studie im Auftrag des GKV-Spitzenverbands in Kooperation mit der BARMER GEK. Berlin: IGES Institut für Gesundheits- und Sozialforschung GmbH 2012. www.gkv-spitzenverband.de/media/dokumente/krankenversicherung_1/aerztliche_versorgung/verguetung_und_leistungen/klassifikationsverfahren/9_Endbericht_Kodierqualitaet_Hauptstudie_2012_12-19.pdf (last accessed 07 November 2023).
17. Sielk M, Altiner A, Janssen B, Becker N, Pilars M, Abholz HH: Prävalenz und Diagnostik depressiver Störungen in der Allgemeinarztpraxis. Ein kritischer Vergleich zwischen PHQ-D und hausärztlicher Einschätzung. *Psychiatr Prax* 2009; 36: 169–74.
18. Reitzle L, Köster I, Tuncer O, Schmidt C, Meyer I: Entwicklung und interne Validierung von Falldefinitionen für die Prävalenzschätzung mikrovaskulärer Komplikationen des Diabetes in Routinedaten. *Gesundheitswesen* 2023; 10.1055/a-2061–6954 (online ahead of print).
19. Piontek K, Shedden-Mora MC, Gladigau M, Kuby A, Löwe B: Diagnosis of somatoform disorders in primary care: diagnostic agreement, predictors, and comparisons with depression and anxiety. *BMC Psychiatry* 2018; 18: 361.
20. Mitchell AJ, Vaze A, Rao S: Clinical diagnosis of depression in primary care: a meta-analysis. *Lancet* 2009; 374: 609–19.
21. Kwicien R, Kopp-Schneider A, Blettner M: Concordance analysis: part 16 of a series on evaluation of scientific publications. *Dtsch Arztebl Int* 2011; 108: 515–21.
22. Feinstein AR, Cicchetti DV: High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol* 1990; 43: 543–9.

Corresponding author

Felicitas Vogelgesang
Robert Koch-Institut
Abteilung für Epidemiologie und Gesundheitsmonitoring
Fachgebiet Psychische Gesundheit
General-Pape-Str. 62–66, 12101 Berlin, Germany
VogelgesangF@rki.de

Cite this as:

Vogelgesang F, Thamm R, Frerk T, Grobe TG, Saam J, Schumacher C, Thom J: The agreement between diagnoses as stated by patients and those contained in routine health insurance data—results of a data linkage study. Dtsch Arztebl Int 2024; 121: 141–7. DOI: 10.3238/arztebl.m2023.0250

Supplementary material

eMethods Section, eTables:

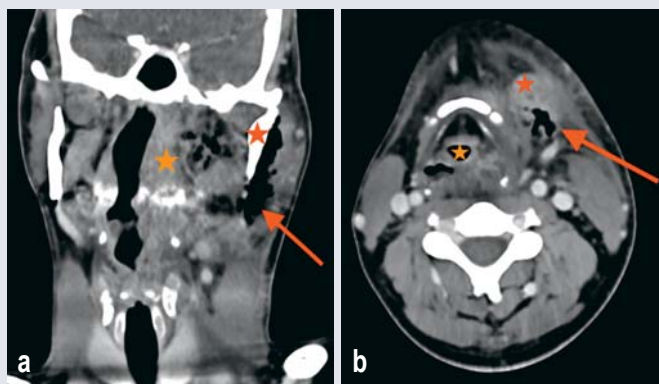
www.aerzteblatt-international.de/m2023.0250

CLINICAL SNAPSHOT



Cervicofacial Necrotizing Fasciitis of Non-Odontogenic Origin

A 38-year-old male patient with no preexisting diseases presented as an emergency with swelling in the region of the left cervicofacial transition. He had been experiencing difficulty swallowing for 3 days. The patient history included occasional drug use (amphetamines). Clinical examination confirmed a pressure-sensitive swelling in this patient, who had tachycardia and a high fever. Laboratory tests showed elevated inflammatory parameters (leukocytes $27.51 \times 10^9/L$, CRP 337 mg/L). Computed tomography revealed a pronounced inflammatory event with gas inclusions in the left perimandibular area arising from the submandibular gland and consistent with cervicofacial necrotizing fasciitis (CNF) (Figure a, b). Multiple surgical debridements were performed, which in combination with perioperative antibiotic therapy (piperacillin/tazobactam), led to an abatement of the inflammatory event. CNF is a life-threatening soft tissue infection and frequently of odontogenic origin. Cervicofacial manifestations are rare (4%). Immuno-deficiency and drug use are risk factors for CNF. Immediate surgical intervention together with antibiotic therapy form the treatment pillars of this clinical picture (40% mortality rate). This case report highlights the fact that CNF can also be of non-odontogenic origin.



Figure

- Coronal computed tomography shows an inflammatory process in a street-like configuration with gas inclusions (red arrow) at the left perimandibular cervicofacial transition (red asterisk) accentuated over a maximum extension in length of 11 cm with accompanying displacement (orange asterisk) of the oropharynx in the contralateral direction.
- Axial computed tomography reveals increased contrast enhancement in the region of the left submandibular gland (red asterisk) with perifocal gas inclusions (red arrow) and accompanying oropharyngeal displacement (orange asterisk) in the contralateral direction.

Images provided by: Abteilung für Radiologie, Bundeswehrkrankenhaus Hamburg-Wandsbek

Dr. med. Florian Dudde, Dr. med. dent. Martin Kreibich, Prof. Dr. med. Dr. med. dent. Kai-Olaf Henkel

Abteilung für Mund-, Kiefer- und plastische Gesichtschirurgie, Bundeswehrkrankenhaus Hamburg-Wandsbek, floriandudde@gmx.de

Conflict of interest statement: The authors state that no conflict of interest exists.

Translated from the original German by Christine Rye.

