


RESEARCH ARTICLE

Open Access



# Estimating the prevalence of depression associated with healthcare use in France using administrative databases

Antoine Filipovic-Pierucci<sup>1,2\*</sup> , Solène Samson<sup>2</sup>, Jean-Paul Fagot<sup>2</sup> and Anne Fagot-Campagna<sup>2</sup>

## Abstract

**Background:** Quantitative indicators are needed in order to define priorities, plan policies and evaluate public health interventions in mental health. The aim of this study was to assess the contribution of a large and exhaustive French national administrative database to study and monitor treated depression by comparing the prevalence and characteristics of the population using significant healthcare resources for depression as identified by different estimation methods and sources and to discuss the advantages and drawbacks of these methods.

**Methods:** This study included the French population covered by the main health insurance scheme in 2012 (Régime général, 86% of the insured French population). Data were extracted from the French health insurance claim database (SNIIRAM), which contains information on all reimbursements, including treatments and hospital stays in France. The following distinct sources of the SNIIRAM were used to select persons with depression: diagnoses of long-term or costly conditions, data from national hospital claims and data concerning all national health insurance reimbursements for drugs.

**Results:** In 2012, we included 58,753,200 individuals covered by the main health insurance scheme; 271,275 individuals had full coverage for depression; 179,470 individuals had been admitted to a psychiatric hospital and 66,595 individuals admitted to a general hospital with a diagnosis of depression during a 2-year timeframe and 144,670 individuals had more than three reimbursements for antidepressants during the study year (with a history of hospitalisation for depression during the past 5 years). Only 16% of individuals were selected by more than one source.

**Conclusions:** We propose an algorithm that includes persons recently hospitalised for depression, or with a history of hospitalisation for depression and still taking antidepressants, or with full coverage for depression as a specific long-term or costly condition, yielding a prevalence estimate of 0.93% or 544,105 individuals. Changes in the case selection methodology have major consequences on the frequency count and characteristics of the selected population, and consequently on the conclusions that can be drawn from the data, emphasizing the importance of defining the characteristics of the target population before the study in order to produce relevant results.

**Keywords:** Depression, Epidemiology, Administrative databases, Methodology, France, Prevalence

\* Correspondence: antoine.filipovic-pierucci@urc-eco.fr

<sup>1</sup>URC-Eco, Health economics and health policy research unit, AP-HP, Hôtel Dieu, Galerie B1–3ème étage, 1 Place du Parvis Notre Dame, Paris 75004, France

<sup>2</sup>CNAMTS (National Health Insurance), Paris, France



## Background

According to the World Health Organisation, 14% of the global burden of disease is attributed to mental health disorders [1], among which unipolar depression is the second leading cause of life years with disability worldwide. Depression per se is a major contributor to the global burden of disease [2, 3] and is associated with a high economic burden [4, 5], making it a major public health priority.

Quantitative indicators are needed in order to define priorities, plan policies and evaluate public health interventions in mental health. Few studies have estimated the prevalence of depression in the general population in France, and they were based on medical interviews, symptom questionnaires, self-reported diagnoses, treatments, or full health insurance coverage with a diagnosis of depression in administrative databases [6]. Furthermore, as studies conducted according to different methodologies can provide different perspectives on a particular issue, multiple sources of information may be complementary to more accurately study depression [7].

Data about hospital healthcare use in psychiatry are now almost exhaustively available in France with the recent improved quality of the national hospital claim databases in psychiatry [8], in addition to other data sources from general hospitals or insurance claims. The contribution of a large and exhaustive national administrative database to study and monitor treated depression should be assessed, especially in terms of the potential contribution of combining information from different sources. Although such studies already exist in North America [9–13], their findings may not apply to the French setting with its specific healthcare system and information sources. Furthermore it may be of interest to compare United States or Canadian results with data from another setting.

The French national health insurance scheme, CNAMTS, has developed a tool to study the medical and economic burden and the care pathway of multiple chronic diseases, including psychiatric disorders [14]. Various methods were tested to select people who had received care for depression, which yielded different populations. The most appropriate and accurate method may vary depending on whether the study goal is epidemiological monitoring, care pathway description or medico-economic analysis. As an insurance scheme, the goal of the CNAMTS study was to estimate the population receiving healthcare as well as the amount of healthcare resources allocated to depression, and was therefore designed to select a population using significant healthcare resources as a result of this disease. This study was a first step before analysing factors such as trends over time and cost patterns, regional health profiles, care pathways, outcomes, comorbidities, etc.

The objective of this study was (1) to examine the prevalence and characteristics of the population using significant healthcare resources for depression as identified by different estimation methods and sources and (2) to discuss the advantages and drawbacks of these methods.

