MIND THE GAP: A RETROSPECTIVE STUDY OF DISCREPANCIES IN SELF-REPORTED AND ADMINISTRATIVE DATABASE-IDENTIFIED MENTAL HEALTH ISSUES IN SLOVENIA

RETROSPEKTIVNA ŠTUDIJA O NESKLADJIH MED SAMOPOROČANIMI IN V ADMINISTRATIVNIH ZBIRKAH PODATKOV UGOTOVLJENIMI TEŽAVAMI V DUŠEVNEM ZDRAVJU V SLOVENIJI

Matej VINKO [1,2*] , Andreja KUKEC [2,3] , Lijana ZALETEL-KRAGELJ [2,3]

[1] Faculty of Medicine, University of Ljubljana, Vrazov trg 2, 1000 Ljubljana, Slovenia
[2] National Institute of Public Heath, Trubarjeva cesta 2, 1000 Ljubljana, Slovenia
[3] Chair of Public Health, Faculty of Medicine, University of Ljubljana, Zaloška cesta 4, 1000 Ljubljana, Slovenia

ABSTRACT

Keywords

Mental health

Epidemiology

Data linkage

Self-report

Routinely collected health data

Background
This study assessed discrepancies between self-reported and administrative data sources in identifying mental health issues in Slovenia, and investigated associated socio-demographic factors.

Methods
Data were linked from the 2019 Slovenian European Health Interview Survey (EHIS; n=9,900) and national health administrative databases capturing inpatient hospitalisations, outpatient prescription drugs and mental health-related sick leave. Mental health issues were identified in EHIS by self-report and in administrative databases using diagnostic codes and medication claims. Socio-demographic factors were obtained from EHIS. Discrepancies were assessed and multinomial logistic regression was used to analyse the association between these factors and the source of case identification.

Results
Of the 9,900 EHIS respondents, 1,336 (13.5%) self-reported mental health issues, while 1,675 (16.9%) were identified in administrative databases. Only 613 individuals (4.6% of the total sample) were identified in both sources. Older age was associated with being identified in both data sources and administrative data only compared to not being identified. Females and unemployed persons were more likely than males and employed persons to be identified as having mental health issues, regardless of the data source. Compared to those with primary education or lower, individuals with higher education were less likely to be identified in administrative data only or in both data sources.

Conclusions
discrepancies exist between self-reported and administrative data sources in identifying mental health issues. Discrepancies are associated with socio-demographic factors and may lead to different interpretations of population mental health. This study underscores the importance of cautiously interpreting self-reported and administrative health data in public health.

IZVLEČEK

Ključne besede

duševno zdravje

epidemiologija

povezovanje podatkov

samoporočanje

rutinsko zbrani zdravstveni podatki

Uvod
Z raziskavo smo ocenili neskladja v prepoznavi posameznikov s težavami v duševnem zdravju pri uporabi samoporočanih in administrativnih podatkov v Sloveniji ter raziskali povezane socialno-demografske dejavnike.

Metode
Podatki so bili pridobljeni iz Nacionalne raziskave o zdravju in zdravstvenem varstvu (EHIS) iz leta 2019 (n = 9.900) in treh zdravstvenih administrativnih baz podatkov, ki zajemajo bolnišnične obravnave, izdajo ambulantnih zdravil na recept ter koriščenje bolniškega staleža. Težave z duševnim zdravjem so bile v raziskavi EHIS opredeljene s samoporočanjem, v administrativnih bazah podatkov pa z uporabo diagnostičnih kod in izdanih receptov za zdravila. Socialno-demografski dejavniki so bili pridobljeni iz raziskave EHIS. Neskladja v prepoznavi posameznikov s težavami v duševnem zdravju in povezanost socialno-demografskih dejavnikov so bila analizirana z uporabo polinomne regresijske analize.

Rezultati
Od 9.900 anketirancev v raziskavi EHIS jih je 1.336 (13,5 %) samoporočalo o težavah z duševnim zdravjem, medtem ko jih je bilo 1.675 (16,9 %) prepoznanih v administrativnih bazah podatkov. Le 613 posameznikov (4,6 % celotnega vzorca) je bilo prepoznanih hkrati v obeh virih. Starejša starost je bila povezana s prepoznavo v obeh virih podatkov in samo v administrativnih podatkih v primerjavi z neprepoznavo v katerem koli viru podatkov. Ženske in brezposelni posamezniki so bili bolj verjetno kot moški in zaposleni prepoznani kot osebe s težavami v duševnem zdravju, ne glede na vir podatkov. V primerjavi s tistimi z osnovnošolsko ali nižjo izobrazbo so bili posamezniki z višjo izobrazbo manj verjetno prepoznani samo v administrativnih podatkih ali v obeh virih podatkov.

