Drug Saf
https://doi.org/10.1007/s40264-017-0618-y

ORIGINAL RESEARCH ARTICLE

An Automated System Combining Safety Signal Detection and Prioritization from Healthcare Databases: A Pilot Study

Mickael Arnaud[1] · Bernard Bégaud[1,2] · Frantz Thiessard[3,4] · Quentin Jarrion[1,2] · Julien Bezin[1,2] · Antoine Pariente[1,2] · Francesco Salvo[1,2]

Abstract

Introduction Signal detection from healthcare databases is possible, but is not yet used for routine surveillance of drug safety. One challenge is to develop methods for selecting signals that should be assessed with priority.

Aim The aim of this study was to develop an automated system combining safety signal detection and prioritization from healthcare databases and applicable to drugs used in chronic diseases.

Methods Patients present in the French EGB healthcare database for at least 1 year between 2005 and 2015 were considered. Noninsulin glucose-lowering drugs (NIGLDs) were selected as a case study, and hospitalization data were used to select important medical events (IME). Signal detection was performed quarterly from 2008 to 2015 using sequence symmetry analysis. NIGLD/IME associations were screened if one or more exposed case was identified in the quarter, and three or more exposed cases were identified in the population at the date of screening. Detected signals were prioritized using the Longitudinal-SNIP (L-SNIP) algorithm based on strength (S), novelty (N), and potential impact of signal (I), and pattern of drug use (P). Signals scored in the top 10% were identified as of high priority. A reference set was built based on NIGLD summaries of product characteristics (SPCs) to compute the performance of the developed system.

Results A total of 815 associations were screened and 241 (29.6%) were detected as signals; among these, 58 (24.1%) were prioritized. The performance for signal detection was sensitivity = 47%; specificity = 80%; positive predictive value (PPV) 33%; negative predictive value = 82%. The use of the L-SNIP algorithm increased the early identification of positive controls, restricted to those mentioned in the SPCs after 2008: PPV = 100% versus PPV = 14% with its non-use. The system revealed a strong new signal with dipeptidylpeptidase-4 inhibitors and venous thromboembolism.

Conclusion The developed system seems promising for the routine use of healthcare data for safety surveillance of drugs used in chronic diseases.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s40264-017-0618-y) contains supplementary material, which is available to authorized users.

✉ Francesco Salvo
francesco.salvo@u-bordeaux.fr

[1] University of Bordeaux, Inserm, Bordeaux Population Health Research Center, Pharmacoepidemiology Team, UMR 1219, 146 Rue Léo Saignat, BP 36, 33000 Bordeaux Cedex, France

[2] CHU de Bordeaux, Pôle de santé publique, Service de pharmacologie médicale, 33000 Bordeaux, France

[3] University of Bordeaux, Inserm, Bordeaux Population Health Research Center, ERIAS Team, UMR 1219, 33000 Bordeaux, France

[4] CHU de Bordeaux, Pôle de santé publique, Service d'Information médicale, 33000 Bordeaux, France

Published online: 28 November 2017

Key Points

Automated signal detection from healthcare databases is a new avenue for drug safety monitoring. However, the huge number of signals expected to be detected requires a signal prioritization process.

In this study, an automated system, adapted to longitudinal healthcare data, and combining detection and prioritization of safety signals, was developed and applied to drugs used in chronic diseases, with noninsulin glucose-lowering drugs as a case study.

The developed system provided a promising performance, suggesting that it could be used routinely for selecting the most relevant safety signals, and it identified a new signal concerning dipeptidylpeptidase-4 inhibitors and venous thromboembolic events.

1 Introduction

The major drug safety issues that have arisen in recent decades all concern drugs used for the treatment of chronic diseases and adverse events (AEs) a priori not evocative of a drug causation. For example, rofecoxib, a non-steroidal anti-inflammatory drug expected to convey a much lower risk of gastrointestinal bleeding, was found to increase the risk of myocardial infarction once millions of people had already been exposed [1, 2]. The proton pump inhibitors, which are used to prevent or treat various gastrointestinal disorders such as peptic ulcer and its complications, were among the most prescribed drugs in the mid-2000s, notably as they were deemed to have a good safety profile. However, they were secondarily shown to increase the risk of community-acquired pneumonia [3, 4] and osteoporosis-related fractures [5–7]. In the late 2000s, important safety alerts emerged concerning the thiazolidinediones, a class of noninsulin glucose-lowering drugs (NIGLDs), and showed that rosiglitazone was associated with an increased risk of cardiovascular events [8, 9], while pioglitazone could induce bladder cancers [10–12].

Current systems of drug surveillance, which are essentially based on spontaneous reporting, are relatively inefficient when it comes to detecting signals involving diseases or events relatively common in the general population and for which a drug causation is not a priori suspected. The access to huge longitudinal data through healthcare databases is useful in this respect by allowing large cohorts to be followed over time. Research in this area has burgeoned for over a decade with the launching of several initiatives worldwide [13–17], some of which have focused on signal detection applied to healthcare databases [18–21]. Even if they are potentially helpful for identifying safety signals that would have been ignored using traditional methods, data mining methods are known to detect a huge numbers of statistical associations representing potential signals that must be assessed to be confirmed or disproved. If signal detection is performed routinely, a method that can help in selecting priority signals would be needed. Routine signal detection thus needs to be combined with an automated prioritization process.