## Methods

### Study population

This study was restricted to the population covered by the main health insurance scheme (*Régime général*), which represented 86% of the insured French population in 2012, as not all information was available for the other insurance schemes at that time. Almost the entire French population is insured [15]. All persons with at least one reimbursement for medical care in 2012 were included in the study. Same-sex twins were excluded from the study because of linkage problems specific to this population.

### Data

Data were extracted from the French health insurance claim database (SNIIRAM), which contains information on all reimbursements, including treatments, diagnostic and therapeutic procedures, and hospital stays in France for almost the entire population. The database is not public; analyses by CNAMTS were allowed by CNIL, the independent French administrative regulatory body whose mission is to ensure that data privacy law is applied to the collection, storage, and use of personal data [16]. Diagnoses related to outpatient visits and results of procedures are not reported in the database [17].

### Selection of persons with depression

The following distinct sources of the SNIIRAM were used to select persons with depression:

Source 1: Diagnoses of long-term or costly conditions (*Affections de Longue Durée*, ALD). Patients with specific long-term or costly conditions may require full coverage for all their condition-related health expenditures upon request by their family doctor and after approval by a health insurance fund medical officer (*médecin-conseil*) [18].

Source 2: Data from national hospital claims (*Programme de Médicalisation des Systèmes d'Information*, PMSI) for all inpatient and day-case admissions in public and private general and psychiatric hospitals, containing medical diagnoses defined as ICD-10 codes. In both general and psychiatric hospitals, a principal diagnosis is defined as the main reason for admission, while associated diagnoses provide information about

conditions that significantly influenced care during the hospital stay [19].

Source 3: Data concerning all national health insurance reimbursements for drugs, laboratory tests and outpatient medical procedures. Individuals receiving reimbursements for antidepressants (N06A section of the ATC classification except for oxitriptan) can be identified. However, these databases do not contain direct information about the diagnosis justifying the prescription, and these drugs are not specific for depression, as they can also be prescribed for other conditions (bipolar disorders, anxiety or chronic pain). An antidepressant prescription is typically valid 1 month.

All three sources were not considered to be equally reliable for identifying patients with depression. Reliability of the sources was assessed as follows: for the purpose of identifying individuals suffering from depression, full coverage for depression as a specific long-term or costly condition (source 1) was more reliable than the hospital claims database (source 2), which was more reliable than reimbursement for antidepressants (source 3). In the hospital claims database, associated diagnoses reported during general hospital stays were assumed to be less reliable than those reported during psychiatric hospital stays. The reasons underlying this classification of source reliability included (1) the mode of acquisition of the information (diagnoses resulting from medical interviews were regarded as more reliable than hospital diagnostic codes sometimes coded by non-medical staff, themselves regarded as a more reliable diagnostic markers than prescription drugs) and (2) what was at stake when the information was coded (hospital diagnostic codes that had no consequence on costs were regarded as less reliable than codes influencing costs or giving access to benefits). These reasons are described and discussed more thoroughly in the Merits and drawbacks of the various methods section of the Discussion section of this article.

Accordingly, five estimation methods with decreasing order of reliability were defined. ICD-10 codes F32 to F39 were used in all estimation methods to identify depression (either as a full health coverage code or as a principal or associated diagnosis). At least three reimbursements for antidepressants were used to identify treatment by antidepressant. Hospital stays in the last 5 years with a principal or associated diagnosis of depression were used to identify principal diagnosis history and associated diagnosis history of depression respectively.

- Method A (Full coverage for depression): Selection of individuals with full coverage for depression as a specific long-term or costly condition during the study (source 1);

- Method B (Hospitalisation for depression): Selection of individuals with depression as *principal* or *associated* diagnosis in a psychiatric hospital stay or as *principal* diagnosis in a general hospital stay using two timeframes: (a) the current calendar year and (b) the last two calendar years (source 2). Calendar years were used for technical reasons.
- Method C (Current antidepressant treatment + History of hospitalisation during the past 5 years): Selection of individuals treated by antidepressant *and* with a general hospital *principal* diagnosis history of depression or a psychiatric hospital *principal* or *associated* diagnosis history of depression (combination of sources 2 and 3);
- Method D (Hospitalisation in a general hospital with an associated diagnosis of depression): Selection of individuals with depression as *associated* diagnosis in a general hospital stay using two timeframes: (a) the current calendar year and (b) the last two calendar years (source 2);
- Method E (Current antidepressant treatment + History of hospitalisation in a general hospital with an associated diagnosis of depression during the past 5 years): Selection of individuals treated by antidepressant *and* with a general hospital *associated* diagnosis history of depression (combination of sources 2 and 3).

Individuals with a hospital diagnosis of bipolar disorder (ICD-10 codes F30 or F31) in the last 5 years or a specific treatment for bipolar disorder (lithium, divalproex or valpromide) were not included in the study.