Zaključki
Pri prepoznavi težav z duševnim zdravjem obstajajo znatna neskladja med samoporočanjem in administrativnimi viri podatkov. Neskladja so povezana z različnimi socialno-demografskimi dejavniki in lahko vodijo do različnih interpretacij duševnega zdravja prebivalstva. Ta študija poudarja pomen previdne interpretacije samoporočanih in administrativnih zdravstvenih podatkov v javnem duševnem zdravju.

*Correspondence: matej.vinko@nijz.si

1 INTRODUCTION

Mental health is an area of growing public health importance. Despite successful global efforts in reducing the burden of other non-communicable diseases, the burden of mental disorders and self-harm behaviour remains stable (1). The persistent burden of mental disorders and self-harm underscores the importance of valid and reliable population prevalence data for public health decision-making.

Population surveys and health administrative databases are common sources for mental disorder prevalence data (2). Existing research indicates significant differences in the populations identified as having mental disorders by surveys and those identified through health administrative databases (3-6). For example, studies show substantial discordance between data sources: Edwards et al. (3) found only 19.4% agreement between survey and administrative diagnoses of mood and anxiety disorders in Ontario. Davis et al. (5) reported limited agreement (Cohen's kappa 0.24-0.46) even between different self-report methods in the UK Biobank. While O'Donnell et al.'s (6) trend analysis in Canada noted a convergence over time, administrative prevalence remained consistently higher than self-report. Even when self-report is used in online registries, Sordo Vieira et al. (4) found that while it can indicate general psychopathology, in-depth interviews are still needed for specific disorder identification.

However, this evidence is limited to a small number of studies, each applying a different methodology. These discrepancies may result due to variations in population coverage, types of health services included, classifications used to register data and classification algorithms or the way mental health disorder is defined (7). In this study, we use the term 'mental health issues' to encompass a broad range of mental health experiences, including diagnosable mental disorders (as defined by clinical criteria) as well as subthreshold manifestations and transient distress that may not meet diagnostic criteria but still impact wellbeing. This distinction is important because it allows us to examine the full spectrum of mental health in the population, rather than focusing solely on diagnosed conditions.

With the aim of assessing the discrepancies between self-reported and administrative data sources in identifying mental health issues in Slovenia, the objectives of the present study were: to assess the discrepancies in self-reported and database-identified groups of cases with mental health issues in Slovenia, linking data from a large population-based survey and three health administrative databases; and to gain insight into which socio-demographic factors contribute to the identification of a case in either of the data sources or in both of them.

2 METHODS

The study was designed as a retrospective study using data from the 2019 Slovenian European Health Interview Survey (EHIS) (8), conducted by the National Institute of Public Health (NIPH), and from three national health administrative databases, also maintained by the NIPH. This approach allowed us to compare individual-level data on mental health service utilisation with self-reported mental health status, providing a comprehensive assessment of the agreement between these data sources.

2.1 Data sources

2.1.1 Definition of the baseline case group

The study baseline case group was derived from respondents of the EHIS (8). The survey targeted Slovenian residents aged 15 and over residing in private households. The Statistical Office of the Republic of Slovenia prepared the probability sample, utilising the framework of census districts and the Central Population Register. A two-stage sampling method was employed, resulting in a stratified two-stage sample. Stratification was explicit by size and type of settlement, and implicit by statistical regions. Data collection was conducted through online (CAWI) and personal (CAPI) interviews. Data were analysed in their unweighted form.

2.1.2 Case groups defined on the basis of administrative databases

The study sample was identified from the health administrative databases using unique personal registration numbers, consistent across databases and included in the EHIS dataset. The databases used include the National Hospital Health Care Statistics Database (NHHCS) (9), Outpatient Prescription Drugs Database (OPD) (10), and Absence from Work Database (AFW) (11) for 2018 and 2019. The administrative databases, encompassing records of public health care utilisation for all residents covered by Slovenian compulsory health insurance, have been utilised in previous mental health research (12, 13). Utilisation is recorded as either an episode (of inpatient treatment or sick leave) or a drug claim (obtained through prescription). Identification of the sample from each individual database occurred in a secure room at NIPH. Individual administrative databases data were matched with EHIS.

2.2 Data sources linkage

After identifying the EHIS case group individuals in the health administrative databases, the personal registration numbers were replaced with sequence numbers from EHIS for data privacy. These sequence numbers were randomly allocated to each individual survey participant, ensuring anonymity while maintaining the ability to link data across databases.

2.3 Outcome measures and observed outcomes

2.3.1 The EHIS outcome measure

Data on mental health issues were obtained from the EHIS self-reported health variable ("Having depression in the past 12 months" with two possible answers: yes or no) and two additional variables ("Having anxiety in the past 12 months" and "Having any other mental health issue in the past 12 months"). A positive answer to any of the three questions was recorded as a case of prevalent mental health issue.

