Hospital databases for the identification of adverse drug reactions: A 2-year multicentre study in 9 French general hospitals

Marie-Noëlle Osmont, Adeline Degremont, Hélène Jantzem, Youna Audouard-Marzin, Sébastien Lalanne, Dominique Carlhant-Kowalski, Eric Bellissant, Emmanuel Oger, Elisabeth Polard

Osmont Marie-Noëlle (Orcid ID: 0000-0001-7472-8780)

Hospital databases for the identification of adverse drug reactions: a two-year multicentre study in nine French general hospitals

Running title: Adverse drug reactions and hospital databases

Marie-Noëlle OSMONT[1], Adeline DEGREMONT[1,2], Hélène JANTZEM[3], Youna AUDOUARD-MARZIN[3], Sébastien LALANNE[1], Dominique CARLHANT-KOWALSKI[3], Eric BELLISSANT[1], Emmanuel OGER[1,2], Elisabeth POLARD[1,2]

[1]Pharmacovigilance and pharmacoepidemiology centre, Pharmacology Department, CHU Rennes, France

[2]Univ Rennes, EA 7449 REPERES 'Pharmacoepidemiology and Health Services Research', F 35000 Rennes, France.

[3]Pharmacovigilance centre, CHRU Brest, France

Corresponding author: Marie-Noëlle OSMONT marie-noelle.osmont@chu-rennes.fr

ABSTRACT

Aims

In order to estimate the actual number of adverse drug reactions (ADRs), we used the French medical administrative database ("PMSI") in addition to ADRs spontaneously reported in the French Pharmacovigilance Database (FPVDB).

Methods

Capture-recapture method was applied to these two sources ("PMSI" and FPVDB), checking their independence via a third data source. The study ran from July 1st, 2014 to June 30th, 2016 in nine French general hospitals. From "PMSI", all discharge summaries including a selection of ICD10 codes (10th International Classification of Diseases) related to ADRs were analyzed. This selection was based on the results of a previous study. All ADRs corresponding to these codes, spontaneously reported in the FPVDB, were included.

Results

In "PMSI", 56.9% of hospital stays were related to an ADR (628 out of 1,104). In the FPVDB, we retained 115 cases. A total of 43 ADRs were common to the two databases. In both sources, the most frequently reported ADRs were cutaneous (33.1% and 19.1%) and renal (25.2% and 11.6%). The most frequently suspected drugs were anti-infectives in "PMSI" (31.1%) and antineoplastic drugs in the FPVDB (30.4%). Using the capture-recapture method, the estimated number of ADRs was 1,657 [95 % CI: 1,273 to 2,040].

Conclusions

The use of the "PMSI" could constitute an additional tool for the estimation of the actual number of ADRs in French hospitals. A model involving a third data source enabled the independence of the two sources ("PMSI" and FPVDB) to be checked before applying the capture-recapture method.

Keywords: databases; ICD10; adverse drug reaction reporting systems, pharmacovigilance

.

What is already known about this subject:

- There is evidence of under-reporting of ADRs to spontaneous reporting systems, including serious ADRs.
- Capture-recapture method is an option for estimating the total number of ADRs from two or more sources (including spontaneous reporting database).

What this study adds:

- Our study was multicentre, involving several general hospitals and a tertiary university hospital. Only single-centre studies using the capture-recapture method to estimate a number of ADRs occurring in French university hospitals have previously been described in the literature.
- We identified serious ADRs from two sources (a hospital database, using carefully chosen ICD10 codes, and a spontaneous reporting database), their independence was checked using a model involving a third data source (Clinical Data Warehouse). To our knowledge, the concomitant use of these three sources for the capture-recapture method to detect ADRs has not yet been described in the literature.
- The use of the PMSI over the study period enabled the number of ADRs identified to increase six-fold (compared to spontaneous reporting alone).

.

INTRODUCTION

In France, adverse events related to care provision are an important cause of morbidity. Out of 8,754 hospitalized patients, 450 adverse events related to the care provided were observed in the 2005 ENEIS survey, half occurring during hospitalization, and one third being preventable. It is known that some of these events were adverse drug reactions (ADRs) [1,2]. A prospective study named EMIR was conducted in France on a representative, randomly-selected sample of medical wards in public hospitals from December 2006 to June 2007 [3]. All patients admitted over a 2-week period were included. Among the 2,692 admissions, 97 were related to an ADR (incidence 3.6%, 95% confidence interval, CI [2.8–4.4]). The estimated annual number of ADR-related hospitalizations in France was 143 915 (95% CI [112 063–175 766]). Overall, according to various studies, estimations of the incidence of ADRs in hospitals vary from 1.7% to 3.6% [3–5].

Spontaneous reporting is the main source of ADR detection in European countries, especially in France. Unfortunately, it is not sufficiently exhaustive to detect all ADRs, because of under-reporting. A systematic review provided evidence of significant and widespread under-reporting of ADRs to spontaneous reporting systems, including serious or severe ADRs [6]. A median under-reporting rate was calculated across the 37 studies and within study subcategories for different methods or settings. The rate was 94% (interquartile range 82-98%). Under-reporting reduces sensitivity because it underestimates the frequency and thereby the impact of a given ADR. In addition, it makes the system more vulnerable to selective reporting, which can introduce considerable bias [7].

The rise of IT networks (Information Technology) and the power of big data analysis should enable an improvement in the relevance of detected signals. Nevertheless, these tools should be considered as additional sources of information rather than as a substitute for spontaneous reporting, which remains the most efficient reference tool [8]. The spontaneous reporting system for ADRs is the cornerstone of the post-marketing monitoring of drug safety and is crucial for rapid signal detection [7]. However this system alone does not make it possible to estimate the actual number of adverse effects.

Estimating the actual number of ADRs can be useful for the authorities (Agence Nationale de Sécurité du Médicament et des produits de santé (ANSM), Agences Régionales de Santé (ARS), etc.) for example, to quantify the additional costs induced by ADRs in a geographical area (hospitalizations, surgery, medical treatments, sequelae, deaths) or to define health priorities and policies.

Queries in the French Medical Information System Program (Programme de Médicalisation des Systèmes d'Information (PMSI)) could be used for ADR detection by selecting codes of interest [9]. Biomedical data collected during the clinical care process can also be reused and analyzed to improve ADR detection. Technologies, such as Clinical Data Warehouses (CDW), have these abilities, but they are not currently deployed in all hospitals. A CDW is available at Rennes University Hospital, thus making the exploitation of clinical data possible. Studies have been conducted at Rennes University Hospital to evaluate the performance of a data-gathering method using a CDW for cases of drug-induced anaphylactic shocks [10,11].