#### Statistical analysis

Populations selected by these different methods were described by age and gender. The incremental contribution of each method (e.g. the contribution of adding method B in the presence of method A) was assessed. Frequency counts were rounded to the nearest 5.

Mortality rates for a given method were calculated as the number of individuals selected by the method who died during the year divided by the total number of individuals selected by the method. Mortality was measured by the administrative vital status of the individuals, as informed by French civil registry data.

Qualitative variables were described by frequency counts, and quantitative variables were described by means. All reported differences were statistically significant with an alpha risk of 5%. Because of the high statistical power of this study, statistically significant differences were only reported and discussed when they were considered to be clinically relevant. Statistical

analysis was performed with SAS software version 9.2 and proportional Venn diagrams were drawn using EulerAPE software [20].

## Results

### Frequency counts by data sources

In 2012, 58,753,200 individuals covered by the main health insurance scheme received at least one reimbursement for any cause [21]. Frequency counts by estimation method are presented in Table 1. In this population, method A identified 271,275 individuals with full coverage for depression; method B identified a total of 229,020 individuals: 109,260 individuals who had been admitted to a psychiatric hospital with a principal or associated diagnosis of depression and 34,630 individuals admitted to a general hospital with a principal diagnosis of depression during the study year. Those counts increased to 179,470 and 66,595, respectively, with aggregate counts over a 2-year timeframe, for a total aggregate count of 229,100 individuals. Method C identified 137,110 individuals with more than three reimbursements for antidepressants during the study year and with a history of hospitalisation for depression during the past 5 years in a psychiatric hospital, 51,420 in a general hospital, for a total of 144,670 individuals with aggregate counts over psychiatric and general hospitals.

Frequency counts and intersections between data sources are presented using a Venn diagram (Fig. 1); only 16% of individuals were selected by more than one method. Individuals taking antidepressants and with a history of hospitalisation for depression (method C) were more frequently selected by another source (44% of these individuals) than individuals with hospitalisation in the past 2 years (method B) and individuals with full coverage for depression (method A), 30 and 20%, respectively, were also selected by another source.

Method D identified 213,010 individuals who were admitted to a general hospital with an associated diagnosis of depression during the current year (340,270 with aggregate counts over a 2-year timeframe, only 14% were detected by another source). Method E identified 266,890 individuals who were treated with antidepressants and who had a history of hospitalisation with an associated diagnosis of depression during the last 5 years (by construction, most of these individuals were already detected by method D).

### Incremental contribution of data sources

The incremental contribution of data sources was studied in order to quantify the specific contribution of a data source in the presence of the other sources. In addition to the 271,275 individuals with full coverage for depression (method A), method B (selecting individuals with a hospital diagnosis except for general hospital

associated diagnoses for depression) added 192,255 individuals to the total (113,435 hospitalised during the current year and another 78,820 hospitalised during the previous year, aggregate counts over psychiatric and general hospitals). Selecting persons taking treatments for depression and with a history of hospitalisation for depression (method C) added another 80,570 individuals (aggregate counts over psychiatric and general hospitals), for a total of 544,105 individuals.

Furthermore, method D (selecting individuals with an associated diagnosis of depression during a general hospital stay) added another 291,925 persons to this total (164,660 individuals hospitalised during the current year and another 127,265 individuals hospitalised during the previous year). Finally, method E added 133,270 individuals (persons with a treatment for depression and a history of hospitalisation for depression defined as a general hospital associated diagnosis of depression in the last 5 years), for a total of 969,300 individuals.

This incremental method yielded prevalence estimates ranging from 0.46% for the most conservative method (method A—only the population with full health insurance coverage for depression) to 1.65% for the less conservative method (using all methods A to E). Intermediate methods that combined full health insurance coverage for depression with hospitalisations for depression, but not including general hospital stays with associated diagnoses of depression yielded estimates ranging from 0.61% (methods A + B) to 0.93% (methods A + B + C).