2.3.2 The administrative databases outcome measures

Cases with prevalent mental health issues were identified from the health administrative databases using the following criteria:

- NHHCS: any hospitalisation with the main or additional diagnosis from the 5th Chapter in the ICD-10 within a period of 12 months before responding to EHIS,
- OPD: any claim with a psycholeptic or psychoanaleptic drug prescribed (all drugs under ATC classification N05A, N05B, N05C, N06A, N06B subgroups were included) within a period of 12 months before responding to EHIS,
- AFW: any sick leave with the main diagnosis from the 5th Chapter in the ICD-10 within a period of 12 months before responding to EHIS.

2.3.3 Observed outcomes

Outcome measures from administrative databases were added to the basic EHIS database as new variables. In the creation of the outcome variable, EHIS was taken into account as one data source, and all three administrative databases together as another. Based on this, a new outcome variable "source of case identification" with the following values was created: 1=case identified in both data sources, 2=case identified in one of the administrative data sources only, 3=case identified in the survey data source only, 0=absence of identification in any of the data sources. A "case identified in EHIS only" (0=no, 1=yes) represented a separate outcome variable.

2.3.4 Socio-demographic characteristics

Data on age, gender, educational status and employment status were obtained from EHIS, representing explanatory factors for the observed outcomes.

2.4 Statistical analysis

Statistical analysis unfolded in three distinct stages. The initial stage involved creating a Venn diagram, using the InteractiVenn tool, to visualise the overlaps of groups where cases were identified (14). The second stage entailed a cross tabulation of the identification of mental health issues from health administrative databases and EHIS. We employed the chi-square test to test the association between identification from the two data sources. To gauge the strength of agreement between the data sources, we utilised the Kappa statistic. In the third stage, we examined the socio-demographic characteristics of individuals identified with mental health issues in one or both sources, or in neither of them. First, we conducted chi-square tests to examine associations among categorical variables across different groups, and ANOVA tests to analyse variations in continuous variables. Finally, we conducted a multinomial logistic regression analysis to investigate the association between explanatory factors and the outcome variable. The final multinomial logistic regression model included all the socio-demographic variables described in section 2.3.4. All variables were entered simultaneously into the model. No stepwise or other automated variable selection procedures were performed. No extrapolation was used. The "absence of identification in any of the data sources" category represented the reference category. We checked for multicollinearity using the Variance Inflation Factor (VIF), ensuring the independent variables in the regression model were not excessively correlated. Categorical independent variables were included in the models as dummy variables. A simple method was applied. Data analysis was carried out using the IBM Statistical Package for the Social Sciences (SPSS) version 25 (15).

3 RESULTS

3.1 Study group description

The EHIS consisted of 9,900 records (response rate 67%), which represented the study group (Table 1).

Table 1. Socio-demographic characteristics of the study group in the study of discrepancies in self-reported and administrative database-identified mental health issues in Slovenia.

Variable	Mean±SD/ N (%)
Age	50.6±19.1
Gender	
Male	4479 (45.2%)
Female	5421 (54.8%)
Education status	
Primary school or lower	1737 (17.5%)
Secondary school	5257 (53.1%)
Tertiary or higher	2906 (29.4%)
Employment status	
Employed, self-employed	4813 (48.9%)
Unemployed	667 (6.8%)
Student	913 (9.3%)
Retired	3332 (33.8%)
Other	127 (1.3%)

3.2 Analysis of overlapping of the case identification sources

A total of 1,336 individuals self-reported mental health issues in EHIS and 1,675 were identified as having mental health issues based on their utilisation of healthcare services. The majority of the latter cases were identified in the OPD (n=1625), while a small percentage was identified in the NHHCS (n=78) and the AFW (n=97). Overlaps between survey and administrative database derived case groups are presented with a Venn diagram (Figure 1).
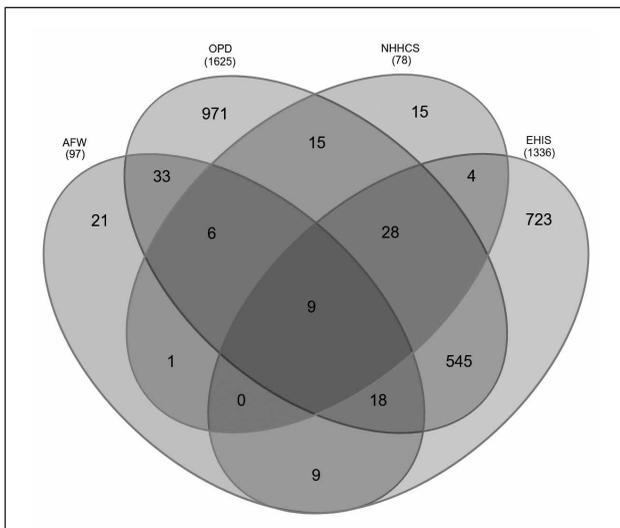

Figure 1. Comparison of number of cases identified as having mental health issues across four data sources (AFW=Absence from Work Database, OPD=Outpatient Prescription Drugs Database, NHHCS=National Hospital Health Care Statistics Database, EHIS=European Health Interview Surveys).