The application of the capture-recapture method to hospital settings using the PMSI and pharmacovigilance databases could contribute to an estimate of the actual impact of ADRs in hospitals. This method is used to provide population estimates from two or more incomplete sources of information. It allows the extent of incomplete ADR case ascertainment to be estimated. In the literature, a few publications have described French studies giving an estimation of the total number of ADRs from two sources (PMSI and spontaneous reporting) using the capture-recapture method [12-14]. The studies that did so showed the lack of exhaustiveness of ADR reporting whatever the sources of data, but the use of the PMSI database improved the detection of ADRs. They underlined the interest of merging data from different databases to fully identify the real impact of ADRs in hospitals.

The aim of our regional study was to estimate the number of ADRs leading to or occurring during hospitalization in nine public general hospitals in western France. Two Regional Pharmacovigilance Centers (RPVC) were involved in this regional study (the only two in the region - Brest and Rennes) which ran from July 1st, 2014 to June 30th, 2016, and used the capture–recapture method with two sources (the PMSI and the French Pharmacovigilance Database (FPVDB) including spontaneously reported ADRs). Their independence was checked by involving a third data source (CDW).

.

METHODS

- **Selection of hospitals**

Sixteen public general hospitals in Brittany, western France, were eligible to participate in the study and were invited to take part. Nine hospitals (56%) were selected on a voluntary basis. Four of them are located in Ille-et-Vilaine, three in Morbihan, one in Côtes d'Armor and one in Finistère. The number of beds in these hospitals ranged from 0 to 200 for 3 hospitals, 200 to 500 for 4 hospitals and more than 500 for 2 hospitals. We limited the study to patients discharged from hospital from July 1st, 2014 to June 30th, 2016.

The Commission Nationale de l'Informatique et des Libertés approved this study in 2016 (decisions no. 1860748, 1865978, 1873496, 1873903, 1876346, 1894267).

- **Data sources**

In the nine hospitals, we used two data sources. These two sources are the following: i/ Diagnosis-related Groups (DRGs) from the PMSI and ii/ spontaneous reports collected by the RPVC of Brittany (Brest and Rennes) and recorded in the French PharmacoVigilance Database (FPVDB).

The French PMSI database generates DRGs that contain administrative (name, gender, birthdate and dates of hospital admission/discharge) and medical (diagnoses and medical or surgical procedures, coded following the International Classification of Diseases 10th revision (ICD10)) data, reported as a standardized medical outcome summary for each hospital stay.

Cases of spontaneously reported ADRs which occurred or required care in the nine hospitals were retrieved from the FPVDB. Since 1985, the FPVDB has been gathering ADR cases occurring in France and reported to the 31 RPVC by health professionals and, since 2011, by patients.

The RPVC are required by law to assess drug safety in hospitals. Prescribers have to report 'serious' ADRs (i.e. resulting in death, requiring patient hospitalization or prolongation of existing hospitalization, resulting in permanent disability or incapacity, or life-threatening) and/or 'unexpected' ADRs (not labelled in the Summary Product Characteristics) to their RPVC.

Each RPVC was in charge of cases from hospitals located in its geographical area: four hospitals for the Brest RPVC and the other five hospitals for the Rennes RPVC.

- **Case definition**

We studied ADRs corresponding to selected ICD10 codes with a date of occurrence or diagnosis from July 1st, 2014 to June 30th, 2016 and cared for in a medical ward in one of the nine hospitals in Brittany. An ADR is a harmful effect suspected to be caused by a drug. The term is normally restricted to late-stage analyses when the association between a medicine and an adverse effect is no longer considered 'non-measurable' or 'uncertain' [15].

- **Selection of cases and identification of duplicates**

- **The PMSI database**

In this study, we chose to focus on certain specific ADRs that concerned ICD10 codes related to dermatological effects (L27.0, L51.1, L51.2), anaphylactic shock (T88.2, T88.6), nephropathy (N14.1, N14.2, N17), hepatic impairment (K71), neuropathy (G62.0) and interstitial lung disease (J70.2, J70.3, J70.4), and either the association of ICD10 code K71 with codes Y40 to Y59 / T88.7, or the association of code N17 with the same codes (appendix table 1). This choice was based on the results of a previous study performed in 2009 in Rennes evaluating queries on various ICD10 codes to the PMSI database to identify serious ADRs [16] and on the existence of another study [17,18] (see Appendix 1).

In each hospital, the department of medical information provided local pharmacovigilance correspondents with lists of medical records identified through the query. The documentation of the cases was carried out by the local pharmacovigilance correspondents. All hospital stays involving a hospitalization summary that included at least one of the ICD10 codes selected were investigated to check whether the adverse event extracted was a suspected

ADR. The correspondent retained the files in which medical records suggested a possible link between an event and a drug as the cause of an ADR. The ADR cases were transferred anonymously (initials, date of birth of the patients) to the RPVC in Brest or Rennes, depending on the geographical area.

Cases were analyzed by the RPVC. Drug causality was assessed using the official French method [19]. This method is algorithmic, and based on the evaluation of eight criteria divided into three groups: chronology, semiology and bibliographic data. One particular feature of the French method of causality assessment is the separate rating of an intrinsic causality score and an extrinsic bibliographic score. A second feature is the independent assessment of intrinsic causality for a given drug in a given case [20].

All the cases were analyzed by a pharmacovigilance expert and then reviewed by another pharmacovigilance expert before recording in the FPVDB. Any disagreement between experts required a consensus session before reaching the final decision. The recording of cases, whether from spontaneous reporting or the PMSI, is an obligation for the RPVC as soon as they become aware of them.

For each false-positive case in the query, the reason for exclusion was specified by the local pharmacovigilance correspondent. After a double check by the RPVC, some cases initially included could be excluded.

- **The FPVDB**

From the FPVDB, we identified all ADRs of interest spontaneously reported by health professionals from the nine hospitals from July 1st, 2014 to June 30th, 2016 (query by SOC: System Organ Class, using MedDRA (Medical Dictionary for Regulatory Activities) classification).

For each spontaneous report, data concerning patient, drug exposure and reaction were available. Drug causality had already been assessed for case registration using the official French method [19].