Cite this as: Dudde F, Kreibich M, Henkel KO: Cervicofacial necrotizing fasciitis of non-odontogenic origin. Dtsch Arztebl Int 2024; 121: 147. DOI: 10.3238/arztebl.m2022.0407

Supplementary material to the article

The Agreement Between Diagnoses as Stated by Patients and Those Contained in Routine Health Insurance Data

Results of a Data Linkage Study

by Felicitas Vogelgesang*, Roma Thamm*, Timm Frerk, Thomas G. Grobe, Joachim Saam, Catharina Schumacher, and Julia Thom

Dtsch Arztebl Int 2024; 121: 141–7. DOI: 10.3238/arztebl.m2023.0250

METHODS

Sampling

The aim of the sampling procedure was to determine a population-based sample of individuals insured by BARMER who were representative for Germany with regard to the characteristics of sex, age, and region of residence. In October 2021, this group of people participated in the postal survey of the OptDatPMH study. Only insured persons aged ≥ 18 years for whom precise information was available regarding birth and sex details, place of residence in Germany, and 92 insured days from October 1 to December 31, 2020 without evidence of subsequent termination of insurance ($n = 6,815,683$) were taken into consideration in the data available for sampling in 2020. All pre-selected insured individuals were allocated to one of a total of 512 strata according to sex, 5-year age group (18–19 years, 20–24 years, 25–29 years, etc., up to 90 years and older), and place of residence (differentiated according to 16 federal states). For the letter regarding the main survey, 26,000 insured persons were randomly selected from the 512 strata in such a way that the proportionate allocation of the strata among those contacted corresponded to the distribution of the average population in Germany in 2020 across the corresponding strata according to the Federal Statistical Office (*Statistisches Bundesamt*). Immediately prior to the letter, BARMER carried out an internal check of the current insurance status of those selected to receive the letter. Insured persons that could no longer be contacted due to having left the health insurance or death, as registered in the meantime, were replaced by other persons from the same stratum if necessary. Thus, in October 2021, precisely 26,000 insured individuals were contacted by letter for the survey. The distribution of structural characteristics of those contacted by letter was representative according to the population data available at the time of the survey. For subsequent analyses, weightings and standardizations were then based on population data for 2021, which, however, did not become available until summer 2022.

The survey

Together with the written invitation to take part in the study, prospective participants received study information, a data-linkage consent form, and an 18-page questionnaire. Based on the pre-test, it was assumed that the questions would take 20–30 min to answer. Two envelopes were provided for return postage in order that the consent form and the questionnaire documents could be further processed separately. Questionnaire documents were read by the aQua Institute;

the forms with written consent to data linkage were processed by the BARMER's trust center.

Response

Following a one-page written reminder in calendar week (CW) 43, followed by a second reminder in CW 48, which once again included all questionnaire documents, 7110 completed questionnaires were received by 11 January 2022 (excluding duplicates). As such, a response rate of 27.3% was achieved. In addition, 6581 questionnaires were accompanied by data linkage consent. A comparison of age and sex details in the survey with information in the routine data showed differences for 23 people, which suggested that a different person had completed the survey documents than the person to whom the questionnaire had been addressed. These individuals were not included in the data linkage, meaning that the data linkage was ultimately carried out for 6558 individuals. Linkage and analysis of the data took place exclusively in BARMER's secure data warehouse.

An analysis of response according to age, sex, and federal state was conducted and will be published separately. Differences between responders and non-responders will be discussed there in more detail. For the present analyses, possible discrepancies due to non-response were adjusted through weighting. A comparison of diagnostic frequency between participants and non-participants was also performed and will be published separately.

Weighting

Why use weighting?

Older rather than younger people tend to take part in surveys. This group not only has more but also different diseases compared to younger people. Moreover, the analyses show that agreement between the two data sources worsens with increasing age. If the different responses in the age groups had not been adjusted by weighting, this would have resulted in a skewed picture in which overall agreement would have been poorer and the proportions of disease higher. Mindful of the fact that a sample of persons insured by a health insurance carrier is not a population-wide, representatively drawn sample (for example, a sample drawn from residents' registration offices), no nationwide prevalences are reported, but rather, proportions are shown that have been adjusted for selective willingness to participate according to age, sex, region of residence, and education.

Application of weighting

The weighting adjusts to the German population by sex, age group (same grouping as in the sampling procedure), and region. This process is iterative and takes into consideration the following rough allocations, differentiated according to region:

- Federal state by age groups
- West–East by education
- Nielsen areas and district type.

Assignment of diseases to ICD-10 codes

Assignment was based on the results of previous methodological projects conducted at the Robert Koch Institute (RKI) and on the expertise of the co-authors from the aQua Institute.

eTable 1

Proportion of medical diagnoses as stated by patients in the survey and/or contained in routine data as well as the agreement in diagnoses between the two data sources using Cohen's κ, PABAK, sensitivity, and specificity*

Disease/disorder		Proportion (% , weighted) and number (n) with a diagnosis of the disease/disorder												
		In both data sources			Only in the survey		Only in the routine data							
		%	n		%	n	%	n	Sensitivity % [95% CI]	Specificity % [95% CI]	κ [95% CI]	PABAK [95% CI]	n Total	
	ICD-10 codes													
Hypertension and medication	I10–I15	29.4	2242	3.5	215	6.5	488	82.0 [80.1; 83.9]	94.6 [93.7; 95.4]	0.78 [0.76; 0.80]	0.80 [0.78; 0.82]	6367		
Hypertension	I10–I15	26.6	1959	3.4	196	9.0	686	74.7 [72.8; 76.6]	94.8 [94.0; 95.6]	0.72 [0.70; 0.74]	0.75 [0.73; 0.77]	6213		
PTSD	F43.1	0.6	34	3.0	175	0.2	17	74.4 [61.1; 87.8]	97.0 [96.5; 97.5]	0.27 [0.18; 0.35]	0.94 [0.92; 0.95]	6105		
Diabetes mellitus	E10–E14	8.0	614	0.3	19	3.1	251	72.4 [68.8; 75.9]	99.7 [99.5; 99.8]	0.81 [0.78; 0.84]	0.93 [0.92; 0.94]	6414		
Obesity	E66	6.4	406	5.1	291	5.2	341	55.3 [51.2; 59.5]	94.3 [93.4; 95.1]	0.50 [0.46; 0.54]	0.79 [0.78; 0.81]	5965		
Heart failure	I50	2.2	174	3.9	299	1.8	139	55.0 [48.8; 61.1]	96.0 [95.4; 96.5]	0.41 [0.36; 0.46]	0.89 [0.87; 0.90]	6114		
Anxiety disorder	F40, F41	3.7	194	3.9	235	3.3	226	52.4 [46.5; 58.3]	95.9 [95.2; 96.5]	0.47 [0.41; 0.52]	0.86 [0.84; 0.87]	6135		
Depression	F32, F33.0–F33.3, F33.8–F33.9, F34.1 (excluding F30, F31)	7.1	407	3.5	188	8.2	569	46.5 [42.8; 50.2]	95.8 [95.0; 96.6]	0.48 [0.44; 0.52]	0.77 [0.75; 0.79]	6080		
Coronary heart disease	I20–I25	2.9	241	1.4	103	4.6	371	38.6 [34.3; 42.9]	98.5 [98.1; 98.9]	0.46 [0.42; 0.51]	0.88 [0.87; 0.89]	6121		
Alcohol-related disorder	F10	0.4	24	0.6	38	1.0	67	29.6 [17.2; 42.1]	99.3 [99.1; 99.6]	0.33 [0.20; 0.47]	0.97 [0.96; 0.97]	6432		
Mental disorder	F00–F99	8.3	482	1.4	66	28.8	1863	22.3 [20.2; 24.3]	97.8 [97.1; 98.4]	0.24 [0.21; 0.26]	0.40 [0.37; 0.43]	6188		
Somatoform disorder	F45	1.4	89	3.1	185	8.0	537	14.4 [11.0; 17.9]	96.7 [96.0; 99.3]	0.14 [0.10; 0.19]	0.78 [0.76; 0.80]	6047		

* sorted in order of descending sensitivity; the data relate to the 12 months preceding the time of the survey or to the 4th quarter of 2020–3rd quarter of 2021

The proportion of individuals identified in at least one of the two data sources is calculated as the sum of the three proportions listed.

