Assessing mental health from registry data : what is the best proxy?

Assessing mental health from registry data: what is the best proxy?

Simon Gabriël Beerten*[1], MD, MSc

Robby De Pauw*[2], PhD

Gijs Van Pottelbergh[1], MD, PhD

Lidia Casas[3,4], MD, PhD

Bert Vaes[1], MD, PhD

(*) joint first authors

[1] *Department of Public Health and Primary Care, KU Leuven, Leuven, Belgium*

[2] *Sciensano, Brussels, Belgium*

[3] *Social Epidemiology and Health Policy, Department of Family Medicine and Population Health, University of Antwerp, Antwerp, Belgium*

[4] *Institute for Environment and Sustainable Development, University of Antwerp, Antwerp, Belgium*

Corresponding author: Simon Gabriël Beerten, MD, MSc

Kapucijnenvoer 7 blok h – bus 7001

3000 Leuven

E-mail: simon.beerten@kuleuven.be

Telephone: +32 16 19 40 50

Keywords: medical informatics, mental health, epidemiology

Word count: 2 327 words

ABSTRACT

Objective: Medical registries frequently underestimate the prevalence of health problems compared with surveys. This study aimed to determine the registry variables that can serve as a proxy for variables studied in a mental health survey.

Materials and methods: Prevalences of depressive symptoms, anxiety and psychoactive medication use from the 2018 Belgian Health Interview Survey (HIS) were compared with same-year prevalences from INTEGO, a Belgian primary care registry. Participants aged 15 and above were included. We assessed correlation using Spearman's rho (SR), and agreement using the intraclass correlation coefficient (ICC). We also calculated the limits of agreement (LOAs) for each comparison. HIS questions about depressive symptoms, anxiety and psychoactive medication use were compared with the following variables from INTEGO: symptom codes, diagnosis codes, free text, antidepressant/benzodiazepine prescriptions and the combinations symptom + diagnosis codes and symptom + diagnosis codes + free text, wherever relevant.

Results and discussion: Correlation between the HIS and INTEGO was generally high, except for anxiety. Agreement ranged from fair to poor, but increased when combining certain variables, by including free text, or by increasing the prescription frequency to resemble chronic use. Agreement remained poor when comparing questions about anxiety. Prevalences from INTEGO were mostly underestimates.

Conclusion: The external validity of medical registries can be poor, especially compared with survey data. A considerate choice of variables and prescription chronicity is needed to accurately use a registry as a surveillance tool for mental health.

BACKGROUND AND SIGNIFICANCE

Health registries are databases that collect clinical patient information from a specific healthcare setting (1). Some databases pertain to hospitals, while others collect data from a primary care setting. These data are usually extracted from an electronic medical record (EMR) kept by the general practitioner (GP), in a continuous fashion (2).

The modalities of each registry are of course adapted to its intended purpose, but many primary care registries include, besides general patient information, data on symptoms, diagnoses, prescriptions and laboratory test results. Most variables are entered in a coded fashion in the EMR by the GP for easy retrieval and for (inter)national comparison (2,3), for instance using the International Classification of Primary Care (ICPC) for the coding of diagnoses (4). Registries are very useful for measuring the burden of disease in a population (by calculating incidences and prevalences, for example), for health surveillance or for identifying areas of care in need of improvement (5). In many countries, information from EMRs is indeed actively used in disease surveillance projects (6), such as for chronic diseases (7) or COVID-19 (8).

There are, however, multiple challenges associated with this approach, which are situated on various levels, such as on an organizational, technical or data quality level (6).

Because the information in health registries is usually not collected for the purpose of specific epidemiologic research questions, using these registries to answer such questions may lead to measurement error and complex interpretation of the results (7). There are two principal issues in this context: registration differences and incomplete information. First, there could be differences in coding behavior between GPs or between regions, leading to differences in measures of occurrence such as prevalence and incidence (9).

Second, because the data collection is not driven by a predefined research question, some creativity is required to make use of the collected variables to answer such questions. Indeed, not all data on the variables that we need is actually present in a certain database, so we have to combine certain variables or use them as a proxy for the variable that we are looking for. Furthermore, information on exposures and potential confounders is also frequently lacking (10). Data accuracy, therefore, can be a real issue (11).