- **Capture-recapture method**

The capture–recapture method provides an estimation of the population affected, and is particularly useful when the investigator clearly has incomplete data available from two or more sources [21]. In this study, we used 2-source capture-recapture with the following two incomplete data sources: the FPVDB and the PMSI database. These sources were used to count the number of ADRs identified in the FPVDB (capture), the number of ADRs identified in the PMSI database query (recapture) and the number of ADRs common to both sources, identified by matching on patient initials and date of birth, date and characteristics of the ADR and the drugs involved. The detailed method is available in Appendix 2 [21-24].

Thanks to pharmacovigilance expertise, each case was diagnosed accurately and was within the time-space unit under study; the "matching" of cases that appears in different sources was carried out appropriately.

The following conditions for applying the capture-recapture method were checked: homogeneity of catchability, case matching, true cases, cases identified as having occurred in an identical period and area and finally, independent sources.

The capture-recapture method based on two sources requires the assumption of independence, which is not straightforward. In order to further evaluate the independence between our two sources, a second analysis integrating a third data source was conducted at Rennes University Hospital. The third source was the eHOP Clinical Data Warehouse which is available in this hospital. For this analysis, only cases of drug-induced anaphylactic shocks were included.

For log-linear modeling, we used the PROC GENMOD procedure available on SAS 9.4 software (SAS Institute, Cary, NC., USA).

RESULTS

- **Numbers of ADRs**

From the FPVDB, we found 115 cases spontaneously reported by health professionals from the nine hospitals during the study period and corresponding to the ICD10 codes. The health professionals who spontaneously reported cases were mainly physicians (82%) and to a lesser extent pharmacists (18%).

From the PMSI database, during the study period, among 1,104 hospital stays detected by one of the selected ICD10 codes, we identified and validated 628 hospital stays related to an ADR (56.9%). Thus 476 were excluded: 67.2% because we did not find any mention of an ADR, 10.5% because the name of the drug was not specified, 7.6% because the ADR had occurred before the study period and was not the reason for the stay (historical event), 6.6% because there was insufficient information, 4.6% because the same ADR was identified several times, 3.3% because there was no discharge report and 0.2% because it concerned a clinical trial.

A total of 43 ADRs were common to the two databases, registered in both the PMSI database and FPVDB.

- **Patient characteristics** (Appendix figure 1)

Patients from the PMSI database (48.5% male) were aged from 7 months to 101 years (median age, 72 years).Patients from the FPVDB (51.4% male) were aged from 8 to 93 years (median age, 66.5 years). Cases in common corresponded to patients (46.5% male) aged from 17 to 93 years (median age, 69 years).

- **Characteristics of the ADRs**

Only the ADRs reaching a proportion of ≥ 5% (classified according to System Organ Classes (SOC)) are shown in Figure 1.

Among the ADRs selected, the two most frequently reported ADRs were “Skin and subcutaneous tissue disorders” and “Renal and urinary disorders” in both sources but in reverse order, i.e. the most frequently reported ADRs were first “Renal and urinary disorders” for the PMSI database whereas it was “Skin and subcutaneous tissue disorders” for the FPVDB. Among the selected ADRs investigated, the proportions of hepatobiliary and respiratory disorders in the two sources ranged from 5 to 10%. In the PMSI, immune system disorders were also detected in the same proportion. Only neurological disorders were found in a proportion of less than 5%. It should also be noted that non-investigated ADRs (gastrointestinal disorders, general disorders and hematological disorders) were reported in association with ADRs that were specifically sought.

- **Suspected drug classes**

Figure 2 shows the distribution of drugs suspected according to the Anatomical Therapeutic Chemical (ATC) classification (first level). The most frequently suspected drugs were anti-infectives in the PMSI database and antineoplastic drugs in the FPVDB.

- **Estimation of the total number of ADRs using the capture-recapture method**

The conditions for applying the capture-recapture method were met. Results for the capture-recapture model are presented in Table 2. The estimated actual number of ADRs was 1,657 [95% confidence interval 1,273 to 2,040].The inclusion in the statistical analysis of a third data source (CDW) made it possible to assess the independence of the two sources. This evaluation was based on the use of three sources for the detection of drug-induced anaphylactic shock: spontaneous reporting, PMSI and a CDW known as eHOP (appendix figure 2). It can be recalled that different models of interaction between sources were evaluated. The results are presented in Table 1 and explanations are available in appendix 3.

The performances of the different models suggested that the PMSI and spontaneous reporting (FPVDB) were independent, and the interaction between these two sources appeared minimal. We therefore thought that the condition “sources are independent” was fulfilled.

DISCUSSION

- **Contributions of the PMSI**

In this study, the use of the PMSI for the detection of ADRs, which had already been done in university hospitals in Brittany, was extended on regional scale to general hospitals on a voluntary basis. As is already the case with

university hospitals, using the PMSI in general hospitals appeared as complementary to spontaneous reporting, allowing the under-reporting to be compensated for with the ADRs under study. The use of the PMSI over the study period enabled the number of ADRs identified to be multiplied by 6 (115 cases identified from spontaneous reporting and 585 new cases from the PMSI database, leading to a total of 700 cases).

Using the two sources proved complementary in terms of cases identified (very little overlap between sources). Furthermore, although the main types of ADR (classified by SOC, Figure 1), the drugs involved (per ATC classes, Figure 2) and the population types (Appendix figure 1) were fairly similar for cases from the PMSI and from spontaneous reporting, they nevertheless do not completely superimpose. Thus, the use of the two sources enabled complementary data to be detected.

The ADRs detected in the study via the PMSI were for the majority serious, which was consistent, given the fact that they carried the ICD10 codes selected for the queries, targeting ADRs that had led to hospitalisation or had occurred during hospitalisation. The ADRs mainly concerned expected iatrogenic effects, but certain cases of interest were also identified (for instance a suspicion of anticoagulant-related nephropathy – acute renal failure by intra-tubular precipitation of hematic cylinders - in a setting of dabigatran overdose).

The PMSI enables a more exhaustive collection of ADRs, thus potentially enabling cases that have not been spontaneously reported to be highlighted (but with a large volume of patient files to be analysed). It also enables estimation of the actual number of ADRs in hospitals, as in the present study.