Sensitivity corresponds to the percentage of people with a documented diagnosis in the routine data who also state the respective diagnosis in the survey, in relation to all people who have their diagnosis documented in the routine data. Specificity is the percentage of people without a documented diagnosis in the routine data who also do not state a respective diagnosis in the survey, in relation to all people who do not have a corresponding diagnosis documented in the routine data.

κ, Cohen's kappa measure of agreement; CI, confidence interval; PABAK, prevalence- and bias-adjusted kappa; PTSD, post-traumatic stress disorder

eTable 2

Proportion of medical diagnoses as stated by patients in the survey and/or contained in routine data with regard to the previous 12 months as well as agreement between the two data sources using Cohen's κ, PABAK, sensitivity, and specificity by sex and age group*

Disease/disorder (ICD-10 codes)	Sex Age group	Proportion (%; weighted) and number (n) with a diagnosis of the disease/disorder								Sensitivity % [95% CI]	Specificity % [95% CI]	κ [95% CI]	PABAK [95% CI]	n
		In both data sources		Only in the survey		Only in the routine data								
		%	n	%	n	%	n							
Hypertension + medication (I10-I15)	Total	29.4	2242	3.5	215	6.5	488	82.0 [80.4; 83.6]	94.6 [93.8; 95.3]	0.78 [0.76; 0.80]	0.80 [0.79; 0.82]	6367		
	Men	30.4	1141	3.9	106	6.5	257	82.3 [79.9; 84.7]	93.8 [92.5; 95.2]	0.77 [0.75; 0.80]	0.79 [0.77; 0.82]	2798		
	Women	28.4	1101	3.1	109	6.4	231	81.7 [79.5; 83.8]	95.2 [94.3; 96.2]	0.79 [0.76; 0.81]	0.81 [0.79; 0.83]	3583		
	30–44 Years	10.4	88	3.3	33	3.3	30	76.0 [67.2; 84.9]	96.2 [94.7; 97.6]	0.72 [0.64; 0.80]	0.87 [0.83; 0.90]	1013		
	45–64 Years	30.7	715	4.6	103	5.2	130	85.5 [82.8; 88.1]	92.8 [91.3; 94.3]	0.79 [0.76; 0.81]	0.80 [0.78; 0.83]	2349		
	65–79 Years	58.4	959	2.7	48	12.4	201	82.5 [80.3; 84.7]	90.8 [88.1; 93.5]	0.67 [0.63; 0.71]	0.70 [0.66; 0.73]	1668		
	≥ 80 Years	66.4	471	2.0	16	17.6	122	79.0 [75.3; 82.8]	87.7 [80.5; 94.9]	0.48 [0.40; 0.56]	0.61 [0.54; 0.68]	717		
Hypertension (I10-I15)	Total	26.6	1959	3.4	196	9.0	686	74.7 [72.8; 76.6]	94.8 [94.0; 95.6]	0.72 [0.70; 0.74]	0.75 [0.73; 0.77]	6213		
	Men	27.4	988	3.8	96	9.6	374	74.0 [71.2; 76.9]	94.0 [92.6; 95.3]	0.70 [0.67; 0.73]	0.73 [0.70; 0.76]	2720		
	Women	25.8	971	3.0	100	8.4	312	75.4 [72.8; 78.0]	95.5 [94.6; 96.4]	0.74 [0.71; 0.76]	0.77 [0.75; 0.80]	3493		
	30–44 Years	9.5	83	3.4	33	4.0	31	70.4 [59.6; 81.1]	96.1 [94.6; 97.6]	0.68 [0.59; 0.77]	0.85 [0.81; 0.89]	990		
	45–64 Years	28.8	644	4.5	95	7.2	174	80.0 [77.0; 83.0]	93.0 [91.6; 94.5]	0.74 [0.71; 0.78]	0.77 [0.74; 0.80]	2275		
	65–79 Years	51.0	822	2.5	42	19.4	308	72.4 [69.5; 75.4]	91.6 [89.0; 94.2]	0.55 [0.51; 0.59]	0.56 [0.52; 0.61]	1629		
	≥ 80 Years	60.1	402	1.7	12	23.3	168	72.1 [68.1; 76.1]	89.8 [82.9; 96.7]	0.41 [0.33; 0.48]	0.50 [0.43; 0.57]	692		
PTSD (F43.1)	Total	0.6	34	3.0	175	0.2	17	74.4 [62.3; 86.6]	97.0 [96.4; 97.5]	0.27 [0.17; 0.36]	0.94 [0.92; 0.95]	6105		
	Men	0.2	4	2.3	58	0.1	2	81.2 [44.1; 100]	97.6 [96.9; 98.4]	0.16 [0.00; 0.33]	0.95 [0.94; 0.97]	2678		
	Women	1.0	30	3.7	117	0.4	15	73.0 [59.0; 87.0]	96.3 [95.5; 97.1]	0.31 [0.21; 0.42]	0.92 [0.90; 0.94]	3427		
	30–44 Years	1.0	9	2.7	23	0.3	4	77.4 [52.9; 100]	97.3 [96.1; 98.5]	0.39 [0.16; 0.63]	0.94 [0.92; 0.97]	975		
	45–64 Years	0.6	15	4.1	91	0.2	7	73.1 [55.4; 90.7]	95.8 [94.8; 96.9]	0.20 [0.10; 0.31]	0.91 [0.89; 0.93]	2210		
Diabetes mellitus (E10-E14)	Total	8.0	614	0.3	19	3.1	251	72.4 [68.8; 75.9]	99.7 [99.5; 99.8]	0.81 [0.78; 0.84]	0.93 [0.92; 0.94]	6414		
	Men	10.4	391	0.2	7	3.4	141	75.5 [71.2; 79.7]	99.8 [99.6; 100]	0.83 [0.80; 0.86]	0.93 [0.92; 0.94]	2805		
	Women	5.7	223	0.4	12	2.8	110	67.5 [61.8; 73.2]	99.6 [99.4; 99.8]	0.77 [0.72; 0.81]	0.94 [0.93; 0.95]	3609		
	30–44 Years	3.0	26	0.3	4	0.8	9	79.3 [65.4; 93.2]	99.7 [99.4; 100]	0.85 [0.75; 0.95]	0.98 [0.97; 0.99]	1022		
	45–64 Years	7.3	160	0.3	5	2.2	44	76.5 [69.8; 83.1]	99.7 [99.5; 100]	0.84 [0.79; 0.89]	0.95 [0.94; 0.96]	2363		
Obesity (E66)	65–79 Years	18.0	287	0.6	7	6.6	114	73.0 [68.1; 77.9]	99.3 [98.7; 99.9]	0.79 [0.75; 0.83]	0.86 [0.83; 0.88]	1668		
	≥ 80 Years	17.7	137	0.4	2	11.1	84	61.6 [54.0; 69.1]	99.5 [98.7; 100]	0.69 [0.62; 0.76]	0.77 [0.72; 0.82]	716		
	Total	6.4	406	5.1	291	5.2	341	55.3 [51.1; 59.5]	94.3 [93.5; 95.0]	0.50 [0.46; 0.54]	0.79 [0.78; 0.81]	5965		
	Men	5.5	144	4.7	110	5.4	174	50.4 [43.5; 57.4]	94.7 [93.6; 95.9]	0.46 [0.40; 0.53]	0.80 [0.77; 0.83]	2563		