Assessing mental health from such a database is one of those challenges. Some prevalence studies done on mental health are survey-based, and they might capture subclinical or other forms of mental health issues unlikely to be brought to the attention of a physician (12). On the other hand, not every health issue brought up in surveys might necessarily be pathological or brought to the attention of a physician (12). In fact, a 'gold standard' in this context does not really exist, and medical registries are often taken as this standard when considering chronic diseases, but in the context of mental health this choice is far from clear (12).

Even though the data sources might be vastly different in terms of data collection methods and provenance of health information, it might nevertheless be useful to search for a common ground between these two, given the fact that both are frequently used in disease surveillance and can have their own advantages. Given the wealth of data usually included in registries (such as medication and comorbidities), they could prove very useful in mental health surveillance, when measurement error can be efficiently tackled. This can be done by comparing indicators across data sources (7).

For example, earlier work from the UK, focusing on the definition of mental health using codelists, found that depression was most often defined by diagnosis codes and/or prescriptions, whereas anxiety tended to be defined mostly using a mix of diagnosis and

symptom codes (13). Other comparisons in literature focus more on the prevalence of comorbidities in chronic disease (14) than on mental health, where the choice of indicators is less straightforward (15).

Objective

In order to fill this gap, in this case study, we compare depression and anxiety prevalence calculations from a mental health survey in Belgium with registry data from a large Belgian primary care database, using multiple strategies to select appropriate variables. Our main aim was to see which combination of variables from the database led to a prevalence calculation similar to that from the Belgian health survey.

MATERIALS AND METHODS

Study design

Databases

In this case study, we will compare two sources of data.

The first dataset is INTEGO. This database contains clinical data of over 600.000 primary care patients all over Flanders, Belgium, including but not limited to diagnoses, prescriptions and lab results. Medication in INTEGO is coded according to the Anatomical Therapeutic Chemical Classification System (ATC). The most relevant ATC codes for depression and anxiety are N05B (anxiolytics), N05C (hypnotics) and N06A (antidepressants). Diagnoses are coded according to the ICPC version 2 (ICPC-2). The full background and methodology of the INTEGO database is described elsewhere (3). The main purpose for the creation of this dataset, back in 1994, was to act as a tool for disease surveillance and improvement of clinical practice in primary care.

The second dataset is the Health Interview Survey (HIS), conducted by the Belgian public health institute Sciensano (16) and consisting of a comprehensive health assessment across different domains, one of which is mental health (17,18). The main purpose of this survey is to identify the main health problems in the population, thereby informing and guiding public policy. The survey is conducted every five years, with 2018 being the latest completed survey. Assessment occurs either in the form of a face-to-face interview or by self-completed questionnaire. Detailed protocols are available online (18). For this study, we analyzed the Flemish sample.

The data collection period for the 2018 HIS survey ran from February 1, 2018 until January 26, 2019. Given that some HIS questions probed medication use and mental health issues for up to one year before, the study period in INTEGO was chosen between February 1, 2017 and December 31, 2018.

Study population

The population for this case study included any patient 15 years or older. Patients in INTEGO come from general practices across Flanders (3) respondents of the HIS were sampled from the entire Belgian population, separately for every region (18). Only the results from Flanders were included in this study.

Results are stratified by sex and age group (15-24, 25-34, 35-44, 45-54, 55-64, 65-74 and 75+ years old). The denominator used for INTEGO was the yearly contact group, i.e. the total number of patients visiting a practice in the network at least once in a given year (19). The yearly contact group of 2018 was chosen as the denominator in this study. The denominator for the HIS was the total size of the respondent cluster (18). The outcome measure for each database was the prevalence proportion per stratum. Prevalence proportions in the HIS were weighted according to the sampling design.

Methodology

Variables

In what follows, we will attempt to define what is the best strategy to approach the prevalence calculations from the HIS in terms of specific variable selection from the INTEGO database. (**Table 1**), together with the expressions used to look for free text indicators (**Table 2**).

Table 1: Variables available in INTEGO for developing a case definition of depression and anxiety.