- **Estimation of the total numbers of ADRs using the capture-recapture method**

The numbers of ADRs in the two data collection sources (spontaneous reporting and PMSI) enabled the use of the capture-recapture method, after having checked the independence of these sources. The estimation of the total number of ADRs under study was 1,657 [1,273 ; 2,040] in the general hospitals involved in the project over the study period. The percentage of spontaneously reported ADRs was about 7% (115 out of 1,657). The percentage of ADRs detected via the PMSI was about 38% (628 out of 1,657).

In the literature, a few publications have described French studies giving an estimation of the total number of ADRs from the two sources (PMSI and spontaneous reporting) using the capture-recapture method [12-14]. They were all single-centre studies (conducted at Toulouse university hospital). In the study by Lugardon *et al.* the use of the PMSI led to the detection of 274 cases of adverse effects versus 151 via spontaneous reporting. A total of 52 cases were found via the two collection modes. According to the capture-recapture method, the total estimated number of serious adverse reactions was 796 [638 ; 954]. This corresponded to an estimation of 81% under-reporting. In the study by Durrieu *et al.* on the detection of adverse effects in paediatrics, 60 cases were identified via the PMSI and 172 via spontaneous reporting. There were 28 cases common to the two sources. The capture-recapture method enabled the number of adverse effects to be estimated at 717 [513 ; 921]. This corresponded to an estimation of 76% under-reporting in this paediatric population.

The rate of under-reporting estimated in our study (93,1%) seems to concur with figures reported in the literature [6], notwithstanding the two studies previously cited.

- **Strengths of the study**

In order to assess independence between the two sources in our study, an analysis integrating a third data source was carried out. It corresponded to a query filed to the eHOP database, available at Rennes university hospital. This analysis allowed the independence of the sources to be checked before applying the capture-recapture method.

Our study was regional and multicentre, involving several general hospitals. Other authors have described single-centre studies providing an estimation of the total number of ADRs based on the two sources via the capture-recapture method [12-14]. Among them, Durrieu *et al.* suggested the usefulness of implementing a multicentre study.

The population concerned in the present study corresponded to all the patients for whom the date of hospitalisation discharge was between July 2014 and June 2016 in the nine health units taking part in the project. This population

was spread across the different geographical zones in the region. The type of unit in which this population was cared for was variable: it was important to have diversity in the type of unit in the study.

ICD10 codes for the different ADRs, for which detection had been tested in university hospitals across Brittany, were used. They enabled a wider detection of ADRs of different types, most of them serious and involving a variety of drugs.

Cases detected by the PMSI were made known via personalized and repeated feedback to physicians. This feedback was achieved through the local pharmacovigilance correspondents. The hypothesis was that this information, which is both context-specific and personalized, would make physicians more aware and encourage them to spontaneously report iatrogenic events.

The use of the capture-recapture method with three sources demonstrated that neither the eHOP nor the PMSI sources were independent. Thus, in the hospitals with both data sources, their combined use is not relevant. These tools (PMSI and eHOP) nevertheless need to be taken into account, but rather as sources of additional information. Spontaneous reporting remains the reference tool to detect signals [8].

- **Limitations of the study**

Selection of the ICD10 codes

For the selected ICD10 codes, the percentage of included cases related to ADRs was 56.9%. Concerning the particular ICD10 code N17 associated with ICD10 codes from Y40 to Y59 / T88.7 (used to look for drug-induced acute renal failure), the percentage of included cases dropped to 35.9%. This led to a re-examination of patient files and finally to the exclusion of these cases, for the most part (64.1%), because they were not drug-induced. Fortunately, for most ICD10 codes, the percentage of included cases was over 60%. Although the PMSI enables a more exhaustive collection of ADRs, it involves a time-consuming analysis of the hospital discharge reports.

Catchability of the cases of ADR

ADRs may not have the same probability of being identified by each source. ADRs related to a new drug or unexpected ADRs are more likely to be reported to a pharmacovigilance structure [12]. One limitation linked to the PMSI is its objective, which is above all medico-economic. Health monitoring, and pharmacovigilance in particular, are not intrinsically part of the PMSI objectives. The coding of diagnoses, although it follows an official methodological guide, can be channelled towards maximising the value of hospital stays. As a result, there may be heterogeneity of coding between the different health units.

Furthermore, compared to other terminologies, such as the Systematized NOmenclature of MEDicine – Clinical Terms (SNOMED-CT), the ICD10 has a certain degree of granularity of concepts and a limited cover, which can have an impact on the precision and the exhaustiveness of the coded data.

Finally, some ADRs occurring during hospitalization (and corresponding to the selected ICD10 codes) were not included in our study if observed by the patient or a general practitioner after discharge from hospital. Nevertheless most ADRs occurring during hospitalization are spontaneously reported by hospital practitioners.

Causality assessment

Although the causality definition is based on the French method [19], it is worth noting that there could be a bias linked to the local pharmacovigilance correspondent's expertise. Indeed, only the cases retained after the correspondent's analysis were sent to the RPVC. In rare instances, certain cases initially included by the correspondent had to be excluded after a second RPVC analysis. However the cases excluded by the correspondent could not be subjected to RPVC analyses for reasons of regulatory access to the files. Certain cases excluded by the local correspondent may have been included by the RPVC as validated ADRs.

Missing data

Some cases were excluded because of insufficient data. A drug-related hypothesis had been raised, but the available data did not enable the adverse reactions to be linked to a specific drug, but only to a drug class. These supposed ADRs were not taken into consideration in the study, which may have had an impact on the estimate of ADRs.

Conclusion

Given the results from this study, the use of the PMSI could constitute an additional tool for the detection of ADRs on a regional scale in general hospitals, thus supplementing spontaneous reporting. The detection of ADRs via the PMSI, or via hospital databases more generally, and the estimation of the total numbers of ADRs using the capture-recapture method, could prove useful for the hospitals concerned and for the different health institutions on regional or national scale. Indeed, these quantitative and qualitative databases could constitute an indicator for the follow-up of drug-related iatrogenic events in a given geographical area. They concern ambulatory (an ADR leading to hospitalisation) and hospital (an ADR occurring during hospitalisation) iatrogenic events. A follow-up over time of this indicator could enable a retrospective analysis of the evolution of the numbers of ADRs or of a type of ADR in particular. A significant variation of the indicator could provide a more in-depth analysis of the causes of ADR occurrences. A follow-up of this type of indicator could be chosen to guide and prioritise strategies, health policies and specific communication strategies on iatrogenic events identified via the chosen codes.