### Population characteristics by data sources

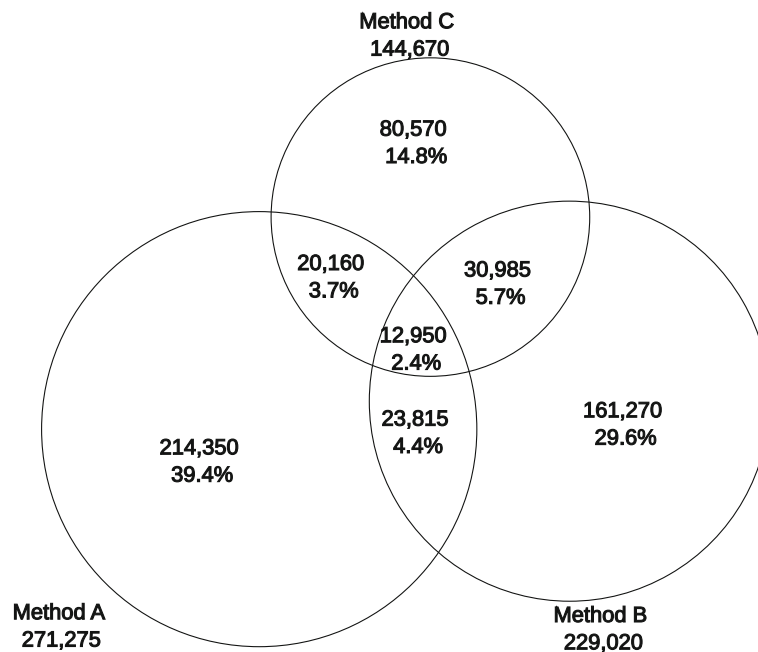
Population characteristics varied according to the data source (Table 1). Individuals with full coverage for depression as a specific long-term or costly condition, those taking antidepressants with a history of hospitalisation for depression and those hospitalised with an associated diagnosis of depression were older than the individuals selected by the other sources (58–62% vs. 50–54%). The proportion of persons taking antidepressants was higher in the population admitted to specialised psychiatric hospitals and the population with full coverage for depression compared to the population hospitalised in general hospitals (72–75% respectively vs. 55%). The mortality rate during the study period was twice as high for persons selected only on the basis of associated diagnoses reported by general hospitals than for other populations (7–9% vs. 2–4%).

The population identified by methods A + B + C was described (Table 2). A majority of the population had reimbursements for antidepressants or anxiolytics (74.3 and 61.3% respectively), and around 30% has reimbursements for neuroleptics or hypnotics. The total national health insurance spending for this population

**Table 1** Frequency counts and characteristics of individuals who received care for depression, derived from various data sources, in the SNIIRAM in 2012

Selection method	Source	Frequency count	Incremental count	Cumulative count	Cumulative prevalence (%)	Age (mean ± SD)	% males	% antidepressants	Mortality rate in 2012 (%)
A	Full coverage for depression	271,275	271,275	271,275	0.46	58 ± 17	28	72	1.6
B	Hospitalisation for depression								
	1 year								
	Psychiatric hospital (principal or associated diagnosis)	109,260	88,650	359,930	0.61	50 ± 17	36	75	1.7
	General hospital (principal diagnosis)	34,630	24,785	384,710	0.65	52 ± 17	37	60	3.4
	2 years								
	Psychiatric hospital (principal or associated diagnosis)	102,940	56,890	441,600	0.75	50 ± 18	36	70	2.1
	General hospital (principal diagnosis)	34,050	21,940	463,530	0.79	52 ± 18	35	55	3.6
C	Current antidepressant treatment + History of hospitalisation during the past 5 years								
	Psychiatric hospital (principal or associated diagnosis)	137,110	59,300	522,830	0.89	54 ± 17	32	-	1.7
	General hospital (principal diagnosis)	51,420	21,270	544,105	0.93	58 ± 18	30	-	2.6
D	Hospitalisation in a general hospital with an associated diagnosis of depression								
	1 year	213,010	164,660	708,760	1.21	61 ± 19	33	62	8.9
	2 years	192,720	127,270	836,030	1.42	60 ± 19	32	58	7.0
E	Current antidepressant treatment + History of hospitalisation in a general hospital with an associated diagnosis of depression during the past 5 years	266,890	133,270	969,300	1.65	62 ± 20	27	-	4.0

Incremental count = number of additional individuals selected by the data source; Cumulative count = number of individuals selected by the combination of data sources, to this point. % Antidepressants = proportion of individuals with at least three reimbursements for antidepressants during the study year. Data sources are classified from most reliable to least reliable



**Fig. 1** Numbers and proportions of individuals who received care for depression, derived from various data sources, in the SNIIRAM 2012. Method A: full coverage for depression as a specific long-term or costly condition. Method B: hospitalisation for depression in a general hospital (principal diagnosis of depression) or psychiatric hospital (principal or associated diagnosis of depression) during the last 2 years. Method C: taking antidepressants and a personal history of hospitalisation for depression in the last 5 years. Individuals selected exclusively on the basis of associated diagnoses in general hospital stays or personal history of hospitalisation for depression (3–5 years previously) were not included in this diagram. Numbers are rounded to the nearest 5

was 6283 million euros, with more than half this cost in hospital fees (3697 million euros). The proportion of individuals eligible for specific coverage for the less well-off (*Couverture Médicale Universelle Complémentaire*, used as a marker of socioeconomic status) was 14.4%, more than double the national average (6.6%). The calculation of CMUC rates was restricted to the population below 60 because individuals are eligible to different benefits after 60.