Numerous algorithms have been developed for the automated prioritization of safety signals detected from spontaneous reporting databases, and some of these are currently used routinely [22–35]. They are all based—entirely or in part—on three key aspects: the strength, novelty, and public health impact of the safety signal. The importance of these aspects was outlined in 1999 by Waller and Lee with the SNIP concept 'Strength, Novelty, and clinical Importance of the signal, and potential for Prevention' [22]. However, experiments in prioritization of safety signals detected in longitudinal data are scarce. The EU-ADR consortium developed an approach that aimed at identifying the signals that were less likely to result from biases, which can be considered as a kind of signal prioritization. Nevertheless, it did not take into account some of the key aspects for prioritization (e.g., the public health impact), and its performance was not assessed [36].

The objective of this pilot study was to develop and assess an automated system combining the detection and prioritization of safety signals identified in healthcare databases and involving treatments of chronic diseases.

2 Methods

2.1 Data Source

The *Echantillon Généraliste des Bénéficiaires* (EGB) claims database is a 1/97th permanent representative sample of the population covered by the French national healthcare insurance system [37]. It contains individual, anonymous, and comprehensive outpatient drug reimbursement and hospitalization data from beneficiaries of the general scheme, which includes salaried workers and their dependents, unemployed, and retired salaried workers, with the exception of civil servants and non-working students (77% of the French population). Drugs are coded according to the anatomical therapeutic chemical (ATC)

classification and hospitalization diagnoses according to the International Classification of Diseases, 10th revision (ICD-10).

2.2 Population, Exposure, and Event Definitions

All persons present in the EGB for at least 1 year between 2005 and 2015 were considered.

NIGLDs were selected as examples of treatments of chronic diseases. Reimbursements were considered as surrogates for drug exposure. Active substances were identified through ATC classification of blood glucose-lowering drugs, excluding insulins (ATC code: A10B). In the event of fixed combinations of NIGLDs, the exposure was considered in each corresponding active substance (e.g., a fixed combination 'metformin and sitagliptin' was considered for both exposure to metformin and to sitagliptin). For each patient and for each NIGLD considered, only the first reimbursement was selected if it occurred after a 12-month run-in period, to ascertain that it was an incident drug exposure.

Hospitalization diagnoses were considered as surrogates for adverse events and were identified by ICD-10 codes grouped according to the first three characters (e.g., I26: pulmonary embolism). The ICD-10 codes were aligned to those included in the important medical event (IME) terms list of the MedDRA® dictionary using the unified medical language system tool to select the ICD-10 that needed to be monitored (i.e., events that per se can result in death, are life-threatening, or cause prolonged hospitalization or persistent disability; see Electronic Supplementary Material 1) [38]. Similarly to the incident drug exposure, it was assumed that the first occurrence of a hospitalization associated with a diagnosis related to the ICD-10 code of interest, if it occurred after a 12-month run-in period, could be considered as an incident event. This approach has been validated in previous studies using sequence symmetry analysis [39, 40].

2.3 Signal Detection

Sequence symmetry analysis (SSA) was selected for signal detection analyses on the basis of a literature review that focused on the comparison of performance of methods used for that purpose [41]. SSA is dedicated to longitudinal data and developed for large-scale standardized applications; it provided good detection performance, notably by considering self-controlled analyses and controlling for temporal trends, which allows the detection of spurious associations to be partly preserved.

SSA compares the number of patients who presented, over a given time window, the sequence 'drug in first, event in second' during the study period with that of patients who presented the reverse sequence 'event in first, drug in second' [42, 43]. The crude ratio of these two types of sequences is, by essence, not affected by confounders that are constant over time, but is sensitive to changes in prescription and hospitalization trends. To address this issue, an adjustment for correcting such temporal trends is applied by dividing the crude sequence ratio by the null-effect sequence ratio. The latter is the sequence ratio that would have been expected from the trends if the drug and the event were considered independently. The corrected sequence ratio obtained after this adjustment is termed the *adjusted sequence ratio* (ASR) [43].

Signal detection was performed quarterly from 2008 to 2015 considering a time window of 12 months for the pre- and post-drug initiation periods [40]. In each analysis, NIGLD/IME associations were screened if (1) one or more exposed case was identified in the quarter, (i.e., if a patient was hospitalized for the event of interest in the quarter and initiated the NIGLD of interest in the previous 12 months) and (2) three or more exposed cases were identified in the population at the date of screening. If no sequence 'event in first, drug in second' was observed, the value 0.49 was assigned to compute the crude sequence ratio [39]. NIGLD/IME associations were considered as signals if the lower limit of the bootstrapped (500 replicates) 95% confidence interval (95% CI) of the ASR exceeded 1 [40].

2.4 Signal Prioritization

The principal algorithms for the prioritization of signals derived from spontaneous reporting data were reviewed [22–35]. The principle of the SNIP strategy developed by Waller and Lee [22] was adapted to build an algorithm for the prioritization of safety signals detected from longitudinal healthcare data, the Longitudinal-SNIP (L-SNIP), which is based on strength of the signal (S), novelty of the signal (N), potential impact of the signal (I), and patterns of drug use (P). The L-SNIP algorithm included a total of 14 criteria related to scientific, medical, and economic aspects and factors potentially conditioning perception of the signal (Table 1):

- Strength: the risk estimate, the minimal risk (the lower limit of 95% CI), precision of the risk estimate;
- Novelty: absence of the signal in the Summary of the Product Characteristics (SPC), drug seniority, and an increase in risk over time;
- Potential impact: the potential number of attributable cases, and the cost of hospitalization for this event;
- Pattern of drug use: the event is not related to the drug indications, the proportion of drug users among frail populations (children, childbearing women), the