Heart failure (I50)	Women	7.3	262	5.4	181	5.0	167	59.4 [53.9; 64.9]	93.8 [92.8; 94.8]	0.52 [0.48; 0.57]	0.79 [0.77; 0.82]	3402
	18–29 Years	2.6	13	4.1	16	3.4	16	43.2 [22.0; 64.4]	95.7 [93.3; 98.1]	0.37 [0.17; 0.57]	0.85 [0.79; 0.91]	632
	30–44 Years	7.2	71	5.7	54	2.7	27	72.4 [62.9; 82.0]	93.7 [91.9; 95.4]	0.58 [0.50; 0.67]	0.83 [0.79; 0.87]	987
	45–64 Years	8.1	196	5.3	120	4.9	105	62.4 [56.0; 68.8]	93.9 [92.8; 95.1]	0.56 [0.50; 0.61]	0.80 [0.77; 0.82]	2260
	65–79 Years	7.0	104	5.5	81	10.2	151	40.8 [34.0; 47.6]	93.4 [91.9; 94.9]	0.38 [0.31; 0.45]	0.69 [0.65; 0.73]	1495
	≥ 80 Years	3.1	22	3.4	20	7.2	42	29.9 [17.1; 42.6]	96.2 [94.4; 98.1]	0.31 [0.18; 0.45]	0.79 [0.73; 0.85]	591
	Total	2.2	174	3.9	299	1.8	139	55.0 [48.1; 61.8]	96.0 [95.4; 96.5]	0.41 [0.36; 0.46]	0.89 [0.87; 0.90]	6114
	Men	2.5	108	4.1	166	2.0	81	55.2 [47.0; 63.3]	95.7 [94.9; 96.5]	0.41 [0.35; 0.48]	0.88 [0.86; 0.89]	2686
	Women	2.0	66	3.7	133	1.7	58	54.7 [44.4; 65.1]	96.2 [95.5; 96.9]	0.41 [0.32; 0.49]	0.89 [0.88; 0.91]	3428
	45–64 Years	1.7	34	2.5	58	0.9	23	65.4 [52.1; 78.8]	97.4 [96.7; 98.2]	0.48 [0.36; 0.60]	0.93 [0.92; 0.95]	2292
	65–79 Years	3.7	66	8.3	131	5.2	69	41.7 [32.3; 51.1]	90.9 [89.4; 92.4]	0.28 [0.21; 0.36]	0.73 [0.69; 0.77]	1554
	≥ 80 Years	12.5	72	13.5	92	7.4	43	62.9 [51.5; 74.3]	83.1 [79.3; 87.0]	0.41 [0.31; 0.52]	0.58 [0.51; 0.66]	630
	Total	3.7	194	3.9	235	3.3	226	52.4 [46.9; 57.8]	95.9 [95.2; 96.5]	0.47 [0.42; 0.52]	0.86 [0.84; 0.87]	6135
	Men	2.8	59	3.1	84	2.0	72	58.5 [48.7; 68.2]	96.7 [95.9; 97.6]	0.50 [0.41; 0.59]	0.90 [0.88; 0.92]	2689
	Women	4.5	135	4.6	151	4.6	154	49.2 [42.2; 56.2]	95.0 [94.1; 95.9]	0.44 [0.38; 0.51]	0.82 [0.79; 0.84]	3446
	18–29 Years	4.7	23	3.7	17	2.4	13	65.8 [46.0; 85.6]	96.0 [93.8; 98.2]	0.57 [0.40; 0.75]	0.88 [0.83; 0.93]	614
	30–44 Years	4.1	41	2.3	23	1.8	19	69.2 [55.4; 82.9]	97.5 [96.5; 98.6]	0.64 [0.53; 0.75]	0.92 [0.89; 0.95]	981
	45–64 Years	3.9	85	5.0	112	3.3	72	54.4 [46.2; 62.5]	94.6 [93.5; 95.7]	0.44 [0.37; 0.51]	0.83 [0.81; 0.86]	2257
	65–79 Years	2.5	33	3.6	52	5.4	87	31.5 [23.0; 40.1]	96.1 [95.0; 97.3]	0.31 [0.22; 0.40]	0.82 [0.79; 0.85]	1618
	≥ 80 Years	2.2	12	4.3	31	5.3	35	29.8 [13.4; 46.1]	95.3 [93.4; 97.2]	0.27 [0.10; 0.43]	0.81 [0.76; 0.86]	665
Depression F32, F33.0–F33.3, F33.8–F33.9, F34.1 (excluding F30, F31)	Total	7.1	407	3.5	188	8.2	569	46.5 [42.8; 50.2]	95.8 [95.0; 96.6]	0.48 [0.44; 0.52]	0.77 [0.75; 0.79]	6080
	Men	5.8	137	2.8	69	6.0	209	49.3 [42.7; 56.0]	96.8 [95.9; 97.8]	0.52 [0.46; 0.59]	0.82 [0.80; 0.85]	2677
	Women	8.4	270	4.2	119	10.4	360	44.8 [40.3; 49.3]	94.8 [93.5; 96.1]	0.45 [0.40; 0.50]	0.71 [0.68; 0.74]	3403
	18–29 Years	7.3	41	3.7	21	3.0	18	71.0 [56.9; 85.1]	95.9 [93.8; 98.0]	0.65 [0.52; 0.78]	0.87 [0.82; 0.92]	614
	30–44 Years	7.0	62	4.2	32	4.4	44	61.6 [51.3; 71.8]	95.3 [93.5; 97.1]	0.57 [0.47; 0.67]	0.83 [0.79; 0.87]	967
	45–64 Years	8.8	198	3.8	77	9.5	217	47.9 [42.7; 53.1]	95.4 [94.3; 96.5]	0.49 [0.44; 0.55]	0.73 [0.70; 0.77]	2226
	65–79 Years	5.6	81	2.9	41	11.8	191	32.3 [25.9; 38.7]	96.5 [95.4; 97.7]	0.36 [0.29; 0.43]	0.71 [0.67; 0.75]	1600
	≥ 80 Years	3.6	25	1.9	17	15.9	99	18.4 [11.0; 25.8]	97.6 [96.6; 98.6]	0.22 [0.12; 0.32]	0.64 [0.58; 0.71]	673
	Total	2.9	241	1.4	103	4.6	371	38.6 [33.6; 43.7]	98.5 [98.1; 98.8]	0.46 [0.41; 0.51]	0.88 [0.87; 0.89]	6121
	Men	4.3	183	1.3	48	6.1	250	41.1 [35.2; 47.0]	98.5 [98.0; 99.1]	0.50 [0.44; 0.56]	0.85 [0.83; 0.87]	2659
Coronary heart disease (I20–I25)	Women	1.6	58	1.5	55	3.1	121	33.3 [24.7; 41.9]	98.4 [98.0; 98.9]	0.38 [0.29; 0.47]	0.91 [0.89; 0.92]	3462
	45–64 Years	2.3	48	0.6	15	3.0	67	43.2 [32.8; 53.6]	99.3 [98.9; 99.7]	0.54 [0.43; 0.64]	0.93 [0.91; 0.94]	2290
	65–79 Years	6.7	112	2.9	43	12.6	198	34.8 [28.5; 41.2]	96.4 [95.3; 97.6]	0.39 [0.32; 0.46]	0.69 [0.65; 0.73]	1567
	≥ 80 Years	11.8	80	7.1	41	15.8	99	42.9 [34.0; 51.7]	90.2 [86.9; 93.5]	0.37 [0.27; 0.46]	0.54 [0.47; 0.62]	622
	Total	2.9	241	1.4	103	4.6	371	38.6 [33.6; 43.7]	98.5 [98.1; 98.8]	0.46 [0.41; 0.51]	0.88 [0.87; 0.89]	6121