**Discussion**

**Main results**

Estimates of the prevalence of people using significant healthcare resources for depression in France vary according to the method used when analysing administrative databases, ranging from 0.46 to 1.65% from the most conservative to the broadest method, respectively, and from 0.61 to 0.93% when using intermediate methods. We propose an algorithm to study the population with significant healthcare use during the past 2 years that includes persons recently hospitalised for depression, or with a history of hospitalisation for depression and still taking antidepressants, or with full coverage for depression as a specific long-term or costly condition (method A, B and C). This algorithm yields a prevalence estimate of 0.93% or

544,105 individuals in the major health insurance scheme.

Population surveys provide much higher prevalence estimates of major depressive disorder during a 12-month period, ranging from 5 to 8%: 5.0% [22], or 6.0% [23] and 7.8% [24], depending on the survey. These estimates are much higher than our highest estimate of 1.65% over a 2-year period. These discrepancies are similar to those reported in a review and assessment of the performance of administrative data for depression surveillance in North America [9] that found that, while these data provide estimates with high specificity and positive predictive value, their sensitivity may remain low. This gap can be explained by differences in definition and methodology: whereas surveys are designed to identify individuals within the entire population presenting clinical criteria of major depressive disorder, the present study selected individuals using significant healthcare resources, representing only a fraction of the population with major depressive disorder. As psychotherapy is not currently reimbursed in France, patients with major depressive disorder may well receive appropriate ambulatory care with or without antidepressant treatment and without requiring hospitalisation or full health insurance coverage for depression. These patients are not included in our estimates.

**Table 2** Characteristics of the population identified by the selected algorithm (method A + B + C) in the SNIIRAM in 2012

Frequency count	544,105
Age	
Mean	54.6
IQR	43–64
Males (%)	32.1
Mortality rate in 2012 (%)	2.0
CMUC <sup>a</sup>	14.4
Comorbidities	
Cardiovascular diseases	14.9
Diabetes	11.4
Cancers	9.1
Neurological or degenerative diseases	9.0
Inflammatory diseases or HIV	4.0
End-stage kidney disease	0.2
Liver or pancreas diseases	3.9
Treated (%) <sup>b</sup>	
Antidepressants	74.3
Anxiolytics	61.3
Neuroleptics	25.9
Hypnotics	36.8
Any of the above	83.9
Amount reimbursed (millions of euros)	
Ambulatory care <sup>c</sup>	946
Pharmaceuticals	624
Hospital	3697
Sick leave	435
Disability allowance	581
Total costs	6283

<sup>a</sup> Full health coverage for individuals with an annual income below the poverty threshold. Analysis restricted to the population younger than 60

<sup>b</sup> At least three reimbursements

<sup>c</sup> Excluding pharmaceuticals

### Merits and drawbacks of the various methods

Methods based exclusively on attribution of full health insurance coverage for depression may fail to identify a large proportion of the target population, by mostly selecting individuals for whom the disease represents a significant clinical and financial burden. Even in this situation, full coverage is less likely to be requested when it has already been attributed for another chronic disease (such as a cardiovascular disease, diabetes or cancer) or if the person is eligible for specific coverage for the less well-off (*Couverture Médicale Universelle Complémentaire*) or has complementary private insurance to cover the additional economic burden. Full health insurance coverage for depression is therefore likely to be an insufficiently sensitive marker; this property has already been

demonstrated for other diseases such as cancer [25]. This finding was confirmed by our analysis, in which most people hospitalised for depression did not have full health insurance coverage. However, as full coverage for severe chronic depression is attributed by a health insurance physician at the request of the patient's doctor, and as it must be renewed on request every 5 years, this data source is probably highly specific [18].

Principal diagnoses from hospital stays are probably reliable markers of severe depression because they constituted the reason for admission of these individuals [19]. Even in non-psychiatric hospitals, they correspond to forms of depressions that are deemed to be sufficiently severe to justify admission. Furthermore, principal diagnoses play an important role in calculation of the cost of the hospital stay and are consequently subject to strict scrutiny by both hospital medical record summarisers and national health insurance. When used as markers of a history of severe depression, principal diagnoses allow the selection of a population with sufficiently severe depression to justify a hospital stay in the previous 5 years and who still need treatment with antidepressants.

The value of associated diagnoses during a psychiatric hospital stay may be similar to that of the principal diagnosis, as this diagnosis was reported by a psychiatrist. On the other hand, associated diagnoses from general hospital stays are likely to be less reliable. Although associated diagnoses should theoretically only provide information about illnesses that played a significant role during the hospital stay, they are also widely used to report history or concomitant diseases that, on their own, would not have justified hospitalisation [26]. Even individuals taking antidepressants and presenting an associated diagnosis of depression could actually be primarily treated for a severe organic disease. This possibility was confirmed by our data showing that this population had a higher mortality rate than the other groups. The inclusion of this population markedly increased prevalence estimates compared to the relative contribution of other more reliable data sources. While this population may suffer from milder forms of depression, individuals with associated diagnoses reported during general hospital stays should probably therefore not be included in this study, which was designed to select the population using significant healthcare resources for severe depression.