Table 1 Criteria included in prioritization algorithm

Criteria	Definition	Weight	Categories	Score
Strength of signal				
Risk estimate	Value of ASR	3	> 3	1
			1.5–3	0.5
			< 1.5	0
Minimal risk	Value of lower limit of 95% CI of ASR	4	> 2	1
			1.5–2	0.5
			< 1.5	0
Precision of risk estimate	Range of 95% CI of ASR	2	< 3	1
			3–6	0.5
			> 6	0
Novelty of signal				
Absence in SPC	Association not mentioned in the French version of the SPC corresponding to the year of analysis	4	Yes	1
			No	0
Drug seniority	Duration of drug on market at time of analysis	2	< 6 y	1
			6–10 y	0.5
			> 10 y	0
Increase in risk over time	Evolution of ASR between time of analysis and three previous detection analyses	1	> 50%	1
			0–50%	0.5
			< 0%	0
Impact of signal				
Potential number of attributable cases	Number of attributable cases potentially related to drug use based on the prevalence of use of the drug and the incidence of the event at the year of analysis	2	> 1000 cases	1
			200–1000 cases	0.5
			< 200 cases	0
Cost of hospitalization for event	Mean cost of hospitalization associated with diagnosis corresponding to event	2	> €6000	1
			€3000–6000	0.5
			< €3000	0
Patterns of drug use				
Event not related to drug indications	Event not mentioned in SPC as indication of drug	4	Yes	1
			No	0
Drug use in vulnerable population (1): children	Proportion of children aged 0–15 years among drug users	3	≥ 10%	1
			< 10%	0
Drug use in vulnerable population (2): childbearing women	Proportion of women aged 15–49 years among drug users	3	≥ 10%	1
			< 10%	0
Prevalence of drug use	Prevalence of drug use in year of analysis	2	> 1/100 persons	1
			0.1/100–1/100 persons	0.5
			< 0.1/100 persons	0
Incidence of drug use	Incidence of drug use in year of analysis	2	> 1/1000 persons	1
			0.1/1000–1/1000 persons	0.5
			< 0.1/1000 persons	0
Increase in incidence of drug use over time	Evolution of incidence of drug use between year of analysis and 2 previous years	1	> 20%	1
			0–20%	0.5
			< 0%	0

ASR adjusted sequence ratio, *95% CI* 95% confidence interval, *SPC* summary of product characteristics

prevalence of drug use, the incidence of drug use, and the increase in the incidence of drug use over the 2 years preceding that of the analysis.

All the criteria were grouped into two or three categories, each being associated with a score of 0, 0.5, and 1. Thereafter they were weighted using a coefficient that ranged from 1 to 4 according to their assumed importance for a decision-making process (Table 1). All these steps (i.e., selection, scoring, and weighting of the criteria) were performed on the basis of a consensus between three senior experts in pharmacovigilance and pharmacoepidemiology (BB, FS, AP) and after the analysis of the existing literature. The signals in the top 10% of the L-SNIP scores, which corresponded to the weighted sum of the scores of the 14 criteria, were identified as *high priority* for the validation process.

2.5 Assessment of the Performance of Signal Detection and Prioritization

All the potential NIGLD/IME associations with sufficient power to detect a relative risk of 2 based on the drug and event prevalence estimated in the EGB database were considered [44].

For measuring the performance of the developed system for signal detection and prioritization, an original reference set was built. The annual versions of the SPCs between 2008 and 2017 were reviewed to be able to identify the drug/IME associations that constituted positive controls of our reference set. The negative controls were randomly selected from the other associations to avoid bias in performance assessment when a large imbalance between the number of positive and negative controls is present [45]. The selection of the negative controls was performed to have a ratio of negative/positive controls of three. The performance for the signal detection was determined using the sensitivity (Se), specificity (Sp), positive predictive value (PPV), and negative predictive value (NPV). For measuring the added value of the signal prioritization for the identification of relevant signals, the positive controls were restricted to those that were mentioned in the SPCs after 2008. The assessment measured the ability of the system to identify these signals before they were mentioned in the SPCs and the performance was compared with that obtained when the L-SNIP algorithm was not used (i.e., when considering only the single signal detection).

In addition, the prioritized signals were screened to find any potentially relevant new safety issue. The relevance of these signals was initially judged by the three senior experts (BB, FS, AP) based on their experience and the results of the signal detection. For signals that were considered relevant, complementary analyses were performed to ascertain whether they were specific to the drug or the drug class, and to the event or the group of related events. All analyses were performed using SAS Enterprise Guide 7.1 (SAS Institute, Inc., Cary, NC, USA).

3 Results

3.1 Trends of Signal Detection and Prioritization Between 2008 and 2015

A total of 920 NIGLD/IME associations with sufficient power to be detected in EBG were considered. Among them, 815 associations were screened at least once between the first quarter of 2008 (2008-Q1) and the end of the study period (2015-Q4). A total of 241 (29.6%) signals were detected during the study period (51 signals [6.3%] only once, and 190 [23.3%] at least twice); among these, 58 signals (24.1%) were identified as of high priority (22 signals [9.1%] prioritized only once, and 36 [14.9%] at least twice).

Over the study period, the number of NIGLD/IME associations screened increased from 216 in 2008-Q1 to 384 in 2011-Q2, before decreasing to 257 associations screened in 2015-Q4 (Fig. 1). Similarly, the number of signals detected increased between 2008-Q1 and 2011-Q4, from 35 to 73, and decreased up to 56 signals in 2015-Q4. The number of signals identified as of high priority fluctuated between four signals in 2008-Q1 and ten signals in 2009-Q4, 2011-Q2, 2011-Q3, and 2013-Q1.

3.2 Performance of Signal Detection and Prioritization

The reference set included 15 positive controls and 45 negative controls (see Electronic Supplementary Material 2). The signal detection identified 21 controls, including seven among the positive ones, which corresponded to the following performance: Se of 47%; Sp of 69%; PPV of 33%; NPV of 80% (Table 2).