ACKNOWLEDGMENTS

Funding for this study was provided by a grant from the French National Agency for Medicines and Health Products Safety (ANSM) (grant number AAP-2014-014).

We would like to thank pharmacovigilance correspondents: Réjane Bessard, Eric Branger, Marie Desoil, Philippe Jaccard, Ronan Largeau, Nicole Le Gall, Béatrice Marie DitDinard, Sabrina Normand and Christine Souhaité.

We would like to thank correspondents of the medical information departments: Dominique Balzac, Emeline Baron, Paolo Bercelli, Matthieu Dauny, Jeanne-Marie Germain, Isabelle Labarre and Karine Vannier-Correia.

We would like to thank Guillaume Bouzillé (eHOP CDW).

Contributions of Authors statement:

Conceptualization: Marie-Noëlle Osmont, Dominique Carlhant-Kowalski, Emmanuel Oger, Elisabeth Polard. Supervision: Marie-Noëlle Osmont. Methodology: Adeline Degremont, Emmanuel Oger. Extraction of data from the administrative database: Marie-Noëlle Osmont, Hélène Jantzem, YounaAudouard-Marzin. Final analysis: Marie-Noëlle Osmont,SébastienLalanne, Adeline Degremont. Writing – original draft preparation: Marie-Noëlle Osmont. Writing – review and editing: Adeline Degremont, Eric Bellissant, Emmanuel Oger, Elisabeth Polard.

Conflict of interest: None

Data availability statement: The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

.

REFERENCES

1. Philippe Michel, Jean-Luc Quenon, Ahmed Djihoud, Sophie Tricaud-Vialle, Anne-Marie De Sarasqueta et Sandrine Domecq - CCECQA - Avec la collaboration de Brigitte Haury et Chantal CASES - Direction de la recherche, des études, de l'évaluation et des statistiques (DREES). Les événements indésirables graves liés aux soins observés dans les établissements de santé : premiers résultats d'une étude nationale - Ministère des Solidarités et de la Santé [Internet]. [cited 2019 Jun 24]. Available from: https://drees.solidarites-sante.gouv.fr/etudes-et-statistiques/publications/etudes-et-resultats/article/les-evenements-indesirables-graves-lies-aux-soins-observes-dans-les
2. Michel P, Minodier C, Lathelize M, Moty-Monnereau C, Domecq S, Chaleix M, et al. Les événements indésirables graves associés aux soins observés dans les établissements de santé - Ministère des Solidarités et de la Santé [Internet]. [cited 2019 May 13]. Available from: https://drees.solidarites-sante.gouv.fr/etudes-et-statistiques/publications/les-dossiers-de-la-drees/dossiers-solidarite-et-sante/article/les-evenements-indesirables-graves-associes-aux-soins-observes-dans-les
3. Bénard-Laribière A, Miremont-Salamé G, Pérault-Pochat M-C, Noize P, Haramburu F. Incidence of hospital admissions due to adverse drug reactions in France: the EMIR study. Fundam Clin Pharmacol. 2015;29(1):106–11.
4. Carrasco-Garrido P, de Andrés LA, Barrera VH, de Miguel GÁ, Jiménez-García R. Trends of adverse drug reactions related-hospitalizations in Spain (2001-2006). BMC Health Serv Res. 2010;10(1):287.
5. de Almeida SM, Romualdo A, de Abreu Ferraresi A, Zelezoglo GR, Marra AR, Edmond MB. Use of a trigger tool to detect adverse drug reactions in an emergency department. BMC Pharmacol Toxicol 2017; 18(1):71.
6. Hazell L, Shakir SAW. Under-Reporting of Adverse Drug Reactions. Drug Saf. 2006; 29(5):385–96.
7. Biagi C, Montanaro N, Buccellato E, Roberto G, Vaccheri A, Motola D. Underreporting in pharmacovigilance: an intervention for Italian GPs (Emilia-Romagna region). Eur J Clin Pharmacol. 2013;69(2):237–44.
8. Vial T. French pharmacovigilance: Missions, organization and perspectives. Therapie. 2016;71(2):143–50.
9. Hohl CM, Karpov A, Reddekopp L, Doyle-Waters M, Stausberg J. ICD-10 codes used to identify adverse drug events in administrative data: a systematic review. J Am Med Inform Assoc JAMIA. 2014;21(3):547–57.
10. Osmont M-N, Campillo-Gimenez B, Metayer L, Jantzem H, Rochefort-Morel C, Cuggia M, et al. Peranesthesic Anaphylactic Shocks: Contribution of a Clinical Data Warehouse. Therapie. 2015 Oct 16
11. Bouzillé G, Osmont M-N, Triquet L, Grabar N, Rochefort-Morel C, Chazard E, et al. Drug safety and big clinical data: Detection of drug-induced anaphylactic shock events. J Eval Clin Pract. 2018;24(3):536–44.
12. Lugardon S, Desboeuf K, Fernet P, Montastruc J-L, Lapeyre-Mestre M. Using a capture-recapture method to assess the frequency of adverse drug reactions in a French university hospital. Br J Clin Pharmacol. 2006; 62(2):225–31.
13. Durrieu G, Batz A, Rousseau V, Bondon-Guitton E, Petiot D, Montastruc JL. Use of administrative hospital database to identify adverse drug reactions in a Pediatric University Hospital. Eur J Clin Pharmacol. 2014;70(12):1519–26.
14. Montastruc J-L, Lugardon S, Desboeuf K, Fernet P, Lapeyre-Mestre M. Use of the capture-recapture method to assess the frequency of 'serious' adverse drug reactions: experience of Toulouse University Hospital. Bull Acad Natl Med. 2008;192(2):421–30; discussion 430-431.
15. World Health Organization. Uppsala Monitoring Centre. Glossary of pharmacovigilance terms. https://www.who-umc.org/global-pharmacovigilance/global-pharmacovigilance/glossary/. Accessed 3 December 2019.
16. Osmont M-N, Cuggia M, Polard E, Riou C, Balusson F, Oger E. Use of the PMSI for the detection of adverse drug reactions. Therapie. 2013;68(4):285–95.

.

17. Bouget J, Balusson F, Scailteux L-M, Maignan M, Roy P-M, L'her E, et al. Major bleeding with antithrombotic agents: a 2012-2015 study using the French nationwide Health Insurance database linked to emergency department records within five areas - rationale and design of SACHA study. Fundam Clin Pharmacol. 2019;33(4):443–62.