3.3 Derivation of the "source of case identification" observed outcome

Cross tabulation of the groups identified as cases in administrative databases or self-reporting mental health issues in the survey shows significant discrepancies in group representation between the two data sources (Table 2). Socio-demographic characteristics of the groups are presented in Table 3.

Out of 1,336 self-reported cases from EHIS only 613 of them (45.9%) had a record in the administrative databases in the 12 months preceding their participation in EHIS. And out of 1,675 individuals identified as cases in administrative databases only 613 (36.6%) self-reported having mental health issues in that same time period. Overall, the results suggest a modest level of agreement between self-assessment and administrative databases, indicating that the concordance between the two data sources is limited (Kappa=0.302; p<0.001). In total, 24.2% of the total sample was identified experiencing mental health issues from at least one data source in the period of 12 months preceding their participation in EHIS, out of which 25.6% were identified in both data sources.

Table 2. Identification or self-identification of mental health issues in three administrative databases of the National Institute of Public Health in the study of discrepancies in self-reported and administrative database-identified mental health issues in Slovenia.

At least 1 record in any database	Self-reporting of mental health issues (survey)		
	Present	**Absent**	**Total**
Present	613 (45.9%)	1062 (12.4%)	1675 (16.9%)
Absent	723 (54.1%)	7502 (87.6%)	8225 (83.1%)
Total	1336 (100.0%)	8564 (100.0%)	9900 (100.0%)

$p_{chi-square}$ <0.001

Table 3. Socio-demographic characteristics of the case groups identified with mental health issues in three administrative databases of the National Institute of Public Health in the study of discrepancies in self-reported and administrative database-identified mental health issues in Slovenia.

Variable	Survey and administrative n=613	Administrative n=1062	Survey n=723	Neither n=7502	Total n=9.900	p-value
	Mean±SD/ N (%)					
Age	57.8±2.0	63.2±15.8	42.8±19.7	49.0±18.6	50.6±19.1	<0.001
Gender						
Male	181 (29.5%)	367 (34.6%)	257 (35.5%)	3674 (49.0%)	4479 (45.2%)	<0.001
Female	432 (70.5%)	695 (65.4%)	466 (64.5%)	3828 (51.0%)	5421 (54.8%)	
Education status						
Primary school or lower	142 (23.2%)	251 (23.6%)	172 (23.8%)	1172 (15.6%)	1737 (17.5%)	<0.001
Secondary school	327 (53.3%)	593 (55.8%)	323 (44.7%)	4014 (53.5%)	5257 (53.1%)	
Tertiary or higher	144 (23.5%)	218 (20.5%)	228 (31.5%)	2316 (30.9%)	2906 (29.4%)	
Employment status						
Employed, self-employed	197 (32.1%)	301 (28.4%)	330 (46.1%)	3985 (53.4%)	4813 (48.9%)	<0.001
Unemployed	98 (16.0%)	70 (6.6%)	78 (10.9%)	421 (5.6%)	667 (6.8%)	
Student	21 (3.4%)	15 (1.4%)	157 (21.9%)	720 (9.6%)	913 (9.3%)	
Retired	289 (47.1%)	656 (61.8%)	144 (20.1%)	2243 (30.1%)	3332 (33.8%)	
Other	8 (1.3%)	19 (1.8%)	7 (1.0%)	93 (1.2%)	127 (1.3%)	

3.4 Analysis of multinomial regression analysis

The results showed that higher age is significantly associated with being identified as having mental health issues in administrative source only, and administrative and survey data sources compared to not being identified in any of the sources (Table 4). Females were more than twice as likely as males to be identified as having mental health issues in both data sources (OR=2.23), only in administrative data (OR=1.77), and only in survey data (OR=1.69) compared to not being identified in any of the sources. Compared to individuals with primary school education or lower, those with secondary school education were less likely to be identified as having mental health issues only in survey data (OR=0.73) and those tertiary or higher education were less likely to be identified only in administrative data (OR=0.75) compared to not being identified in any of the sources. Unemployed individuals were about four times more likely than employed individuals to be identified in both data sources (OR=4.07), and they were also more likely to be identified only in administrative data (OR=1.82) and only in survey data (OR=2.20). Students were more likely than employed individuals to be identified only in survey data (OR=2.06) compared to not being identified in any of the sources. The Likelihood Ratio Test compared the final model to a model with only the intercept, and the significant result (p<0.001) indicated that the predictors in the model significantly improved the fit. To assess the goodness-of-fit of the multinomial logistic regression model, we employed both Pearson's test, which yielded a statistically significant result (p=0.006), and Deviance test, the result of which was non-significant (p=0.999).