Extending the selected hospital stay timeframe from 1 to 2 years increased prevalence estimates by 0.14%. However, it is likely that depression which was sufficiently severe to require hospitalisation was not completely resolved 1 year later, therefore justifying the inclusion of hospital stays during the previous 2 years. Since this study was based on calendar years, actual

intervals between hospital discharge and study inclusion may vary from 1 to 24 rolling months. The proportion of individuals treated by antidepressants in our study was similar for persons hospitalised during the current year and persons hospitalised during the previous year, confirming that this population still suffered from depression or still required treatment to consolidate recovery.

Most individuals were selected by only one method, which highlights the value of combining several data sources, especially in the presence of sources with low sensitivity such as full health insurance coverage for depression. To our knowledge, this is the first study in France using these combined sources to study depression [6].

All ICD-10 codes for depression were used, including codes usually reserved for milder forms of depression (F32.0 and F32.9) to allow for inaccurate diagnosis coding based on the following rationale: severity of depression is determined by whether hospitalisation was deemed necessary more than by the ICD-10 code itself, especially since the objective of this study was to select persons with depression based on healthcare use criteria (such as hospital stays) and not purely on clinical criteria. The same codes have been used to define depression in an electronic health records-based study in the United Kingdom [27]. Individuals with bipolar disorder codes were not included in this study, but persons with bipolar disorder in a depressive phase in whom their specific condition had never been previously diagnosed may have been incorrectly included in all sources (full coverage, treatments or hospitalisations).

Antidepressant drugs are not specific of depression, as they may also be prescribed for chronic pain or anxiety [28, 29]. Furthermore, using drugs as a marker may select individuals with milder forms of depression not resulting in significant healthcare use. For both of these reasons, antidepressant treatment was only considered when it was combined with a history of hospitalisation for depression during the past 5 years, which is both specific of the disease and a better marker of significant healthcare use [29]. As an order of magnitude, 4.4 million people (8.7% of the population) in France were treated with an antidepressant in 2011 [30].

#### Limitations and future directions

As the databases used in this study only contain information on individuals with at least one reimbursement for medical care, this study was unable to detect the population with no contact with the healthcare system during the study period because of personal choices or problems of accessibility to health services. This drawback could constitute a major problem from an epidemiological or a need for care perspective. The proportion of the population without reimbursement in

2013 can be estimated to be around 5%, but varies according to gender, age and health status. It would therefore be difficult to estimate the impact of this population on the results of this study. On the other hand, insurance claim data are appropriate from the perspective of this study, as the primary objective was to select a population using significant healthcare resources related to a particular medical problem, and to assess the amount of resources allocated to this problem. Monitoring of the population with a significant healthcare use is particularly important from a public health, healthcare management and policy planning perspective. Finally, the population covered by the major health insurance scheme (86% of the French population) may not be completely representative of the entire French population. In particular, farmers and self-employed workers are not included in this insurance scheme.

The development of this algorithm is part of a larger work by CNAMTS to identify populations with various diseases in the SNIIRAM. Results from this research are analysed to produce a yearly report on spending per pathology and spending trends over time [21], or to study the national burden of more specific diseases [31]. This data was also used to compare regional profiles of various pathologies in France overseas departments [32], and could be employed more widely at a regional level to better inform local health policies. More specific longitudinal cohort studies are also being developed from this data to analyse pathway of care, outcomes, or comorbidities [33].

#### Conclusion

Combined analysis of administrative databases is an alternative and complementary approach to medical interview-based epidemiological surveys [9, 10], which constitute the gold standard. Medical interview-based surveys can provide estimates that include individuals who had no contact with the healthcare system, who did not seek care for their disease, or for whom the contact cannot be identified as a marker in an administrative database. However, these studies are expensive, must be performed on small random samples and take a long time to be implemented and to provide results. Working on medical records and administrative databases has a low added cost and can rapidly yield results on the recent situation and on secular trends, as the data have already been collected for other reasons. In the present study, we propose an algorithm to study the population with significant healthcare use during the past 2 years that will include persons recently hospitalised for depression, or with a history of hospitalisation for depression and still taking antidepressants, or with full health insurance coverage for depression as a specific long-term or costly condition.



These data provide relevant information to stakeholders for public health, healthcare management or policy planning [34] that must be analysed by taking their limitations into account.

This study shows that changes in the case selection methodology have major consequences on the frequency count and characteristics of the selected population, and consequently on the conclusions that can be drawn from the data. This finding is important for mental health researchers contemplating similar approaches: it emphasizes the importance (1) to define precisely the target population before the study and (2) to carefully consider how and why the administrative data is produced in the first place in order to produce relevant results.