18. Oger E, Botrel MA, Juchault C, Bouget J. Sensitivity and specificity of an algorithm based on medico-administrative data to identify hospitalized patients with major bleeding presenting to an emergency department. BMC Med Res Methodol. 2019;19(1):194.

19. Arimone Y, Bidault I, Dutertre J-P, Gérardin M, Guy C, Haramburu F, et al. Réactualisation de la méthode française d'imputabilité des effets indésirables des médicaments. Therapie. 2011;66(6):517–25.

20. Miremont-Salamé G, Théophile H, Haramburu F, Bégaud B. Causality assessment in pharmacovigilance: The French method and its successive updates.Therapie. 2016;71(2):179-86.

21. Hook EB, Regal RR. Capture-recapture methods in epidemiology: methods and limitations. Epidemiol Rev. 1995;17(2):243–64.

22. Chapman CJ. Some properties of the hypergeometric distribution with applications to zoological censuses. U Calif Public Stat. 1951; l:131–60.

23. Wittes JT. On the bias and estimated variance of Chapman's two-sample capture-recapture population estimate. Biometrics.1972; 28:592–7.

24. Perrocheau et al. Estimation du nombre total de méningites à pneumocoque de l'enfant, par la méthode capture-recapture à 3 sources, France, 2001-2002. BEH. 2006

.

TABLES

Models	Estimated number of cases not captured in any data source (95%CL)	Total number of cases	Deviance	d.f	AIC	BIC
1-2, 2-3, 1-3	0	44	0	0	33.75	33.37
1-2, 1-3	26.0 (5.01 to 134)	70	3.065	1	34.81	34.49
1-2, 2-3	0	44	2.306	1	31.98	31.66
1-3, 2-3	0	44	0.005	1	32.17	31.84
1-2	7.42 (2.29 to 24.1)	51	8.483	2	36.69	36.41
1-3	17.5 (4.21 to 72.7)	61	0.005	2	34.17	33.90
2-3	0	44	2.628	2	30.18	29.91
Independent	6.75 (2.14 to 21.3)	51	8.649	3	35.24	35.02

Source 1 = 'PMSI', Source 2 = 'FPVDB', Source 3 = 'eHOP', d.f = degrees of freedom.

AIC: Akaike Information Criterion, BIC: Bayesian Information Criterion, eHOP: Clinical Data Warehouse

Table 1: Different models of interaction between sources

Estimated values	Maximum likelihood estimator	Nearly unbiased estimator*
Unobserved cell	979	957
Completeness of source PMSI	37.4%	37.9%
Completeness of source FPVDB	6.85%	6.94%
Total population	1679	1657
95% confidence limit	1296 to 2063	1273 to 2040

* taking into account estimation of bias resulting from source dependencies in subgroups (each participating center)

PMSI: Programme de Médicalisation des Systèmes d′Information, FPVDB: French Pharmacovigilance Database

Table 2: Capture-recapture estimate of the entire population using two sources

.

FIGURE LEGENDS

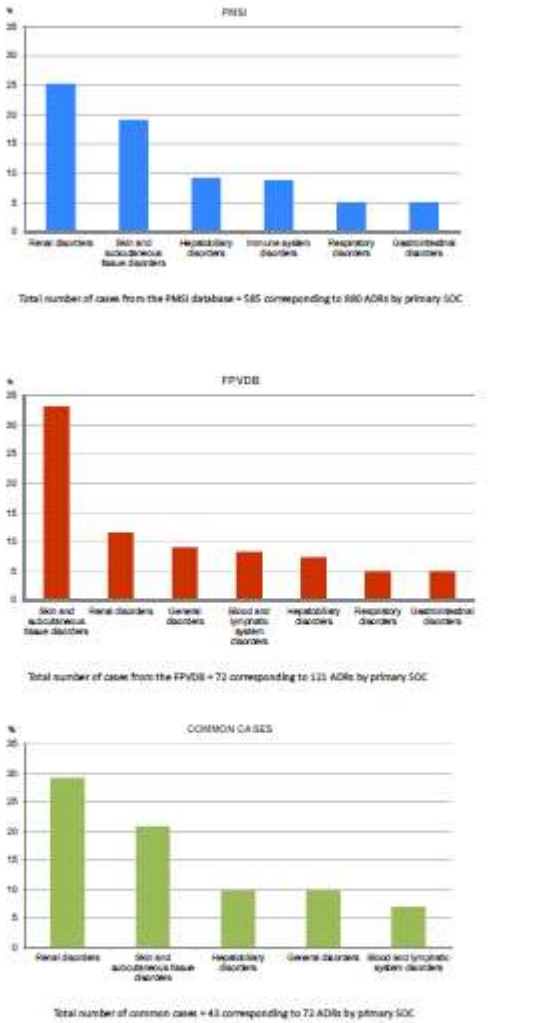


Figure 1: Adverse drug reactions in PMSI, FPVDB and duplicates using System Organ Classes

.