Alcohol-related disorder (F10)	Total	0.4	24	0.6	38	1.0	67	29.6 [17.0; 42.3]	99.3 [99.1; 99.6]	0.33 [0.20; 0.47]	0.97 [0.96; 0.97]	6432
	Men	0.7	15	0.7	21	1.4	47	31.5 [15.7; 47.3]	99.3 [98.9; 99.6]	0.37 [0.20; 0.54]	0.96 [0.95; 0.97]	2807
	Women	0.2	9	0.6	17	0.6	20	25.3 [8.4; 42.1]	99.4 [99.0; 99.8]	0.25 [0.10; 0.41]	0.98 [0.97; 0.99]	3625
	45–64 Years	0.5	13	0.9	17	1.9	38	21.8 [10.3; 33.4]	99.1 [98.6; 99.6]	0.26 [0.13; 0.39]	0.94 [0.93; 0.96]	2364
	65–79 Years	0.4	5	0.6	11	1.7	27	17.4 [2.4; 32.4]	99.4 [99.1; 99.8]	0.23 [0.05; 0.41]	0.96 [0.94; 0.97]	1683
	Total	8.3	482	1.4	66	28.8	1863	22.3 [20.2; 24.3]	97.8 [97.1; 98.4]	0.24 [0.21; 0.26]	0.40 [0.37; 0.43]	6188
	Men	6.5	164	1.7	34	26.2	781	19.9 [16.7; 23.2]	97.5 [96.6; 98.5]	0.22 [0.18; 0.26]	0.44 [0.40; 0.49]	2722
	Women	9.9	318	1.1	32	31.3	1082	24.1 [21.5; 26.7]	98.1 [97.2; 98.9]	0.25 [0.22; 0.28]	0.35 [0.31; 0.39]	3466
	18–29 Years	9.9	57	2.7	11	18.7	107	34.7 [25.5; 43.9]	96.3 [93.9; 98.7]	0.37 [0.27; 0.48]	0.57 [0.49; 0.65]	616
	30–44 Years	8.0	79	1.6	15	23.7	211	25.2 [19.7; 30.7]	97.6 [96.4; 98.9]	0.28 [0.22; 0.35]	0.49 [0.43; 0.56]	981
Mental disorder (F00–F99)	45–64 Years	10.6	246	1.3	24	30.0	678	26.1 [23.1; 29.1]	97.9 [97.0; 98.8]	0.27 [0.24; 0.30]	0.38 [0.33; 0.42]	2275
	65–79 Years	4.7	74	0.7	10	36.5	593	11.5 [8.7; 14.3]	98.8 [98.0; 99.6]	0.12 [0.08; 0.15]	0.26 [0.21; 0.31]	1620
	≥ 80 Years	3.5	26	0.4	6	40.2	274	8.1 [4.6; 11.6]	99.2 [98.5; 99.9]	0.08 [0.04; 0.12]	0.19 [0.11; 0.27]	696
	Total	1.4	89	3.1	185	8.0	537	14.4 [11.0; 17.9]	96.6 [96.1; 97.2]	0.14 [0.10; 0.19]	0.78 [0.76; 0.80]	6047
	Men	1.0	29	2.6	70	6.0	195	14.7 [8.7; 20.6]	97.2 [96.5; 98.0]	0.15 [0.08; 0.23]	0.83 [0.80; 0.85]	2671
	Women	1.7	60	3.6	115	10.0	342	14.3 [10.4; 18.3]	96.0 [95.1; 96.8]	0.14 [0.09; 0.18]	0.73 [0.70; 0.75]	3376
	18–29 Years	0.9	7	1.8	12	6.5	37	12.6 [1.5; 23.7]	98.1 [96.9; 99.3]	0.15 [0.01; 0.29]	0.84 [0.78; 0.89]	611
	30–44 Years	1.3	15	3.1	34	6.4	65	16.5 [8.0; 25.0]	96.6 [95.4; 97.9]	0.16 [0.06; 0.26]	0.81 [0.77; 0.85]	972
	45–64 Years	2.0	49	4.2	86	7.6	178	20.5 [14.3; 26.7]	95.4 [94.3; 96.5]	0.19 [0.12; 0.26]	0.76 [0.74; 0.79]	2192
	65–79 Years	0.9	13	2.6	40	11.3	187	7.6 [3.0; 12.3]	97.1 [96.1; 98.1]	0.07 [0.00; 0.14]	0.72 [0.69; 0.76]	1595
Somatoform disorder (F45)	≥ 80 Years	0.9	5	2.1	13	10.4	70	7.7 [0.5; 14.9]	97.6 [96.0; 99.3]	0.08 [0.03; 0.19]	0.75 [0.69; 0.81]	677

* sorted in order of descending sensitivity; the data relate to the 12 months preceding the survey date or to the 4th quarter of 2020–3rd quarter of 2021

Only the results for age groups with at least 10 cases in the survey and/or in the routine data are shown.