#### Acknowledgments

This research was supported by CNAMTS and at the time of the study Antoine Filipovic-Pierucci was a public health medical resident of Paris hospitals (Assistance Publique-Hôpitaux de Paris). We thank Catherine Quantin for giving us early access to her report on depressive disorders in the SNIIRAM, Hervé Maisonneuve and Bruno Falissard for their advice for submitting this article, Chloé Le Cossec and Ayden Tadjahmady for their comments and suggestions, and Sébastien Rivière for the statistics in Table 2.

#### Funding

SS, JPF and AFC were employed by CNAMTS. AFP was employed by Paris Hospitals. There were no other sources of funding for this research.

#### Availability of data and materials

SNIIRAM individual-level data are not publicly available. Access conditions are stated in the following law: Arrêté du 19 juillet 2013 relatif à la mise en œuvre du Système national d'information interrégimes de l'assurance maladie.

#### Authors' contributions

AFP, SS, JPF and AFC made substantial contributions to conception and design of the study. AFP and SS performed the statistical analysis. AFP and AFC interpreted the results. AFP drafted the manuscript; SS, JPF and AFC revised it critically. All authors read and approved the final version of the manuscript.

#### Competing interests

The authors declare that they have no competing interests.

#### Consent for publication

Not applicable.

#### Ethics approval and consent to participate

All procedures performed in studies involving human participants were in accordance with the ethical standards of the national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Because of the retrospective nature of this study, formal consent is not required.

Received: 8 March 2016 Accepted: 8 December 2016

Published online: 03 January 2017

#### References

- WHO | WHO Mental Health Gap Action Programme (mhGAP). In: WHO. [http://www.who.int/mental\\_health/mhgap/en/](http://www.who.int/mental_health/mhgap/en/). Accessed 25 Jan 2016.
- Murray CJL, Vos T, Lozano R, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet Lond Engl*. 2012;380:2197–223.
- Wittchen HU, Jacobi F, Rehm J, et al. The size and burden of mental disorders and other disorders of the brain in Europe 2010. *Eur Neuropsychopharmacol J Eur Coll Neuropsychopharmacol*. 2011;21:655–79.
- Luppa M, Heinrich S, Angermeyer MC, König H-H, Riedel-Heller SG. Cost-of-illness studies of depression: a systematic review. *J Affect Disord*. 2007;98:29–43.
- Chevreul K, Prigent A, Bourmaud A, Leboyer M, Durand-Zaleski I. The cost of mental disorders in France. *Eur Neuropsychopharmacol J Eur Coll Neuropsychopharmacol*. 2013;23:879–86.
- Chan Chee C, Gourier-Fréry C, Guignard R, Beck F. The current state of mental health surveillance in France. *Santé Publique*. 2011;23 Suppl 6: S13–29.
- Drees, Actes du séminaire recherche, Organisation de l'offre de soins en psychiatrie et santé mentale, Document de travail, Série Études et recherche, n°129, avril, 2014.
- Coldefy M, Nestrigue C, Or Z. Étude de faisabilité sur la diversité des pratiques en psychiatrie. 2012.
- Fiest KM, Jette N, Quan H, St Germaine-Smith C, Metcalfe A, Patten SB, Beck CA. Systematic review and assessment of validated case definitions for depression in administrative data. *BMC Psychiatry*. 2014;14:289.
- Townsend L, Walkup JT, Crystal S, Olfson M. A systematic review of validated methods for identifying depression using administrative data. *Pharmacoepidemiol Drug Saf*. 2012;21 Suppl 1:163–73.
- Singh JA. Accuracy of Veterans Affairs databases for diagnoses of chronic diseases. *Prev Chronic Dis*. 2009;6:A126.
- Alaghebandan R, Macdonald D, Barrett B, Collins K, Chen Y. Using administrative databases in the surveillance of depressive disorders—case definitions. *Popul Health Manag*. 2012;15:372–80.
- Noyes K, Liu H, Lyness JM, Friedman B. Medicare beneficiaries with depression: comparing diagnoses in claims data with the results of screening. *Psychiatr Serv Wash DC*. 2011;62:1159–66.
- CNAMTS. Présentation de la cartographie des pathologies et des dépenses. 2015.
- Chevreul K, Berg Brigham K, Durand-Zaleski I, Hernandez-Quevedo C. France HIT (2015); 2016. <http://www.euro.who.int/en/about-us/partners/observatory/publications/health-system-reviews-hits/full-list-of-country-hits/france-hit-2015>. Accessed 25 Jan 2016.
- Homepage | CNIL. <https://www.cnil.fr/en/home>. Accessed 1 Dec 2016.
- Tuppin P, de Roquefeuil L, Weill A, Ricordeau P, Merlière Y. French national health insurance information system and the permanent beneficiaries sample. *Rev Dépidémiologie Santé Publique*. 2010;58:286–90.
- Païta M, Weill A. Les personnes en affection de longue durée au 31 décembre 2007. *Points Repère*. 2008;20:201–8.
- ATIH. Guide méthodologique de production des informations relatives à l'activité médicale et à sa facturation en médecine, chirurgie, obstétrique et odontologie. Paris: Bulletin officiel; 2012.
- Micallef L, Rodgers P. eulerAPE: Drawing Area-Proportional 3-Venn Diagrams Using Ellipses. *PLoS ONE*. 2014;9:e101717.
- CNAMTS. Rapport au ministre chargé de la sécurité sociale et au parlement sur l'évolution des charges et produits de l'assurance maladie au titre de 2014 (loi du 13 août 2004). *Assurance Maladie*. 2013.
- Sapinho D, Beck F, Chee CC. Prévalence de l'épisode dépressif majeur et co-morbidités. 2009. p. 35–58.
- Lépine J-P, Gasquet I, Kovess V, Arbabzadeh-Bouchez S, Nègre-Pagès L, Nachbaur G, Gaudin A-F. Prévalence et comorbidité des troubles psychiatriques dans la population générale française : résultats de l'étude épidémiologique ESEMeD/MHEDEA 2000/(ESEMeD). *L'Encéphale*. 2005;31: 182–94.
- Lambooy B, Léon C, Guilbert P. Troubles dépressifs et recours aux soins dans la population française à partir des données du Baromètre Santé 2005. *Rev Dépidémiologie Santé Publique*. 2007;55:222–7.
- Colonna M, Chagnoux E, Remontet L, Mitton N, Belot A, Bossard N, Grosclaude P, Uhry Z. Estimations de l'incidence départementale des cancers en France métropolitaine 2008–2010. Étude à partir des données des registres des cancers du réseau Francim et des bases de données médico-administratives. Sainte-Maurice: Institut de veille sanitaire; 2015.
- Quantin C, Roussot A, Besson J, Cottenet J, Bismuth M-J, Cousin F, Soudry-Faure A, Mariet A-S, Collin C. Étude des algorithmes de repérage de la dépression dans le SNIIRAM par le réseau REDSIAM. *Rev Fr Aff Soc*. 2016; 201–25.