Among the positive controls, three were added to the SPCs after 2008. The Se was 33%, Sp was $>80\%$, and NPV was around 95% whether the L-SNIP algorithm was used or not; however, PPV was 100% when the prioritization was combined with the detection, while this value dropped to 14% when only the detection was considered (Table 3).

3.3 Identification of New Relevant Signals

The 58 signals that were identified at least once as of high priority during the study period have been screened in depth, and three of them appeared to be of interest. All

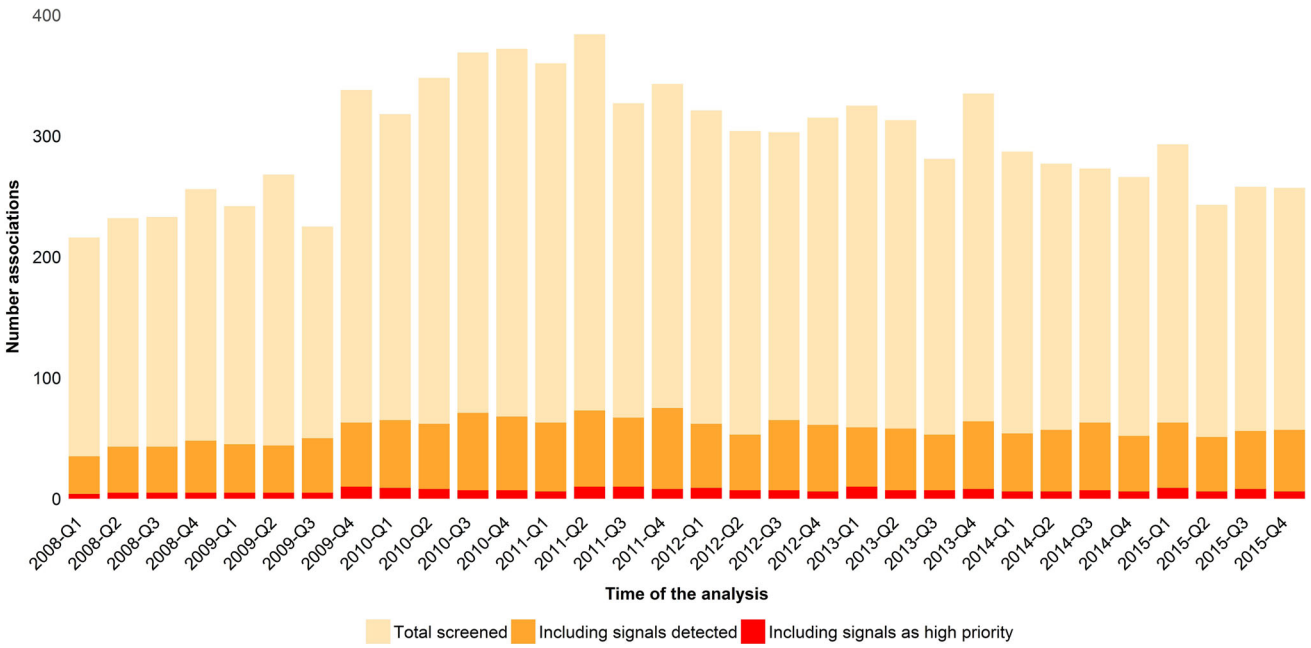

Fig. 1 Global trend of signal detection/prioritization between 2008 and 2015

Table 2 Results of the performance assessment for the signal detection

	Positive controls	Negative controls	
Detected	7	14	PPV = 7/(7 + 14) = 33%
Not detected	8	31	NPV = 31/(31 + 8) = 80%
	Se = 7/(7 + 8) = 47%	Sp = 31/(31 + 14) = 69%	

NPV negative predictive value, *PPV* positive predictive value, *Se* sensitivity, *Sp* specificity

concerned DPP-4 inhibitors and the occurrence of venous thromboembolic events, and shared a similar detection profile with a strengthening of the risk estimate over time (Fig. 2).

The association between vildagliptin and pulmonary embolism was screened over five different quarters during the study period, and was found significant for the last four screenings. It was identified as of high priority in 2015-Q1 (seven cases; ASR 7.3, 95% CI 2.1–13.0), and in 2015-Q4 (8 cases; ASR 8.3, 95% CI 3.1–15.5). Pulmonary embolism was also found to be associated with saxagliptin, the association being identified as of high priority for the second screening in 2013-Q2 (four cases; ASR 8.7, 95% CI 2.1–9.3). The third signal concerned 'venous thromboembolism and thrombosis' (not otherwise specified) with sitagliptin. The signal strengthened over time to finally reach five cases and an ASR of 10.0 (95% CI 1.9–10.4) in 2015-Q3. A fourth signal (not identified as of high priority) was detected during the study period and concerned vildagliptin and phlebitis/thrombophlebitis. This association was screened nine times and remained stable and significant in the last five screenings. In 2015-Q4 it concerned 13 cases with an ASR of 3.1 (95% CI 1.6–11.0) (Fig. 2).

Complementary analyses were performed to further assess this potential safety issue with the use of DPP-4 inhibitors and the risk of venous thromboembolic events. An analysis of DPP-4 inhibitors and arterial thromboembolic events identified two signals during the study period, but their profile of detection fluctuated between significance and non-significance over time. A further search on signals related to venous thromboembolic events retrieved two other signals (one for acarbose and another for gliclazide) that shared a similar profile of detection to that observed with DPP-4 inhibitors but with a lower strength of ASR, and a lower 95% CI value that constantly remained at the limit of significance (see Electronic Supplementary Material 3 and 4).