4 DISCUSSION

The analysis revealed pronounced disparities in mental health issue prevalence estimates obtained from self-reported data compared to health administrative databases, and only a minority of individuals identified as experiencing mental health issues were found in both data sources.

This finding highlights a significant measurement challenge in public mental health research. Nearly a third of individuals self-reported mental health issues but were not identified in the administrative data sources, while almost half that were identified in the administrative data did not report mental health issues in the EHIS. The discordance in two case groups derived from administrative and survey data sources was expected, as it was already identified in previous similar research (5-8).

What separates our research from previous studies is that we have not focused on specific mental health disorders (i.e. mood and anxiety disorders), yet we have still found a substantial part of the population using mental health services while not self-reporting any mental issues in the same time window. While administrative data is generally expected to underestimate the prevalence of mental health conditions, our finding of a higher prevalence in administrative data compared to the EHIS survey, consistent with observations

Table 4. Results of multinomial regression analysis of mental health issue identification across different data sources in the study of discrepancies in self-reported and administrative database-identified mental health issues in Slovenia.

Variable	Administrative and survey			Only administrative			Only survey		
	OR	95% CI	p-value	OR	95% CI	p-value	OR	95% CI	p-value
Age	1.02	1.01-1.03	<0.001	1.03	1.03-1.04	<0.001	0.99	0.98-1.00	0.086
Gender									
Male	1.00			1.00			1.00		
Female	2.23	1.86-2.68	<0.001	1.77	1.44-1.99	<0.001	1.77	1.44-1.99	<0.001
Education status			<0.001			<0.001			<0.001
Primary school or lower	1.00			1.00			1.00		
Secondary school	0.86	0.69-1.08	0.198	0.95	0.79-1.14	0.572	0.73	0.59-0.90	0.004
Tertiary or higher	0.75	0.57-0.99	0.039	0.75	0.60-0.93	0.009	0.88	0.69-1.13	0.310
Employment status			<0.001			<0.001			<0.001
Employed, self-employed	1.00			1.00			1.00		
Unemployed	4.07	3.11-5.34	<0.001	1.82	1.37-2.42	<0.001	2.20	1.67-2.89	<0.001
Student	0.86	0.51-1.45	0.569	0.57	0.32-1.01	0.054	2.06	1.51-2.81	<0.001
Retired	1.41	1.05-1.90	0.022	1.48	1.17-1.87	0.001	0.93	0.69-1.27	0.666
Other	0.92	0.43-1.97	0.836	1.28	0.72-2.19	0.368	0.76	0.34-1.67	0.489

from a similar study (6), underscores the complexity of mental health measurement and the need to further investigate the factors contributing to this discrepancy.

The sizeable group of individuals identified in the healthcare databases yet not self-reporting mental health issues in the EHIS presents an intriguing facet of the measurement discrepancy. We found higher age is associated with being identified as having mental health issues in administrative data sources. The majority of cases identified from administrative databases were captured in OPD. A study on antidepressant prescribing trends in Slovenia showed substantial differences in the prevalence of antidepressant recipients between those aged 54 or lower and those older than 55 years, with prevalence reaching 20% in individuals aged 80 and over (16). Even though neither the referenced or present study delves into the interpretation of possible causes of high consumption of psychoactive drugs among the elderly, the results indicate higher healthcare usage with increasing age. In parallel, older adults are less likely to recognise mental health issues as such, (17) and might be treated for mental health symptoms caused by physical illness, therefore not identifying as having mental health issues in the surveys, because the primary source of poor mental health is associated with a physical illness (18). They could also be unaware of the connection between their symptoms and a diagnosable condition, or this information might not have been effectively communicated to them by healthcare providers. Consequently, while their conditions are reflected in administrative data, their self-reports may

not include mental health concerns. Recall bias could also influence this discrepancy (19). Participants in the EHIS might have difficulty remembering experiencing mental health issues or they might underestimate their past negative experience in the function of coping, although research indicates the latter might not be a common phenomenon (20). Lastly, we have to consider stigma as an important factor in using healthcare services and self-reporting mental health issues. Stigma appears to have a greater impact on self-reporting of mental health issues in surveys compared to the actual usage of mental health care services (21-23). Individuals may be more willing to seek professional help than to disclose mental health problems. Individuals with higher education were less likely to be identified only in administrative data, potentially suggesting a greater comfort level with self-disclosure of mental health issues in this population group. While our study design cannot definitively measure the impact of stigma, these findings hint at a potential interplay between stigma and data source discrepancies. Higher prevalence observed in administrative data might appear counterintuitive, however it highlights the potential influence of prescription drug data and the varying pathways individuals take to access mental healthcare. This underscores the importance of using multiple data sources to capture a more complete picture of mental health in the population.