Variable		Comment
1. Prescriptions	ATC prescription codes for antidepressants (N06A) and anxiolytics/hypnotics (N05B/N05C)	Possible false-positives (e.g. chronic pain, sleep disorders) Mild, non-medicated cases will be missed
2. Symptom codes	ICPC-2 codes P01 ('feeling anxious) and P03 ('feeling depressed')	Accuracy depending on coding practice and diligence

3. Diagnosis codes	ICPC-2 codes P74 ('anxiety disorder') and P76 ('depressive disorder')	Mild cases might not get a diagnosis code; accuracy depending on coding practice
4. Free-text search	Any free text in the patient file indicating depression or anxiety (see **Table 2**)	Very sensitive, but also potentially very non-specific

Abbreviations: ATC = Anatomical Therapeutic Classification, ICPC-2 = International Classification of Primary Care, version 2.

Table 2: Regex expressions used to search for free text indicators.

Expression	Derived search terms	Interpretation	Comment
"neersl|depr|hopel"	neerslachtig ("dejected") depressief ("depressed") hopeloos ("hopeless")	Detects any entry that contains one or more of the terms separated by '|' anywhere in the text	Used to search for indicators of *depression*
"angst|nerv|zenuwa|opgej"	angst(ig) (anxiety/anxious) nerveus (nervous) zenuwachtig (nervous) opgejaagd (agitated)		Used to search for indicators of *anxiety*

Search strategy

The HIS questions that were tested in this case study are presented in **Table 3**.

Table 3: Questions from the Health Interview Survey 2018 (18).

Question 1	During the past 12 months, have you had any of the following diseases or conditions: *Serious gloom or depression for a period of at least 2 weeks*
Question 2	Did you take any medicines for this during the past 12 months?
Question 3	During the past 2 weeks, have you used any sleeping tablets or tranquillizers that were prescribed for you by a doctor?

Question 4	During the past 2 weeks, have you used any antidepressants that were prescribed for you by a doctor?
Question 5	Over the last two weeks, have you been bothered by the following problems: *Feeling nervous, anxious or on edge*
Question 6	In the past few weeks, have you been feeling unhappy or depressed?

Question 1 to 4 are yes-no questions, while questions 5 and 6 are scored on a 4-point frequency Likert scale (respectively: “not at all”, “several days”, “more than half the days”, “nearly every day”, and “not at all”, “no more than usual”, “rather more than usual”, “much more than usual”).

The search strategy for each question is as follows.

- Question 1: This question pertains to the presence of depression in the respondent, lasting for a minimum of two weeks within the past year. For INTEGO, the following ATC and ICPC codes and combinations were used: P03 (feeling depressed, P76 (depression), free text, P76 + P03, P76 + P03 + free text, N06A (antidepressants) 1+ prescription, N06A 3+ prescriptions.

- Questions 2 to 4 in the HIS are about the use of psychopharmaceuticals, either antidepressants or benzodiazepines (here described separately as sleeping tablets and tranquillizers). We compared these questions with the ATC codes N06A, N05B and N05C, respectively, available in the INTEGO dataset.

- Question 5: This question assessed feelings of anxiety in the HIS survey respondents and was tested against the following INTEGO variables or combinations: P01 (feeling anxious), P74 (anxiety disorder), free text, P01 + P74, P01 + P74 + free text, N05B + N05C (benzos).

- Question 6: A variant of question 1, it probed for the presence of depressed mood instead of anxiety. Search strategies for variables were analogous to those for question 5.

Statistical analysis

In order to compare the prevalence calculations between the databases, we assessed both correlation and agreement. Correlation was tested using Spearman's ρ to assess monotonic relationships between the both datasets. Agreement was tested using the two-way 'agreement' intraclass correlation coefficient (ICC), to assess pair-wise differences between the measurements of the two datasets.

Additionally, to have a visual representation of the level of agreement between the databases, we performed Bland-Altman analyses and calculated the corresponding limits of agreement (LOAs) at 95% confidence. The narrower these intervals, the more the measurements between the two datasets align.

All statistical analysis was done in R, version 4.2.3. The *survey* package was used to calculate the weighted prevalence proportions in the HIS. We used the *blandr* and *irr* packages for the Bland-Altman and ICC calculations, respectively. Statistical significance was set at 5%.