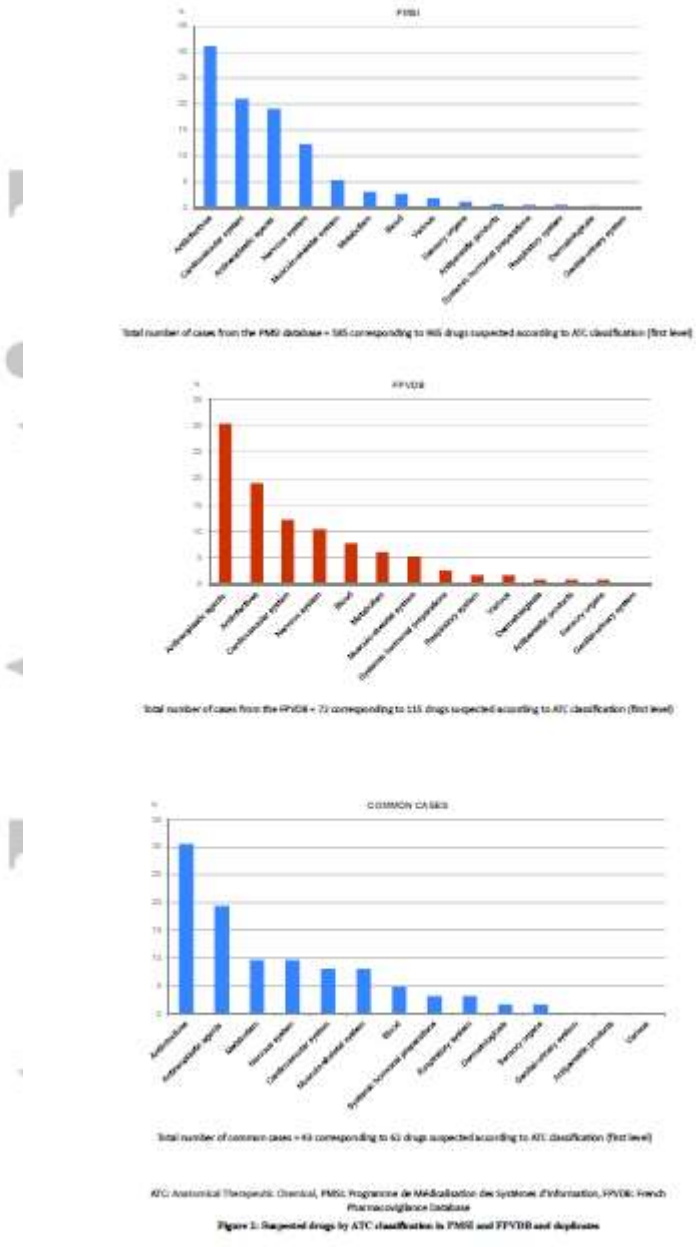


Figure 2: Suspected drugs by ATC classification in PMSI and FPVDB and duplicates

APPENDICES

Appendix 1

In the previous study performed in 2009, it was clearly not feasible to search for all existing ADRs. First, there needed to be specific ICD10 codes. Then the codes tested were to concern potentially serious ADRs that needed to be detected by pharmacovigilance for their potentially negative impact on the benefit-risk balance of a drug. Results of this study showed that the percentage of true positives (ADR cases retained from the patient records extracted) was above 40% for some ICD10 codes (T88.6, L27.0, L51.1, J70.2, J70.4, G62.0, N14.1 and N14.2) as well as for the association of code K71 with codes Y40 to Y59 / T88.7 (used to look for hepatic ADRs) [1]. Queries using these codes of interest were then used in the following years by the Rennes RPVC. The ICD10 codes L51.2, J70.3 and T88.2 and the association of code N17 with codes Y40 to Y59 / T88.7 were added to the queries on the basis of experiences in the RPVC in Brittany (Brest and Rennes).

.

The ICD10 codes for bleeding events with oral anticoagulants (antivitamin K and direct oral anticoagulants) were not selected for this study. Indeed, a cross-sectional study conducted by Rennes University Hospital supported specific data collection and a medical validation approach rather than an ICD10-based algorithm for assessing major bleeding [2,3].

1. Osmont M-N, Cuggia M, Polard E, Riou C, Balusson F, Oger E. Use of the PMSI for the detection of adverse drug reactions. Therapie. 2013;68(4):285–95.

2. Bouget J, Balusson F, Scailteux L-M, Maignan M, Roy P-M, L'her E, et al. Major bleeding with antithrombotic agents: a 2012-2015 study using the French nationwide Health Insurance database linked to emergency department records within five areas - rationale and design of SACHA study. Fundam Clin Pharmacol. 2019;33(4):443–62.

3. Oger E, Botrel MA, Juchault C, Bouget J. Sensitivity and specificity of an algorithm based on medico-administrative data to identify hospitalized patients with major bleeding presenting to an emergency department. BMC Med Res Methodol. 2019;19(1):194.

.

Appendix 2

The total number of ADRs (N) is calculated as: N = P * F/B where F is the total number of ADRs in the FPVDB, P is the total number of cases determined from the PMSI database, and B is the number of patients identified in both these sources.

The maximum likelihood estimator (MLE) and the adjustment originally suggested by Chapman [1] were applied to the two sources of data to obtain the estimate of the total number of ADRs. This estimation was shown by Wittes [2] to have optimal properties under a wide range of conditions, and to decrease the small-sample bias of the MLE. We provided variance-based intervals, acknowledging that this method underestimates both the upper and the lower limits, although the error margin may not be large [3]. The modified formula is as follows:

$N = [(F + 1)(P + 1)/(B + 1)] - 1$

The 95% CI of N is $N \pm 1.96 \sqrt{Var(N)}$

$Var(N) = (F + 1)(P + 1)(F - B)(P - B)/(B + 1)^2 (B + 2)$

Using each participating center as a subgroup, we estimated the bias in the capture-recapture estimation resulting from both source dependencies in subgroups and the differential catchability of cases by each source, as described by Hook [3].

A dependency present in one source increases (positive dependency) or decreases (negative dependency) the probability of being present in another source [4]. When the method is applied to more than two sources, the dependency across sources can be evaluated and controlled for during the analysis by introducing interaction terms between the sources into the model; a log-linear model allows cross-counting data to be analyzed in a contingency table and the number of cases expected in each cell of the contingency table to be estimated (including the empty cell). Thus the estimated value for the number X of cases identified by no source provides an estimate of N and its variance. Eight different log-linear models can be adjusted to the data, depending on the interactions considered. The fit of the model to the data was assessed by likelihood ratio statistics. The choice of the best model was based on the Akaike Information Criterion (AIC).

1. Chapman CJ. Some properties of the hypergeometric distribution with applications to zoological censuses. U Calif Public Stat. 1951; l:131–60.

2. Wittes JT. On the bias and estimated variance of Chapman's two-sample capture-recapture population estimate. Biometrics.1972; 28:592–7.

3. Hook EB, Regal RR. Capture-recapture methods in epidemiology: methods and limitations. Epidemiol Rev. 1995;17(2):243–64.

4. Perrocheau et al. Estimation du nombre total de méningites à pneumocoque de l'enfant, par la méthode capture-recapture à 3 sources, France, 2001-2002. BEH. 2006

Appendix 3

The first model assumed all possible interactions, except for 3 sources (model 1-2, 2-3, 1-3); because there is no degree of freedom in this model, it can be termed as "saturated". Other models are limited to or equivalent to two independent sources (e.g. models 1-2, 1-3, and 2-3) or two independent subsets of sources (models 1-2, 1-3; 1-2, 2-3; and 1-3, 2-3) and can be termed as “full two-source models”.