It is also important to note that while the overall prevalence of mental health issues was higher in the administrative databases, a subset of individuals reported these issues in

the EHIS survey but did not appear in the administrative data. For these individuals, stigmatising attitudes towards help-seeking behaviour, such as negative expectations of professional help or fear of being stigmatised for having a mental disorder, may be a significant barrier to seeking help, as documented in previous research (24, 25). One key factor is likely the individual's perception of their need for healthcare services. Individuals who do not perceive a need for professional help, perhaps due to mild symptoms, minimal functional impairment, or other personal beliefs, are less likely to engage with the healthcare system and thus be recorded in administrative databases. Some individuals might seek help outside of the traditional healthcare system, through alternative therapies or support groups, which would not be reflected in administrative data (17). However, these individuals may still acknowledge and report experiencing mental health issues in surveys. Furthermore, limitations inherent in the health data collection methods themselves might play a role. Incomplete data coverage within the administrative databases is a possibility. For example, individuals using outpatient services without claiming psychoactive medications were not captured in the administrative data sources we used in this study. Additionally, under-diagnosis by healthcare providers could contribute to the discrepancy. Stigma surrounding mental health or a lack of awareness among providers could lead to missed diagnoses, resulting in individuals experiencing mental health issues but not being recorded in administrative databases (26, 27). Also, younger participants of the EHIS survey, students specifically, were less likely to be identified only in administrative and more likely to be identified in only survey data sources. This could be explained by a number of factors previously researched within this age group, such as preference towards informal sources of help and favouring counselling compared to psychiatric or psychological services, (17) or even interpreting and reporting milder forms of distress as mental health issues (28).

This study has some limitations. First, the administrative databases used capture only a part of mental health service utilisation within the formal healthcare system, potentially underestimating service use. Additionally, the accuracy of administrative data relies on consistent and reliable coding practices. Second, the EHIS mental health assessment used in this study is less structured than instruments such as the Composite International Diagnostic Interview and is not commonly used to assess mental disorder prevalence. This approach may encompass a broader range of mental distress, potentially including subthreshold experiences, and may not be directly comparable to studies using validated mental disorder screening tools. However, a broader scope is valuable for understanding the full spectrum of mental health within a population, not solely clinical diagnoses.

The inclusion of subthreshold experiences offers insights that have value in contributing to better understanding of help-seeking behaviour as well as identification of at-risk populations (28, 29). A single-item question on experiencing symptoms or receiving a diagnosis of mental disorder during a certain time period in the past is an approach often used in population surveys in the field of mental health (30) as well as in other fields (31). Secondly, the goodness-of-fit of our multinomial logistic regression model, as indicated by the Pearson test, was less than ideal. While the Deviance test suggested adequate fit, the Pearson test indicated suboptimal goodness-of-fit for our multinomial logistic regression, possibly due to large sample size and outcome class imbalance. A secondary analysis with reduced imbalance improved the Pearson test. Although statistical fit has limitations, the model's face validity and meaningful insights remain valuable. Readers should consider fit limitations when interpreting results. Finally, someone may dispute lack of sensitivity analysis. In fact, this analysis was done, but due to the limits of the journal it could not be presented. However, the results can be obtained from the authors. A key strength of this study lies in its unique access to both self-reported data from the large, representative EHIS survey and comprehensive administrative health databases, linked at the individual level. This approach allows for a more comprehensive understanding of the discrepancies and complexities inherent in identifying mental health issues using different data sources. Finally, by focusing on the Slovenian population within its unique healthcare context of universal coverage and a mixed public-private provider model, this study offers insights relevant to similar health systems where diverse factors may influence mental health help-seeking and reporting.

5 CONCLUSION

Integrating diverse data sources and refining measurement instruments are crucial steps towards a more comprehensive and accurate picture of mental health in our communities. Additionally, a deeper understanding of the factors influencing help-seeking behaviour, and how this relates to mental health and health outcomes, is essential for developing effective interventions and support systems. Future research should further explore the factors influencing discrepancies between data sources and develop strategies to bridge the gap between self-reported experiences and objective health records.

To improve agreement between these data sources, collaborative efforts are needed. Improving the accuracy and completeness of administrative data requires implementing standardised coding practices and encouraging the use of validated clinical assessment tools. This necessitates cooperation between public health researchers, healthcare providers and policymakers.

Enhancing the validity of self-reported data relies on using validated questionnaires with clear instructions, while ensuring respondent privacy. Finally, reconciling discrepancies requires implementing data linkage initiatives and analysing the reasons for discrepancies between data sources.

These efforts, along with initiatives such as the European Health Data Space, which opens possibilities for large-scale, multinational longitudinal studies, are crucial for ensuring data accurately reflect the mental health needs of the population.

ACKNOWLEDGMENTS

CONFLICTS OF INTEREST

The authors have no relevant financial or non-financial interests to disclose.

FUNDING

ETHICAL APROVAL

National Statistics Act, Regulation (EC) No 1338/2008, Commission Regulation (EU) 2018/255, Annual Programme of Statistical Surveys for 2019 provide the legal basis for EHIS data collection. These article also encompass ethical considerations.