RESULTS

The INTEGO population used for this study included 267 665 unique patients (contact group of 2018), of whom 54% were female. Search strategy results for all HIS questions are presented in **Table 4**.

Table 4: HIS questions compared to different INTEGO strategies.

Question 1: Serious gloom or depression for a period of at least 2 weeks			
Variable or combination (INTEGO)	**Spearman's ρ**	**Intraclass correlation coefficient**	**Limits of agreement (95% confidence)**
P03 (feeling depressed)	0.763 (p = .002)	0.026 (p = .320)	[0.010; 0.107]
P76 (depression)	0.653 (p = .014)	0.195 (p = .167)	[-0.006; 0.077]

Free text	0.653 (p = .014)	0.322 (p = .111)	[-0.013; 0.066]
P76 + P03	0.640 (p = .016)	0.270 (p = .134)	[-0.010; 0.070]
P76 + P03 + free text	0.653 (p = .014)	0.386 (p = .085)	[-0.015; 0.061]
N06A (antidepressants) 1+ prescription	0.793 (p = .001)	0.391 (p = .098)	[-0.120; 0.021]
N06A 3+ prescriptions	0.736 (p = .004)	0.583 (p = .012)	[-0.082; 0.039]
Question 2: Medication for gloom or depression during the past 12 months			
N06A, 1+ prescription	0.345 (p = .227)	0.112 (p = .210)	[0.006; 0.543]
N06A, 3+ prescriptions	0.323 (p = .260)	0.099 (p = .218)	[0.071; 0.588]
Question 3: Use of prescribed sleeping tablets or tranquillizers during the past 2 weeks			
N05B + N05C (benzos) 1+ prescription	0.899 (p = <.001)	0.822 (p = .001)	[-0.127; 0.063]
Question 4: Use of prescribed antidepressants during the past 2 weeks			
N06A, 1+ prescription	0.899 (p = <.001)	0.586 (p = .091)	[-0.093; -0.008]
N06A, 3+ prescriptions	0.895 (p = <.001)	0.835 (p = .026)	[-0.053; 0.007]
Question 5: Feeling nervous, anxious or on edge over the last two weeks			
P01 (feeling anxious)	0.178 (p = .542)	0.000 (p = .488)	[0.343; 0.881]
P74 (anxiety disorder)	-0.604 (p = .025)	-0.002 (p = .561)	[0.331; 0.882]
Free text	-0.433 (p = .124)	-0.003 (p = .584)	[0.307; 0.863]
P01 + P74	-0.415 (p = .141)	-0.002 (p = .549)	[0.326; 0.875]
P01 + P74 + free text	-0.455 (p = .104)	-0.004 (p = .607)	[0.301; 0.861]
N05B + N05C	0.468 (p = .094)	0.044 (p = .247)	[0.210; 0.704]
Question 6: Unhappy or depressed in the past few weeks			
P03	-0.222 (p = .445)	-0.000 (p = .538)	[0.371; 0.618]
P76	-0.301 (p = .295)	-0.003 (p = .730)	[0.337; 0.607]
Free text	-0.301 (p = .295)	-0.004 (p = .745)	[0.326; 0.601]
P03 + P76	-0.284 (p = .325)	-0.004 (p = .737)	[0.330; 0.603]
P03 + P76 + free text	-0.301 (p = .295)	-0.004 (p = .748)	[0.322; 0.598]
N06A, 1+ prescription	-0.068 (p = .820)	-0.006 (p = .690)	[0.215; 0.560]
N06A, 3+ prescriptions	-0.042 (p = .892)	-0.004 (p = .679)	[0.253; 0.577]

For question 1, correlation was moderate to high with every strategy and highest for the ATC code N06A (antidepressants), if the criterion of at least one prescription per year was used. The agreement ICC fluctuated and tended to increase when combining strategies. Overall agreement for any strategy was relatively poor. The best combinatory strategy (i.e. combining different variables, and with the highest correlation and agreement, and the narrowest LOAs) was the use of the ICPC-codes P76, P03 and free text occurrences indicative of depression. The best single-variable strategy was N06A (at least 3 prescriptions per year), although with broad LOAs.