4 Discussion

This pilot study aimed at improving the tools used for signal detection in longitudinal healthcare databases by developing an automated system combining both detection and prioritization of signals related to treatments of chronic diseases. In the proposed approach, signal detection was

Table 3 Performance of the developed system according to the presence or absence of the use of the L-SNIP algorithm for signal prioritization

	Positive controls added in the SPC after 2008	Negative controls	
Detection + prioritization			
Identified before mention in the SPC	1	0	PPV = 1/(1 + 0) = 100%
Identified after mention in the SPC or never identified*	2	45	NPV = 45/(45 + 2) = 96%
	Se = 1/(1 + 2) = 33%	Sp = 45/(45 + 0) = 100%	
Detection only			
Identified before mention in the SPC	1	6	PPV = 1/(1 + 6) = 14%
Identified after mention in the SPC or never identified	2	39	NPV = 39/(39 + 2) = 95%
	Se = 1/(1 + 2) = 33%	Sp = 39/(39 + 6) = 100%	

NPV negative predictive value, *PPV* positive predictive value, *Se* sensitivity, *Sp* specificity, *SPC* summary of product characteristics

performed quarterly and the detected signals were subsequently prioritized according to their relevancy by using the L-SNIP algorithm combining 14 criteria based on the strength, novelty, and potential impact of the signal, and on the patterns of use of the drug concerned. The assessment of the developed system showed that it performs satisfactorily, and that the L-SNIP algorithm could be used in order to improve the feasibility of routine signal detection from healthcare databases. The developed system also identified a new signal with the use of the DPP-4 inhibitors and the risk of venous thromboembolic events, which deserves to be investigated with further ad-hoc and more robust studies.

Signal detection was performed quarterly, a periodicity already adopted by the Uppsala Monitoring Center for signal detection from the spontaneous reporting database VigiBase [27]. The European Medicines Agency has used shorter intervals (monthly or twice a month) [33], but this does not seem to perform better than quarterly screening [46]. From a pragmatic perspective, associations were screened only if a new exposed case was identified during the quarter considered and if at least three similar drug–event associations were found in the study population, a strategy previously shown to offer the best compromise [46]. Signal detection was performed using SSA, as it offers an interesting balance between good detection performance, notably by including self-controlled analyses and controlling for temporal trends to minimize the detection of false-positive associations, and ease of use [41].

In this pilot study, the signal detection of the developed system proved interesting, and was comparable to that obtained previously in a validation study of the SSA performed by the AsPEN consortium [40]. However, this result should be considered with caution given the low number of controls included in the reference set. The screening of only the ICD-10 codes related to IMEs according to MedDRA® reduced the number of signals detected and prioritized, even if this number remained large. This was expected, as it is typically encountered with automated safety signal detection from either spontaneous reports or healthcare data. The high number of detected signals reinforces the idea that a form of prioritization of safety signals is needed if routine use of healthcare databases is envisaged. As it is intended to be used as a routine method, events were retrieved through the 'native' ICD-10 classification without any form of grouping of codes related to the same given disease. It is obvious that grouping ICD-10 codes would reduce the number of signals related to spurious or duplicate events (e.g., acute renal failure vs unspecified kidney failure). For routine purposes, it would be worth defining a priori medical events that are likely to be captured in healthcare databases, and then selecting the codes that seem the most appropriate for their identification [47, 48]. The use of grouped codes, at least for the most relevant drug-related disorders such as those proposed by the EU-ADR project [49], could be useful in this perspective.

The added value of the signal prioritization was assessed by measuring the ability of the developed system to identify the NIGLDs/IME associations before they were mentioned in the corresponding SPCs. The assessment showed promising results, as the use of the L-SNIP algorithm allowed a seven-fold increase in the probability of retrieving IME secondarily cited in the SPC among the associations flagged compared with the single signal detection. As for the performance obtained for signal detection, this result has to be considered with caution given the low number of controls. Nevertheless, this study