AVAILABILITY OF DATA AND MATERIALS

Due to the sensitive nature of individual health information and data protection regulations, the raw data cannot be made publicly available. For further information or inquiries regarding the data, please contact the corresponding author.

AI USAGE STATEMENT

During the preparation of this article the authors did not use generative language models.

PREPRINT STATEMENT

The authors confirm that this article has not been posted or published as a preprint prior to submission.

ORCID

Matej Vinko:
https://orcid.org/0000-0003-0396-4510

Andreja Kukec:
https://orcid.org/0000-0002-5267-8409

Lijana Zaletel-Kragelj:
https://orcid.org/0000-0003-1014-7906

REFERENCES

1. GBD 2019 Mental Disorders Collaborators. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990-2019: A systematic analysis for the Global Burden of Disease Study 2019. Lancet Psychiatry. 2022 Feb 1;9(2):137-150. doi: 10.1016/S2215-0366(21)00395-3.

2. Vigo D, Jones W, Dove N, Maidana DE, Tallon C, Small W, et al. Estimating the prevalence of mental and substance use disorders: A systematic approach to triangulating available data to inform health systems planning. Can J Psychiatry. 2022 Feb;67(2):107-116. doi: 10.1177/07067437211006872.

3. Edwards J, Thind A, Stranges S, Chiu M, Anderson KK. Concordance between health administrative data and survey-derived diagnoses for mood and anxiety disorders. Acta Psychiatr Scand. 2020 Apr;141(4):385-395. doi: 10.1111/acps.13143.

4. Sordo Vieira L, Nguyen B, Nutley SK, Bertolace L, Ordway A, Simpson H, et al. Self-reporting of psychiatric illness in an online patient registry is a good indicator of the existence of psychiatric illness. J Psychiatr Res. 2022 Jul;151:34-41. doi: 10.1016/j.jpsychires.2022.03.022.

5. Davis KAS, Cullen B, Adams M, Brailean A, Breen G, Coleman JRI, et al. Indicators of mental disorders in UK Biobank—A comparison of approaches. Int J Methods Psychiatr Res. 2019 Sep;28(3):e1796. doi: 10.1002/mpr.1796.

6. O'Donnell S, Vanderloo S, McRae L, Onysko J, Patten SB, Pelletier L. Comparison of the estimated prevalence of mood and/or anxiety disorders in Canada between self-report and administrative data. Epidemiol Psychiatr Sci. 2016 Aug;25(4):360-369. doi: 10.1017/S2045796015000463.

7. Tannenbaum C, Lexchin J, Tamblyn R, Romans S. Indicators for measuring mental health: Towards better surveillance. Healthc Policy Polit Sante. 2009 Nov;5(2):e177-186.

8. Nacionalni inštitut za javno zdravje. Nacionalna raziskava o zdravju in zdravstvenem varstvu (EHIS 2019) [Internet]. 2020 [cited 2025 Feb 12]. Available from: https://www.nijz.si/sl/podatki/nacionalna-raziskava-o-zdravju-in-zdravstvenem-varstvem-ehis-2019

9. Nacionalni inštitut za javno zdravje. Metodološka pojasnila: Evidenca bolezni, zastrupitev in poškodb, ki zahtevajo zdravljenje v bolnišnici - BOLOB [Internet]. 2016 [cited 2025 Feb 12]. Available from: https://podatki.nijz.si/docs/6_6b_SBO_2016.pdf

10. Nacionalni inštitut za javno zdravje. Metodološka pojasnila: Izdana zdravila na recept (NIJZ 64) [Internet]. 2016 [cited 2025 Feb 12]. Available from: https://podatki.nijz.si/docs/6_6l_RECEPTI_2016.pdf

11. Nacionalni inštitut za javno zdravje. Metodološka pojasnila: Evidenca začasne / trajne odsotnosti z dela zaradi bolezni, poškodb, nege, spremstva in drugih vzrokov (NIJZ 3) [Internet]. 2018 [cited 2025 Feb 12]. Available from: https://podatki.nijz.si/docs/3_3f_BS_2018.pdf

12. Katschnig H, Straßmayr C, Endel F, Berger M, Zauner G, Kalseth J, et al. Using national electronic health care registries for comparing the risk of psychiatric re-hospitalisation in six European countries: Opportunities and limitations. Health Policy. 2019 Nov 1;123(11):1028–1035. doi: 10.1016/j.healthpol.2019.07.006.

13. Korošec D, Vrbnjak D, Štiglic G. Prevalence of mental disorders and high rates of absenteeism from work among healthcare professionals in Slovenia: A retrospective study. BMJ Open. 2023 Dec 1;13(12):e075718. doi: 10.1136/bmjopen-2023-075718.