Questions 2 and 4 were similar, except for the duration of medication use being queried. Question 2, however, only needed an answer if the response to question 1 had been positive. Therefore, medication use for this question was assessed for only those patients identified by the most sensitive variable combination (P76 + P03 + free text) for question 1. Correlation and agreement for question 2 were both very poor, with broad LOAs. On the other hand, correlation for all strategies for questions 3 and 4 were very high, and agreement increased considerably when multiple prescriptions per year were considered, as a proxy for chronic use.

Overall correlation for questions 5 and 6 was very poor and even negative for some strategies. The ICC was distributed around zero, indicating little to no agreement between the HIS and INTEGO. Judging from the LOAs, all INTEGO variables or combinations resulted in profound underestimations.

The Bland-Altman plots for each tested combination of variables, as well as dummy INTEGO datasets and the aggregated data used in this study, can be found in **Appendix**.

We also tested the agreement ICC per age group and sex category, which can be found in Tables A6 to A17. Generally, agreement was higher for female participants and higher in the youngest age categories, depending on the question (see **Appendix**).

DISCUSSION

General

In this study, we found generally high correlation between the HIS and INTEGO, except for questions related to anxiety. Agreement was fair to poor, but could be increased by combining certain variables, by including free text, or by increasing the prescription frequency. On the whole, prevalence calculations from INTEGO were underestimates. This study is, to our knowledge, the first to pilot a possible approach to comparing mental health assessments between a survey and a health registry in Belgium. Conceptually, the design is similar to two studies by Joling et al. and Janssen et al. on depression and anxiety, respectively (20,21). Another Danish paper on the same subject was recently published, albeit using different methodology (22).

In the aforementioned Dutch studies, the main goal was to analyze how good GPs were at recognizing depression and anxiety by comparing their clinical assessments to a paired sample from a reference survey (20,21). The main conclusions were: 1) antidepressant prescriptions are the best single indicator for depression and 2) the combination of ICPC-codes P01, P74 and P79, free text, referrals and prescriptions is the best indicator for anxiety. The authors used Youden's J (sensitivity + specificity – 1) to estimate accuracy (23). For depression, our results were in line with their findings. For anxiety, we found no relevant correlation or agreement at all. In another Spanish study, a combination of diagnoses and prescriptions, but mostly prescriptions, gave higher agreement between a health registry and a survey.(24)

As our study population consisted of two independent samples, sensitivity and specificity calculations were not possible. Instead, we tested two important pillars of

database accuracy: correlation and agreement (25). The latter was also calculated and visualized using Bland-Altman analyses.

The comparison between the databases allowed us to gain insights into how data may vary between sources and which variables are best suited to approach real-life situations. The main conclusion in similar studies is that of underestimation on the side of registries, as they fail to capture subclinical pathology which does not present to physicians but is likely to be picked up in surveys (12,22). Our results confirm this. Very broad questions about feelings of depression and anxiety yielded poor results when tested against the limiting nature of available clinical variables in the registry. This, of course, should be no surprise, considering the fact that transient feelings of low mood or distress are a part of life, experienced by every individual at some point in time (26). Health registries are more likely to capture pathological conditions than everyday occurrences.

Furthermore, our results indicate that the choice of variables matters when using clinical registries. Although they cannot possibly replace surveys, they are nonetheless used as a surveillance tool for many diseases, including mental health problems (5,27). A possible case definition of depression or anxiety, using the INTEGO database, was more accurate when multiple variables were combined. Yearly prescription frequency was also used in other studies to mimic chronic medication use [for example for benzodiazepines (28,29)], as there is no standardized definition yet. Future research using this type of registry has to test and then unify multiple definitions of chronic use (by varying sensitivity and specificity), especially in a primary care population.

Strengths and limitations

The main strength of this study is the inclusion of a wealth of data from two sources, both statistically representative for the Flemish population. Furthermore, the INTEGO

database is situated in primary care, which is a low-threshold setting for patients presenting with most mental health issues.

The main limitation of this study is the fact that respondents were not linked between databases, as in similar studies from the Netherlands and Denmark (20–22). Due to privacy regulations, this would not have been possible to do on such a small scale as in this study. We also did not include hospital data in our study, which could have led to underestimation of disease burden on the side of the registry.