The proportion of individuals identified in at least one of the two data sources is calculated as the sum of the three proportions listed.

Sensitivity corresponds to the percentage of people with a documented diagnosis in the routine data who also state the respective diagnosis in the survey, in relation to all people who have their diagnosis documented in the routine data. Specificity is the percentage of people without a documented diagnosis in the routine data who also do not state a respective diagnosis in the survey, in relation to all people who do not have a corresponding diagnosis documented in the routine data.

κ, Cohen's kappa measure of agreement; CI, confidence interval; PABAK, prevalence- and bias-adjusted kappa; PTSD, post-traumatic stress disorder

eTable 3

Proportion of medical diagnoses as stated by patients in the survey and/or contained in routine data with regard to lifetime and agreement between the two data sources using Cohen's κ, PABAK, sensitivity, and specificity by sex and age group*

Disease/disorder (ICD-10 codes)	Sex Age group	Proportion (%; weighted) and number (n) with a diagnosis of the disease/disorder							Sensitivity % [95% CI]	Specificity % [95% CI]	κ [95% CI]	PABAK [95% CI]	n
		In both data sources		Only in the survey		Only in the routine data							
		%	n	%	n	%	n						
Myocardial infarction (I21, I22)	Total	2.4	165	1.5	103	0.4	30	85.3 [79.7; 90.8]	98.5 [98.2; 98.8]	0.71 [0.66; 0.77]	0.96 [0.95; 0.97]	5710	
	Men	3.8	129	2.2	81	0.5	20	88.6 [83.1; 94.2]	97.7 [97.2; 98.3]	0.73 [0.67; 0.78]	0.95 [0.94; 0.96]	2481	
	Women	1.2	36	0.8	22	0.4	10	77.2 [62.9; 91.4]	99.2 [98.8; 99.6]	0.67 [0.54; 0.80]	0.98 [0.97; 0.99]	3229	
	45–64 Years	2.2	42	0.8	17	0.4	6	84.1 [72.2; 95.9]	99.2 [98.7; 99.6]	0.77 [0.67; 0.87]	0.98 [0.96; 0.99]	2102	
	65–79 Years	5.6	85	3.2	48	0.6	13	90.4 [84.9; 95.8]	96.6 [95.4; 97.7]	0.73 [0.65; 0.81]	0.92 [0.90; 0.95]	1633	
	≥ 80 Years	5.6	38	5.4	38	1.6	11	77.6 [64.7; 90.5]	94.2 [92.1; 96.2]	0.58 [0.46; 0.69]	0.86 [0.82; 0.90]	705	
	Total	37.0	2429	4.3	216	8.9	588	80.6 [79.0; 82.3]	92.0 [90.8; 93.2]	0.73 [0.71; 0.75]	0.74 [0.72; 0.76]	5701	
Hypertension (I10–I15)	Men	38.0	1206	5.1	105	8.9	286	81.0 [78.5; 83.4]	90.4 [88.3; 92.5]	0.72 [0.69; 0.75]	0.72 [0.69; 0.75]	2492	
	Women	36.0	1223	3.6	111	8.8	302	80.3 [78.1; 82.6]	93.4 [92.0; 94.8]	0.75 [0.72; 0.77]	0.75 [0.72; 0.78]	3209	
	18–29 Years	3.7	13	5.0	18	2.4	10	61.2 [37.4; 84.9]	94.7 [92.0; 97.4]	0.46 [0.27; 0.66]	0.85 [0.79; 0.91]	455	
	30–44 Years	15.3	104	5.6	46	4.5	36	77.4 [69.4; 85.4]	93.0 [90.8; 95.2]	0.69 [0.61; 0.77]	0.80 [0.75; 0.85]	799	
	45–64 Years	36.3	756	5.9	116	8.6	182	80.8 [78.0; 83.7]	89.3 [87.3; 91.4]	0.71 [0.67; 0.74]	0.71 [0.67; 0.75]	2095	
	65–79 Years	64.9	1042	1.7	31	14.0	236	82.2 [80.0; 84.4]	92.1 [89.3; 95.0]	0.61 [0.57; 0.66]	0.69 [0.65; 0.72]	1639	
	≥ 80 Years	72.8	514	0.5	5	17.4	124	80.7 [77.5; 84.0]	95.2 [90.7; 99.7]	0.43 [0.35; 0.51]	0.64 [0.58; 0.70]	713	
Diabetes mellitus (E10–E14)	Total	10.3	688	0.8	52	4.8	341	68.0 [64.8; 71.3]	99.0 [98.7; 99.3]	0.75 [0.73; 0.78]	0.89 [0.87; 0.90]	5769	
	Men	13.0	420	0.2	7	5.0	179	72.1 [68.1; 76.2]	99.7 [99.5; 100]	0.80 [0.77; 0.83]	0.90 [0.88; 0.91]	2497	
	Women	7.8	268	1.4	45	4.7	162	62.6 [57.0; 68.2]	98.5 [98.0; 98.9]	0.69 [0.64; 0.74]	0.88 [0.86; 0.90]	3256	
	30–44 Years	4.8	34	1.7	18	1.6	14	75.2 [62.4; 88.0]	98.2 [97.4; 99.0]	0.73 [0.63; 0.83]	0.93 [0.91; 0.96]	813	
	45–64 Years	8.4	165	0.9	24	3.8	72	68.8 [61.8; 75.8]	99.0 [98; 6; 99.5]	0.76 [0.70; 0.81]	0.91 [0.89; 0.93]	2121	
	65–79 Years	20.1	317	0.4	6	8.9	148	69.4 [64.7; 74.0]	99.4 [98.8; 100]	0.75 [0.71; 0.79]	0.81 [0.78; 0.84]	1642	
	≥ 80 Years	21.5	169	0.4	2	13.3	105	61.8 [55.2; 68.4]	99.4 [98.6; 100]	0.67 [0.61; 0.73]	0.73 [0.67; 0.78]	730	
PTSD (F43.1)	Total	1.3	63	4.2	256	1.0	51	57.3 [46.9; 67.7]	95.7 [95.1; 96.3]	0.32 [0.25; 0.39]	0.90 [0.88; 0.91]	5479	
	Men	0.3	8	3.3	86	0.4	10	44.9 [13.7; 76.1]	96.7 [95.8; 97.5]	0.13 [0.00; 0.26]	0.93 [0.91; 0.94]	2397	
	Women	2.3	55	5.0	170	1.6	41	59.3 [48.2; 70.4]	94.8 [93.9; 95.6]	0.38 [0.30; 0.46]	0.87 [0.85; 0.89]	3082	
Heart failure (I50)	Total	5.1	356	4.6	303	4.2	272	55.1 [50.5; 59.7]	95.0 [94.3; 95.6]	0.49 [0.45; 0.53]	0.83 [0.81; 0.84]	5513	
	Men	6.0	226	4.9	163	4.1	140	59.6 [53.6; 65.6]	94.6 [93.7; 95.5]	0.52 [0.47; 0.58]	0.82 [0.80; 0.84]	2420	
	Women	4.3	130	4.3	140	4.2	132	50.1 [42.7; 57.5]	95.3 [94.4; 96.2]	0.45 [0.39; 0.52]	0.83 [0.81; 0.85]	3093	
	45–64 Years	3.1	60	3.4	65	2.3	50	57.7 [47.4; 68.0]	96.4 [95.5; 97.3]	0.49 [0.40; 0.59]	0.89 [0.86; 0.91]	2052	