14. Heberle H, Meirelles GV, Da Silva FR, Telles GP, Minghim R. InteractiVenn: A web-based tool for the analysis of sets through Venn diagrams. BMC Bioinformatics. 2015 May 22;16(1):169. doi: 10.1186/s12859-015-0611-3.

15. SPSS. IBM statistical package for social services (Version 25). Seattle WA IBM; 2018.

16. Cebron Lipovec N, Anderlic A, Locatelli I. General antidepressants prescribing trends 2009–2018 in Slovenia: A cross-sectional retrospective database study. Int J Psychiatry Clin Pract. 2022 Nov 1;26(4):401–405. doi: 10.1080/13651501.2022.2057331.

17. Farrer L, Leach L, Griffiths KM, Christensen H, Jorm AF. Age differences in mental health literacy. BMC Public Health. 2008 Apr 20;8(1):125. doi: 10.1186/1471-2458-8-125.

18. Lorem GF, Schirmer H, Wang CEA, Emaus N. Ageing and mental health: Changes in self-reported health due to physical illness and mental health status with consecutive cross-sectional analyses. BMJ Open. 2017 Jan;7(1):e013629. doi: 10.1136/bmjopen-2016-013629.

19. 19.   Kruijshaar ME, Barendregt J, Vos T, de Graaf R, Spijker J, Andrews G. Lifetime prevalence estimates of major depression: An indirect estimation method and a quantification of recall bias. Eur J Epidemiol. 2005 Jan 1;20(1):103–111. doi: 10.1007/s10654-004-1009-0.

20. Colombo D, Suso-Ribera C, Fernández-Álvarez J, Cipresso P, Garcia-Palacios A, Riva G, et al. Affect recall bias: Being resilient by distorting reality. Cogn Ther Res. 2020 Oct 1;44(5):906–918. doi: 10.1007/s10608-020-10122-3.

21. Gagné T, Henderson C, McMunn A. Is the self-reporting of mental health problems sensitive to public stigma towards mental illness? A comparison of time trends across English regions (2009–19). Soc Psychiatry Psychiatr Epidemiol. 2023;58(4):671–680. doi: 10.1007/s00127-022-02388-7.

22. Stewart H, Jameson JP, Curtin L. The relationship between stigma and self-reported willingness to use mental health services among rural and urban older adults. Psychol Serv. 2015;12(2):141–148. doi: 10.1037/a0038651.

23. Golberstein E, Eisenberg D, Gollust SE. Perceived stigma and mental health care seeking. Psychiatr Serv. 2008 Apr;59(4):392–399. doi: 10.1176/ps.2008.59.4.392.

24. Savage H, Murray J, Hatch SL, Hotopf M, Evans-Lacko S, Brown JSL. Exploring professional help-seeking for mental disorders. Qual Health Res. 2016 Oct 1;26(12):1662–1673. doi: 10.1177/1049732315591483.

25. Roskar S, Bracic MF, Kolar U, Lekic K, Juricic NK, Grum AT, et al. Attitudes within the general population towards seeking professional help in cases of mental distress. Int J Soc Psychiatry. 2017 Nov;63(7):614–621. doi: 10.1177/0020764017724819.

26. Vistorte AOR, Ribeiro WS, Jaen D, Jorge MR, Evans-Lacko S, Mari J de J. Stigmatizing attitudes of primary care professionals towards people with mental disorders: A systematic review. Int J Psychiatry Med. 2018 Jul 1;53(4):317–338. doi: 10.1177/0091217418778620.

27. Henderson C, Noblett J, Parke H, Clement S, Caffrey A, Gale-Grant O, et al. Mental health-related stigma in health care and mental health-care settings. Lancet Psychiatry. 2014 Nov 1;1(6):467–482. doi: 10.1016/S2215-0366(14)00023-6.

28. Foulkes L, Andrews JL. Are mental health awareness efforts contributing to the rise in reported mental health problems? A call to test the prevalence inflation hypothesis. New Ideas Psychol. 2023 Apr;69:101010. doi: 10.1016/j.newideapsych.2023.101010.

29. van Os J. The dynamics of subthreshold psychopathology: Implications for diagnosis and treatment. Am J Psychiatry. 2013 Jul;170(7):695–698. doi: 10.1176/appi.ajp.2013.13040474.

30. OECD. Measuring population mental health [Internet]. OECD; 2023 [cited 2025 Feb 12]. Available from: https://www.oecd-ilibrary.org/social-issues-migration-health/measuring-population-mental-health_5171eef8-en

31. Miroševič Š, Prins J, Bešić N, Borštnar S, Škufca Smrdel AC, Homar V, et al. Symptoms of anxiety and depression among Slovenian breast cancer survivors post-treatment during the COVID-19 pandemic: A cross-sectional study. Zdr Varst. 2024 Mar;63(1):55–62. doi: 10.2478/sjph-2024-0008.