CONCLUSION

This study analyzed correlation and agreement between a health registry and a survey database for different questions regarding depression and anxiety. While generally acceptable correlation and agreement for depression were found, the reverse was true for anxiety.

Therefore, this work should serve as a cautious reminder of the importance of careful variable selection and consideration of prescription chronicity in surveillance and prevalence studies, especially for mental health. The future linkage between respondents over the two databases will provide more robust results.

FUNDING INFORMATION

This work was supported by a research project from the Research Foundation Flanders (reference: G0B4820N).

ETHICAL APPROVAL

The INTEGO procedures used in this study are approved by the ethical review board of the KU Leuven Faculty of Medicine (no. ML 1723) and by the Belgian

Privacy Commission (no. SCSZG/13/079). INTEGO was waived the need for individual informed consent, but has an opt-out procedure for patients who do not wish their data to be included.

All procedures in this study are in accordance with current national guidelines and regulations.

CONFLICTS OF INTEREST

The authors report no conflicts of interest.

SUMMARY POINTS

What was already known on the topic
• Health registries, such as those based on electronic medical records, are important in health surveillance
• In comparison with survey databases, health registries frequently underestimate the burden of disease, also in the context of mental health
What this study added to our knowledge
• For epidemiologic measures of depression and anxiety, health registries and survey databases show generally good correlation
• Inter-database agreement can be poor, especially concerning measures of anxiety
• For health registries, a considerate variable selection is needed for use in mental health surveillance

REFERENCES

1. Solomon DJ, Henry RC, Hogan JG, Van Amburg GH, Taylor J. Evaluation and implementation of public health registries. Public Health Rep. 1991;106(2):142–50.

2. Hiddema-van de Wal A, Smith RJ, van der Werf GT, Meyboom-de Jong B. Towards improvement of the accuracy and completeness of medication registration with the use of an electronic medical record (EMR). Fam Pract. 2001 Jun;18(3):288–91.

3. Truyers C, Goderis G, Dewitte H, Akker M vanden, Buntinx F. The Intego database: background, methods and basic results of a Flemish general practice-based continuous morbidity registration project. BMC Med Inform Decis Mak. 2014 Jun 6;14(1):48.

4. Classification Committee of the World Organization of Family Doctors (WICC). International Classification of Primary Care. Oxford: Oxford University Press; 1997.

5. Pop B, Fetica B, Blaga ML, Trifa AP, Achimas-Cadariu P, Vlad CI, et al. The role of medical registries, potential applications and limitations. Med Pharm Rep. 2019 Jan;92(1):7–14.

6. Aliabadi A, Sheikhtaheri A, Ansari H. Electronic health record–based disease surveillance systems: A systematic literature review on challenges and solutions. J Am Med Inform Assoc. 2020 Dec 9;27(12):1977–86.

7. Mcveigh K, Newton-Dame R, Perlman S, Chernov C, Thorpe L, Singer J, et al. Developing an Electronic Health Record-Based Population Health Surveillance System. 2013 Jul 1;

8. Sheikhtaheri A, Tabatabaee Jabali SM, Bitaraf E, TehraniYazdi A, Kabir A. A near real-time electronic health record-based COVID-19 surveillance system: An experience from a developing country. Health Inf Manag J Health Inf Manag Assoc Aust. 2022 Jul 15;18333583221104212.

9. van den Dungen C, Hoeymans N, Gijsen R, van den Akker M, Boesten J, Brouwer H, et al. What factors explain the differences in morbidity estimations among general practice registration networks in the Netherlands? A first analysis. Eur J Gen Pract. 2008;14 Suppl 1:53–62.

10. Nørgaard M, Ehrenstein V, Vandenbroucke JP. Confounding in observational studies based on large health care databases: problems and potential solutions – a primer for the clinician. Clin Epidemiol. 2017 Mar 28;9:185–93.

11. Barkhuysen P, de Grauw W, Akkermans R, Donkers J, Schers H, Biermans M. Is the quality of data in an electronic medical record sufficient for assessing the quality of primary care? J Am Med Inform Assoc JAMIA. 2014 Aug;21(4):692–8.