Coronary heart disease (I20–I25)	65–79 Years	9.6	153	9.4	145	8.3	123	53.6 [46.9; 60.3]	88.5 [86.6; 90.4]	0.41 [0.35; 0.47]	0.65 [0.60; 0.69]	1553
	≥ 80 Years	21.9	141	9.3	71	16.0	89	57.8 [49.9; 65.7]	85.0 [81.2; 88.8]	0.44 [0.36; 0.53]	0.49 [0.42; 0.57]	662
	Total	7.5	530	2.2	130	6.3	437	54.3 [50.6; 58.1]	97.4 [96.9; 98.0]	0.59 [0.56; 0.63]	0.83 [0.81; 0.85]	5544
	Men	10.6	381	2.0	60	6.5	217	62.1 [57.6; 66.5]	97.7 [96.9; 98.5]	0.67 [0.63; 0.71]	0.83 [0.81; 0.85]	2412
	Women	4.7	149	2.5	70	6.2	220	43.1 [37.0; 49.1]	97.3 [96.5; 98.0]	0.47 [0.41; 0.54]	0.83 [0.81; 0.85]	3132
	45–64 Years	4.8	88	1.5	29	5.1	105	48.4 [40.0; 56.7]	98.3 [97.7; 99.0]	0.56 [0.48; 0.64]	0.87 [0.84; 0.89]	2049
	65–79 Years	17.0	265	3.4	50	12.1	202	58.4 [53.3; 63.5]	95.2 [93.7; 96.7]	0.59 [0.54; 0.64]	0.69 [0.65; 0.73]	1576
	≥ 80 Years	25.0	174	7.3	44	19.7	119	56.0 [49.0; 63.0]	86.7 [82.4; 91.0]	0.44 [0.35; 0.52]	0.46 [0.37; 0.54]	659
	Total	10.7	573	2.8	149	12.1	661	46.9 [43.7; 50.1]	96.3 [95.6; 97.0]	0.50 [0.47; 0.54]	0.70 [0.68; 0.73]	5324
	Men	9.4	204	2.7	59	12.4	314	43.2 [37.9; 48.6]	96.5 [95.4; 97.7]	0.47 [0.41; 0.53]	0.70 [0.66; 0.74]	2272
	Women	11.8	369	3.0	90	11.8	347	49.9 [45.6; 54.3]	96.1 [95.2; 97.0]	0.53 [0.49; 0.57]	0.70 [0.67; 0.73]	3052
	18–29 Years	6.6	19	2.6	10	7.7	27	45.8 [29.2; 62.5]	97.0 [95.0; 98.9]	0.50 [0.34; 0.67]	0.79 [0.72; 0.87]	455
	30–44 Years	13.0	98	2.6	19	8.6	62	60.3 [51.5; 69.2]	96.7 [94.9; 98.5]	0.63 [0.55; 0.72]	0.78 [0.72; 0.83]	785
	45–64 Years	11.2	243	3.4	65	11.8	225	48.6 [43.5; 53.8]	95.6 [94.5; 96.8]	0.51 [0.46; 0.56]	0.70 [0.66; 0.73]	2026
	65–79 Years	12.0	170	2.5	39	17.9	259	40.1 [35.0; 45.2]	96.4 [95.1; 97.6]	0.43 [0.37; 0.48]	0.59 [0.55; 0.63]	1468
Alcohol-related disorder (F10)	≥ 80 Years	6.3	43	2.4	16	16.5	88	27.6 [19.1; 36.1]	96.9 [95.2; 98.5]	0.31 [0.21; 0.42]	0.62 [0.55; 0.70]	590
	Total	1.7	86	1.2	64	2.0	94	46.6 [38.3; 55.0]	98.8 [98.4; 99.1]	0.51 [0.43; 0.58]	0.94 [0.93; 0.95]	5763
	Men	2.7	58	1.6	39	2.8	62	48.7 [37.9; 59.5]	98.4 [97.8; 98.9]	0.53 [0.43; 0.63]	0.91 [0.89; 0.93]	2507
	Women	0.9	28	0.9	25	1.2	32	41.6 [27.1; 56.1]	99.1 [98.6; 99.6]	0.44 [0.31; 0.58]	0.96 [0.95; 0.97]	3256
	45–64 Years	2.3	40	1.7	28	2.3	43	50.6 [39.4; 61.7]	98.2 [97.5; 99.0]	0.52 [0.41; 0.62]	0.92 [0.90; 0.94]	2120
	65–79 Years	2.2	34	1.0	19	1.4	23	60.5 [45.1; 76.0]	99.0 [98.5; 99.5]	0.63 [0.50; 0.76]	0.95 [0.93; 0.97]	1645
Depression (F32, F33.0–F33.3, F33.8–F33.9, F34.1; excluding F30, F31)	Total	12.4	799	5.5	314	14.3	991	46.4 [43.6; 49.2]	92.4 [91.4; 93.5]	0.43 [0.40; 0.47]	0.60 [0.58; 0.63]	6243
	Men	9.2	262	5.2	131	9.8	330	48.2 [43.4; 53.1]	93.5 [92.2; 94.9]	0.46 [0.41; 0.51]	0.70 [0.67; 0.73]	2729
	Women	15.5	537	5.8	183	18.6	661	45.4 [42.0; 48.7]	91.2 [89.7; 92.7]	0.40 [0.36; 0.44]	0.51 [0.48; 0.55]	3514
	18–29 Years	8.8	50	5.6	36	4.8	28	64.7 [51.3; 78.0]	93.6 [91.1; 96.1]	0.57 [0.45; 0.69]	0.79 [0.73; 0.85]	630
	30–44 Years	11.2	113	7.3	64	11.1	121	50.4 [43.6; 57.1]	90.6 [87.9; 93.3]	0.44 [0.36; 0.51]	0.63 [0.58; 0.69]	995
	45–64 Years	15.7	379	6.1	137	15.1	342	51.0 [47.2; 54.9]	91.2 [89.6; 92.7]	0.46 [0.42; 0.50]	0.58 [0.54; 0.61]	2299
Anxiety disorder (F40, F41)	65–79 Years	12.9	203	3.4	54	20.1	333	39.1 [34.5; 43.7]	94.9 [93.2; 96.5]	0.39 [0.33; 0.44]	0.53 [0.48; 0.58]	1632
	≥ 80 Years	7.3	54	2.7	23	26.6	167	21.6 [16.1; 27.2]	96.0 [94.3; 97.6]	0.21 [0.14; 0.28]	0.42 [0.34; 0.49]	687
	Total	7.2	364	3.8	222	11.3	651	38.7 [35.0; 42.4]	95.4 [94.6; 96.1]	0.41 [0.37; 0.44]	0.70 [0.68; 0.72]	5522
	Men	5.3	110	3.5	96	7.1	184	42.7 [35.7; 49.8]	96.0 [95.0; 97.0]	0.44 [0.37; 0.51]	0.79 [0.76; 0.82]	2416
	Women	8.8	254	4.0	126	15.2	467	36.8 [32.7; 40.9]	94.7 [93.7; 95.7]	0.37 [0.33; 0.42]	0.62 [0.58; 0.65]	3106
	18–29 Years	10.0	36	2.0	7	9.1	38	52.3 [37.8; 66.8]	97.5 [95.4; 99.6]	0.58 [0.44; 0.72]	0.78 [0.70; 0.85]	443
	30–44 Years	6.9	59	2.4	18	12.5	102	35.4 [27.4; 43.4]	97.0 [95.4; 98.6]	0.40 [0.31; 0.49]	0.70 [0.64; 0.76]	785