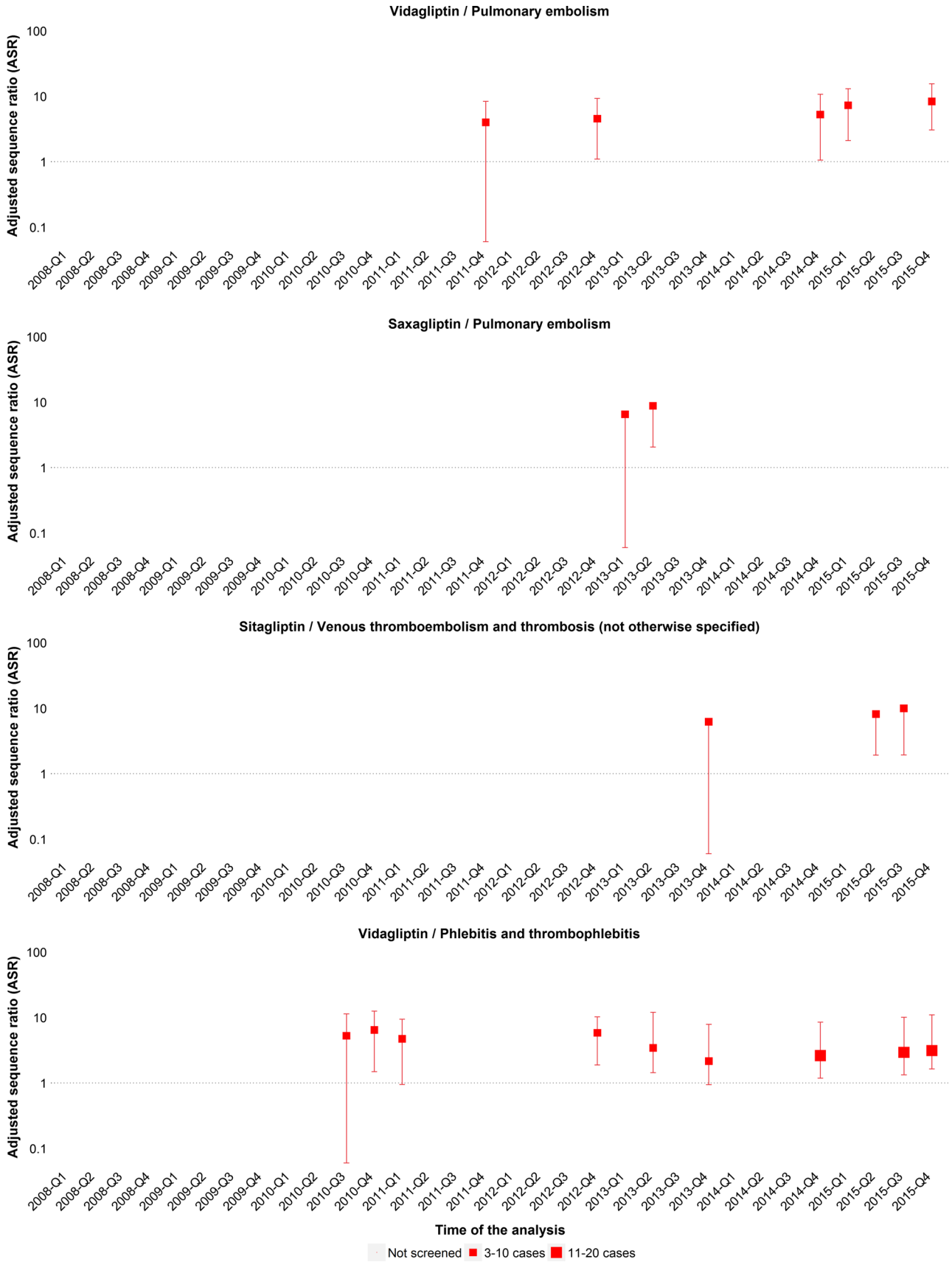

Fig. 2 Results of signal detection for signals related to DPP-4 inhibitors with venous thromboembolism events

indicated that the developed system is able to find new signals, such as venous thromboembolic events related to the use of DPP-4 inhibitors, which seem relevant enough for planning further and urgent investigations. Three different signals concerning pulmonary embolism, venous embolism, and thrombosis were considered high priority at least once during the study period. They also shared a similar profile of detection that could suggest an actual relationship, notably with a strengthening of the risk estimate after the occurrence in patients exposed to DPP-4 inhibitors. Complementary analyses also suggested a possible role of DPP-4 inhibitors in venous thromboembolism: (1) a fourth detected signal concerned phlebitis and thrombophlebitis, but it was not identified as of high priority; (2) there was no signal of arterial thromboembolic events related to DPP-4 inhibitors; (3) only two non-prioritized signals concerned venous thromboembolic events related to other NIGLDs. No pre-clinical or clinical data can support or rule out this association. This absence of external evidence concerning the signals detected in this study, as well as the large number of patients exposed to DPP4-inhibitors [50], emphasizes the need for an urgent and more robust investigation.

The developed system was not able to prioritize cardiac failure related to rosiglitazone in the top 10% of signals, as this event was already mentioned in its SPC in 2008. Nevertheless, this signal was detected from the beginning of the study period, and at least once a year up to the withdrawal of rosiglitazone, with ASR values around 4 (data not shown). The system was not able to detect bladder cancers related to pioglitazone, as the risk found in the literature was quite low [10–12], and the EGB database does not provide sufficient power to detect it. The absence could also be due to a limitation of SSA for the detection of cancers, as this method considered a 12-month period to identify exposed cases. While this period seems appropriate for most of the adverse events that can be captured from healthcare databases, it could be too short for detecting drug-related cancers that have late effects. Consequently, cancers could need specific tailored methods to be detected.

Thirty-two detected signals oscillated between high and low priority. It was typically the case of associations that were mentioned in the SPC after their first detection in EGB. For example, the association between sitagliptin and acute renal failure was detected and prioritized before its mention in the SPC, while it was never re-ranked in the top 10% L-SNIP score after its mention in the SPC. For the other associations (e.g., metformin and malignant neoplasm of pancreas), the oscillations between the two statuses were due to their ranking at the limit of the 10% threshold. Thus, for these signals, a slight modification of the risk estimate or the patterns of drug use from one quarter to another can lead to the prioritization threshold going up or going down. This could be considered as expected, and did not alter the value of the L-SNIP algorithm.

The criteria included in the L-SNIP algorithm were selected on the basis of the information available in healthcare databases and with the view to develop an automated and efficient system of prioritization that can be easily combined with signal detection. Some criteria were derived from the algorithms of prioritization applied to signals detected in spontaneous report databases. For example, the criterion 'drug seniority' was considered whereby the full score was given to drugs marketed for <6 years, as it was previously demonstrated that safety signals concerning drugs that were marketed for ≤ 5 years were associated with their prioritization by the experts of the Dutch Medicines Agency [29]. The criterion 'increase in risk over time' was considered as a novelty criterion, as it is more likely to reflect a new aspect of a known safety signal, and it fits with the definition of a signal provided by Hauben and Aronson in 2009 [51]. The access to longitudinal healthcare data allowed us to consider criteria based on patterns of drug use, such as prevalence or incidence of drug use, which is information that is lacking in algorithms based on spontaneous reporting data. The criterion 'event not related to drug indications' was added to reduce the impact of protopathic and indication biases, which are well known shortcomings of signal detection performed on healthcare databases [41]. Only two criteria focusing on frail populations (children and childbearing women) were retained; the elderly were not included in the algorithm as chronic diseases are frequent in this population, so this criterion would not have been discriminant. In addition, the number of criteria considered in the L-SNIP algorithm was limited to make it more discriminant. In that respect, some criteria classically used for signal prioritization were not included since we considered them to be less relevant in this context (e.g., 'positive rechallenge' [30]) or because they are hardly automatable (e.g., 'biological plausibility' [24]). Criteria relative to scientific and medical issues were given pride of place, since it was assumed that they are of greater importance in the decision-making process than economic factors or public perception. Using a single criterion for economic factors thus seemed reasonable; the cost of hospitalizations induced by the drug was preferred to the cost of the drug itself.