The independent model assumes all sources are independent.

Model 2-3 (FPVDB-eHOP) had the lowest AIC and Bayesian Information Criterion (BIC) values and predicted implausibly, no missing cases, as did the saturated model and the other models (1-2, 2-3 and 1-3,2-3) with close AIC values. Thus the “saturated” model and model 2-3 did not seem relevant because they predicted no case undetected by any of the sources. Other models predicted a total population ranging from 51 to 70. The model predicting missing cases and the lowest AIC value estimated a total population of 61.

The most relevant model seemed to be the one with an interaction 1-3 (PMSI-eHOP): it integrated an interaction between these two sources, which is in favor of the non-independence of these two sources. This appears consistent because the PMSI and eHOP were strongly related; indeed the PMSI data are included in eHOP.

.

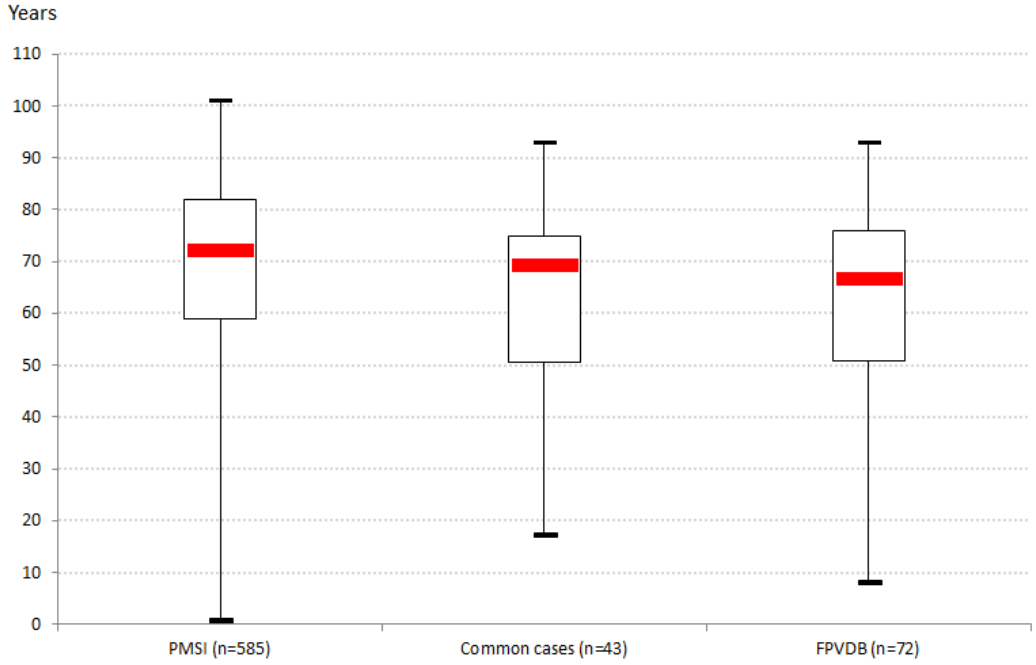
Appendix table 1: List of ICD10 codes used for queries

ICD-10 billing code	Description
Cutaneous adverse reactions	
L27.0	Generalized skin eruption due to drugs and medication
L51.1	Multiform bullous erythema
L51.2	Toxic epidermal necrolysis [Lyell]
Immuno-allergic adverse reactions	
T88.2	Shock due to anaesthesia
T88.6	Anaphylactic shock due to adverse effect of correct drug properly administered
Renal adverse reactions	
N14.1	Nephropathy induced by other drugs and biological substances
N14.2	Nephropathy induced by unspecified drug or biological substance
N17*	Acute renal failure
Hepatic adverse reactions	
K71*	Toxic liver disease
Neurological adverse reactions	
G62.0	Drug-induced polyneuropathy
Respiratory adverse reactions	
J70.2	Acute drug-induced interstitial lung disorders
J70.3	Chronic drug-induced interstitial lung disorders
J70.4	Drug-induced interstitial lung disorders, unspecified

* In combination with an ICD-10 code Y40-59 (“Drugs, medication and biological substances causing adverse effects in therapeutic use”) or T88.7 (“Unspecified adverse effect of drug or medication”)

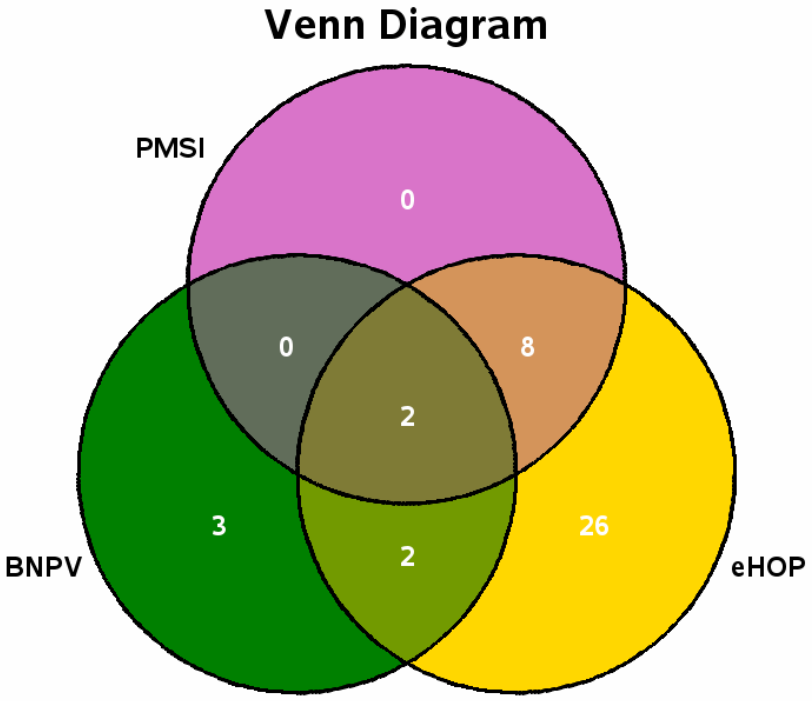
.

Appendix figure 1: Age distribution of patients by source



PMSI: Programme de Médicalisation des Systèmes d'Information, FPVDB: French Pharmacovigilance Database

Appendix figure 2: Survey restricted to Rennes University Hospital for anaphylactic shock, with three sources



Word count: 4075 words