Mental disorder (F00–F99)	45–64 Years	7.8	159	5.0	99	10.2	216	43.4 [38.2; 48.7]	93.9 [92.6; 95.3]	0.42 [0.37; 0.47]	0.70 [0.66; 0.73]	2026
	65–79 Years	5.6	83	4.4	71	13.4	219	29.3 [23.5; 35.1]	94.5 [93.2; 95.8]	0.29 [0.22; 0.36]	0.64 [0.60; 0.68]	1590
	≥ 80 Years	4.7	27	3.3	27	11.6	76	28.6 [17.9; 39.2]	96.1 [94.4; 97.8]	0.31 [0.19; 0.43]	0.70 [0.64; 0.76]	678
	Total	14.6	779	0.5	30	56.3	3186	20.6 [19.1; 22.1]	98.2 [97.5; 99.0]	0.12 [0.11; 0.13]	-0.14 [-0.17; -0.11]	5584
	Men	12.0	275	0.7	17	50.2	1262	19.3 [16.9; 21.8]	98.3 [97.3; 99.3]	0.14 [0.12; 0.16]	-0.02 [-0.07; 0.03]	2455
	Women	17.0	504	0.4	13	62.1	1924	21.5 [19.5; 23.5]	98.1 [96.9; 99.3]	0.09 [0.08; 0.11]	-0.25 [-0.29; -0.21]	3129
	18–29 Years	18.0	66	1.0	5	45.0	206	28.6 [21.5; 35.8]	97.3 [94.8; 99.8]	0.21 [0.14; 0.28]	0.08 [-0.03; 0.19]	446
	30–44 Years	14.6	119	0.4	4	55.9	437	20.7 [16.8; 24.7]	98.6 [97.1; 100]	0.13 [0.09; 0.16]	-0.13 [-0.21; -0.05]	783
	45–64 Years	17.2	363	0.7	13	55.9	1150	23.5 [21.3; 25.8]	97.5 [96.2; 98.9]	0.13 [0.11; 0.15]	-0.13 [-0.17; -0.09]	2061
	65–79 Years	11.2	182	0.2	5	60.6	955	15.6 [13.3; 18]	99.2 [98.4; 99.9]	0.09 [0.07; 0.11]	-0.22 [-0.27; -0.16]	1599
	≥ 80 Years	7.2	49	0.2	3	65.6	438	9.9 [6.6; 13.1]	99.3 [98.4; 100]	0.05 [0.03; 0.07]	-0.32 [-0.40; -0.24]	695
	Total	4.6	245	2.7	151	28.2	1605	13.9 [12.1; 15.8]	96.0 [95.3; 96.8]	0.12 [0.10; 0.15]	0.38 [0.35; 0.41]	5436
	Men	2.9	67	2.4	62	19.6	514	13.0 [9.5; 16.5]	96.9 [95.9; 97.8]	0.14 [0.09; 0.18]	0.56 [0.52; 0.60]	2387
	Women	6.1	178	2.9	89	36.2	1091	14.4 [12.2; 16.6]	95.0 [93.7; 96.3]	0.11 [0.08; 0.13]	0.22 [0.18; 0.26]	3049
Somatoform disorder (F45)	18–29 Years	3.4	15	2.4	10	21.5	95	13.7 [6.1; 21.3]	96.7 [94.4; 99.1]	0.14 [0.03; 0.25]	0.52 [0.42; 0.62]	437
	30–44 Years	4.5	39	2.2	18	28.5	243	13.5 [9.0; 18.1]	96.7 [95.0; 98.5]	0.13 [0.07; 0.19]	0.39 [0.31; 0.46]	778
	45–64 Years	6.7	133	3.3	67	27.6	562	19.4 [16.1; 22.7]	95.0 [93.7; 96.3]	0.17 [0.13; 0.21]	0.38 [0.34; 0.42]	1976
	65–79 Years	3.0	48	2.6	42	31.3	494	8.8 [6.0; 11.7]	96.0 [94.6; 97.5]	0.06 [0.02; 0.10]	0.32 [0.27; 0.37]	1565
	≥ 80 Years	2.0	10	1.9	14	32.4	211	5.9 [1.9; 9.8]	97.1 [95.0; 99.3]	0.04 [-0.02; 0.10]	0.31 [0.23; 0.40]	680

* sorted in order of descending sensitivity; the data relate to lifetime or to the 4th quarter of 2011–3rd quarter of 2021
Only the results for age groups with at least 10 cases in the survey and/or in the routine data are shown.
The proportion of individuals identified in at least one of the two data sources is calculated as the sum of the three proportions listed.
Sensitivity corresponds to the percentage of people with a documented diagnosis in the routine data who also state the respective diagnosis in the survey, in relation to all people who have their diagnosis documented in the routine data. Specificity is the percentage of people without a documented diagnosis in the routine data who also do not state a respective diagnosis in the survey, in relation to all people who do not have a corresponding diagnosis documented in the routine data.
κ, Cohen's kappa measure of agreement; CI, confidence interval; PABAK, prevalence- and bias-adjusted kappa; PTSD, post-traumatic stress disorder