To combine the 14 criteria included in the L-SNIP algorithm, each was first transformed into a score based on a consensual analysis of the scientific literature by three senior experts in pharmacovigilance and pharmacoepidemiology. Techniques classically used to perform this task (e.g., normalization [32], ranking [28]) were not applicable, as the selected criteria referred to different dimensions of the signal; that is, criteria specific to the signal itself (e.g., 'minimal potential risk'), or only to the drug involved in the signal (e.g., 'seniority of the drug'). A consensus for weighting the criteria

was also sought by considering their relevancy for the decision-making process. Among the five weighting processes retrieved in the literature [24, 28, 30–32], the one considering a 1:4 ratio between the least and the most important criterion [24] was chosen. The other processes use a much leaner ratio (up to the second decimal), making clinical relevancy of the weights questionable. To limit the number of signals to be managed, a threshold, arbitrarily set to the top 10% of the highest values of the priority score obtained from the signals detected in the quarter, was introduced to focus on signals of prime interest. The use of different cut-offs at 15% and 20% did not increase the number of prioritized signals (data not shown), while larger cut-offs seemed unsuitable for an algorithm that aimed to select the most relevant signals among the huge number of those detected.

5 Conclusion

The developed system performed well and its potential application for routine signal detection and prioritization seems promising. The L-SNIP algorithm seemed to correctly prioritize the relevance of signals detected. Further research in a wider range of drug classes and using different definitions of events is needed to definitely validate this system. Meanwhile, an investigation is urgently required to support or rule out the strong signal between the use of DPP-4 inhibitors and the risk of venous thromboembolic events demonstrated in this study.

Compliance with ethical standards

Conflicts of interest Mickael Arnaud, Bernard Bégaud, Frantz Thiessard, Quentin Jarrion, Julien Bezin, Antoine Pariente, and Francesco Salvo have no conflict of interest directly relevant to the content of this study.

Funding This study is part of the Drugs Systematized Assessment in real-liFe Environment (DRUGS-SAFE) research program that is funded by the French Medicines Agency (*Agence Nationale de Sécurité du Médicament et des Produits de Santé*, ANSM). The program aims at providing an integrated system allowing the concomitant monitoring of drug use and safety in France. The potential impact of drugs (e.g., non-insulin glucose-lowering drugs), the frailty of populations, and the seriousness of risks drive the research program. This publication represents the views of the authors and does not necessarily represent the opinion of the French Medicines Agency.

References

1. Mukherjee D, Nissen SE, Topol EJ. Risk of cardiovascular events associated with selective COX-2 inhibitors. JAMA. 2001;286:954–9.
2. Lévesque LE, Brophy JM, Zhang B. The risk for myocardial infarction with cyclooxygenase-2 inhibitors: a population study of elderly adults. Ann Intern Med. 2005;142:481.
3. Laheij RJ, Sturkenboom MC, Hassing R-J, Dieleman J, Stricker BH, Jansen JB. Risk of community-acquired pneumonia and use of gastric acid–suppressive drugs. JAMA. 2004;292:1955–60.
4. Johnstone J, Nerenberg K, Loeb M. Meta-analysis: proton pump inhibitor use and the risk of community-acquired pneumonia. Aliment Pharmacol Ther. 2010;31:1165–77.
5. Yang Y-X, Lewis JD, Epstein S, Metz DC. Long-term proton pump inhibitor therapy and risk of hip fracture. JAMA. 2006;296:2947–53.
6. Vestergaard P, Rejnmark L, Mosekilde L. Proton pump inhibitors, histamine H2 receptor antagonists, and other antacid medications and the risk of fracture. Calcif Tissue Int. 2006;79:76–83.
7. Targownik LE, Lix LM, Metge CJ, Prior HJ, Leung S, Leslie WD. Use of proton pump inhibitors and risk of osteoporosis-related fractures. Can Med Assoc J. 2008;179:319–26.
8. Lipscombe LL, Gomes T, Lévesque LE, Hux JE, Juurlink DN, Alter DA. Thiazolidinediones and cardiovascular outcomes in older patients with diabetes. JAMA. 2007;298:2634–43.
9. Nissen SE, Wolski K. Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes. N Engl J Med. 2007;356:2457–71.
10. Neumann A, Weill A, Ricordeau P, Fagot JP, Alla F, Allemand H. Pioglitazone and risk of bladder cancer among diabetic patients in France: a population-based cohort study. Diabetologia. 2012;55:1953–62.
11. Tuccori M, Filion KB, Yin H, Yu OH, Platt RW, Azoulay L. Pioglitazone use and risk of bladder cancer: population based cohort study. BMJ. 2016;352:i1541.
12. Hillaire-Buys D, Faillie J-L. Pioglitazone and the risk of bladder cancer. BMJ. 2012;344:e3500.
13. Coloma PM, Schuemie MJ, Trifirò G, Gini R, Herings R, Hippisley-Cox J, et al. Combining electronic healthcare databases in Europe to allow for large-scale drug safety monitoring: the EU-ADR Project. Pharmacoepidemiol Drug Saf. 2011;20:1–11.
14. Wisniewski AFZ, Bate A, Bousquet C, Brueckner A, Candore G, Juhlin K, et al. Good signal detection practices: evidence from IMI PROTECT. Drug Saf. 2016;39:469–90.
15. Platt R, Carnahan RM, Brown JS, Chrischilles E, Curtis LH, Hennessy S, et al. The U.S. Food and Drug Administration's Mini-Sentinel program: status and direction. Pharmacoepidemiol Drug Saf. 2012;21:1–8.
16. Stang PE, Ryan PB, Racoosin JA, Overhage JM, Hartzema AG, Reich C, et al. Advancing the science for active surveillance: rationale and design for the Observational Medical Outcomes Partnership. Ann Intern Med. 2010;153:600–6.
17. Andersen M, Bergman U, Choi N-K, Gerhard T, Huang C, Jalbert J, et al. The Asian Pharmacoepidemiology Network (AsPEN): promoting multi-national collaboration for pharmacoepidemiologic research in Asia. Pharmacoepidemiol Drug Saf. 2013;22:700–4.
18. Ryan PB, Stang PE, Overhage JM, Suchard MA, Hartzema AG, DuMouchel W, et al. A comparison of the empirical performance of methods for a risk identification system. Drug Saf. 2013;36:143–58.
19. Schuemie MJ, Gini R, Coloma PM, Straatman H, Herings RMC, Pedersen L, et al. Replication of the OMOP experiment in Europe: evaluating methods for risk identification in electronic health record databases. Drug Saf. 2013;36:159–69.
20. Pratt N, Andersen M, Bergman U, Choi N-K, Gerhard T, Huang C, et al. Multi-country rapid adverse drug event assessment: the Asian Pharmacoepidemiology Network (AsPEN) antipsychotic and acute hyperglycaemia study. Pharmacoepidemiol Drug Saf. 2013;22:915–24.