12. 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):360–9.

13. Carreira H, Williams R, Strongman H, Bhaskaran K. Identification of mental health and quality of life outcomes in primary care databases in the UK: a systematic review. BMJ Open. 2019 Jul 2;9(7):e029227.

14. Violán C, Foguet-Boreu Q, Hermosilla-Pérez E, Valderas JM, Bolíbar B, Fàbregas-Escurriola M, et al. Comparison of the information provided by electronic health records data and a population health survey to estimate prevalence of selected health conditions and multimorbidity. BMC Public Health. 2013 Mar 21;13(1):251.

15. McVeigh KH, Newton-Dame R, Chan PY, Thorpe LE, Schreibstein L, Tatem KS, et al. Can Electronic Health Records Be Used for Population Health Surveillance? Validating Population Health Metrics Against Established Survey Data. EGEMS Wash DC. 2016;4(1):1267.

16. Sciensano. About Sciensano. [Internet]. Available from: https://www.sciensano.be/en/about-sciensano.

17. Gisle L, Drieskens S, Demarest S, Heyden J. Geestelijke gezondheid: gezondheidsenquête 2018. Brussels, Belgium: Sciensano; 2018.

18. Sciensano. sciensano.be. [cited 2023 Apr 16]. Health Interview Survey. Available from: https://www.sciensano.be/en/projects/health-interview-survey

19. Bartholomeeusen S, Kim CY, Mertens R, Faes C, Buntinx F. The denominator in general practice, a new approach from the Intego database. Fam Pract. 2005 Aug 1;22(4):442–7.

20. Joling KJ, van Marwijk HWJ, Piek E, van der Horst HE, Penninx BW, Verhaak P, et al. Do GPs' medical records demonstrate a good recognition of depression? A new perspective on case extraction. J Affect Disord. 2011 Oct;133(3):522–7.

21. Janssen EHC, van de Ven PM, Terluin B, Verhaak PFM, van Marwijk HWJ, Smolders M, et al. Recognition of anxiety disorders by family physicians after rigorous medical record case extraction: results of the Netherlands Study of Depression and Anxiety. Gen Hosp Psychiatry. 2012 Oct;34(5):460–7.

22. Weye N, McGrath JJ, Lasgaard M, Momen NC, Knudsen AK, Musliner K, et al. Agreement between survey- and register-based measures of depression in Denmark. Acta Psychiatr Scand [Internet]. [cited 2023 May 1];n/a(n/a). Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/acps.13555

23. Youden WJ. Index for rating diagnostic tests. Cancer. 1950;3(1):32–5.

24. Orueta JF, Nuño-Solinis R, Mateos M, Vergara I, Grandes G, Esnaola S. Monitoring the prevalence of chronic conditions: which data should we use? BMC Health Serv Res. 2012 Oct 22;12:365.

25. LIU J, TANG W, CHEN G, LU Y, FENG C, TU XM. Correlation and agreement: overview and clarification of competing concepts and measures. Shanghai Arch Psychiatry. 28(2):115–20.

26. Galderisi S, Heinz A, Kastrup M, Beezhold J, Sartorius N. Toward a new definition of mental health. World Psychiatry. 2015 Jun;14(2):231–3.

27. McBrien KA, Souri S, Symonds NE, Rouhi A, Lethebe BC, Williamson TS, et al. Identification of validated case definitions for medical conditions used in primary care electronic medical record databases: a systematic review. J Am Med Inform Assoc JAMIA. 2018 Nov 1;25(11):1567–78.

28. Coteur K, Mamouris P, Vaes B, Van Nuland M, Matheï C, Schoenmakers B. Evolution of benzodiazepine receptor agonist prescriptions in general practice: A registry-based study. Front Public Health. 2022 Jan 1;10:1014734.

29. Kurko T a. T, Saastamoinen LK, Tähkäpää S, Tuulio-Henriksson A, Taiminen T, Tiihonen J, et al. Long-term use of benzodiazepines: Definitions, prevalence and usage patterns - a systematic review of register-based studies. Eur Psychiatry J Assoc Eur Psychiatr. 2015 Nov;30(8):1037–47.