27. Chang C-K, Hayes RD, Broadbent M, Fernandes AC, Lee W, Hotopf M, Stewart R. All-cause mortality among people with serious mental illness (SMI), substance use disorders, and depressive disorders in southeast London: a cohort study. *BMC Psychiatry*. 2010;10:77.
28. Sramek JJ, Zarotsky V, Cutler NR. Generalised anxiety disorder: treatment options. *Drugs*. 2002;62:1635–48.
29. Quantin C, Goldberg M. Journées EMOIS Nancy 2015 : Session E2–Groupe de travail “Troubles mentaux et du Comportement (GT-TMC)” REDSIAM. [https://www.canal-u.tv/video/canal\\_u\\_medecine/journees\\_emois\\_nancy\\_2015\\_session\\_e2\\_groupe\\_de\\_travail\\_troubles\\_mentaux\\_et\\_du\\_comportement\\_gt\\_tmc\\_redsiam.17874](https://www.canal-u.tv/video/canal_u_medecine/journees_emois_nancy_2015_session_e2_groupe_de_travail_troubles_mentaux_et_du_comportement_gt_tmc_redsiam.17874). Accessed 1 Feb 2016.
30. Fagot J-P, Cuerq A, Samson S, Fagot-Campagna A. Cohort of one million patients initiating antidepressant treatment in France: 12-month follow-up. *Int J Clin Pract*. 2016;70(9):744–51.
31. Tuppin P, Rivière S, Rigault A, et al. Prevalence and economic burden of cardiovascular diseases in France in 2013 according to the national health insurance scheme database. *Arch Cardiovasc Dis*. 2016;109:399–411.
32. Filipovic-Pierucci A, Rigault A, Fagot-Campagna A, Tuppin P. Health status of populations living in French overseas territories in 2012, compared with metropolitan France: An analysis of the national health insurance database. *Rev Epidemiol Sante Publique*. 2016;64(3):175-83.
33. Fagot J-P, Samson S, Merlière J, Van Der Erf S, Gabach P, Fagot-Campagna A. Prévalences des pathologies somatiques associées à des troubles psychiatriques ou des traitements par psychotropes dans le SNIIRAM et le PMSI. Bordeaux: Congrès Adelphe-SFSP; 2013.
34. Koné Péfoyo AJ, Rivard M, Laurier C. Importance de la surveillance en santé publique et utilité des données administratives. *Rev Dépidémiologie Santé Publique*. 2009;57:99–111.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