21. Kulldorff M, Dashevsky I, Avery TR, Chan AK, Davis RL, Graham D, et al. Drug safety data mining with a tree-based scan statistic. Pharmacoepidemiol Drug Saf. 2013;22:517–23.
22. Waller PC, Lee EH. Responding to drug safety issues. Pharmacoepidemiol Drug Saf. 1999;8:535–52.
23. Waller P, Heeley E, Moseley J. Impact analysis of signals detected from spontaneous adverse drug reaction reporting data. Drug Saf. 2005;28:843–50.
24. Seabroke S, Wise L, Waller P. Development of a novel regulatory pharmacovigilance prioritisation system: an evaluation of its performance at the UK medicines and healthcare products regulatory agency. Drug Saf. 2013;36:1025–32.
25. Meyboom RH, Lindquist M, Egberts AC, Edwards IR. Signal selection and follow-up in pharmacovigilance. Drug Saf. 2002;25:459–65.
26. Ståhl M, Lindquist M, Edwards IR, Brown EG. Introducing triage logic as a new strategy for the detection of signals in the WHO Drug Monitoring Database. Pharmacoepidemiol Drug Saf. 2004;13:355–63.
27. Lindquist M. Use of triage strategies in the WHO signal-detection process. Drug Saf. 2007;30:635–7.
28. Caster O, Juhlin K, Watson S, Norén GN. Improved statistical signal detection in pharmacovigilance by combining multiple strength-of-evidence aspects in vigiRank: retrospective evaluation against emerging safety signals. Drug Saf. 2014;37:617–28.
29. Van Puijenbroek EP, Van Grootheest K, Diemont WL, Leufkens HG, Egberts AC. Determinants of signal selection in a spontaneous reporting system for adverse drug reactions. Br J Clin Pharmacol. 2001;52:579–86.
30. Levitan B, Yee CL, Russo L, Bayney R, Thomas AP, Klincewicz SL. A model for decision support in signal triage. Drug Saf. 2008;31:727–35.
31. Jamekornkul C, Chaisumritchoke ST. Developing a signal triage algorithm for Thai national adverse drug reaction database. Thai J Pharm Sci. 2016;40(3):153–7.
32. Salvo F, Raschi E, Moretti U, Chiarolanza A, Fourrier-Réglat A, Moore N, et al. Pharmacological prioritisation of signals of disproportionate reporting: proposal of an algorithm and pilot evaluation. Eur J Clin Pharmacol. 2014;70:617–25.
33. European Medicines Agency. Guideline on good pharmacovigilance practices (GVP): module IX—signal management. 2012. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/06/WC500129138.pdf. Accessed 18 Aug 2017.
34. United States Food and Drug Administration. Classifying significant postmarketing drug safety issues. Fed Regist. 2012. https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM295211.pdf. Accessed 16 Nov 2017.
35. Sottosanti L, Ferrazin F. Italian pharmacovigilance system. Manns Pharmacovigil. 3rd ed. Chichester: Wiley; 2014.
36. Coloma PM, Schuemie MJ, Trifirò G, Furlong L, van Mulligen E, Bauer-Mehren A, et al. Drug-induced acute myocardial infarction: identifying "prime suspects" from electronic healthcare records-based surveillance system. PLoS ONE 2013;8:e72148. https://doi.org/10.1371/journal.pone.0072148.
37. Bezin J, Duong M, Lassalle R, Droz C, Pariente A, Blin P, et al. The national healthcare system claims databases in France, SNIIRAM and EGB: Powerful tools for pharmacoepidemiology. Pharmacoepidemiol. Drug Saf. 2017. http://doi.wiley.com/10.1002/pds.4233. Accessed 2 Jun 2017.
38. European Medicines Agency. Inclusion/exclusion criteria for the "Important medical events" list. 2017. http://www.ema.europa.eu/docs/en_GB/document_library/Other/2016/08/WC500212100.pdf. Accessed 16 Nov 2017.
39. Tsiropoulos I, Andersen M, Hallas J. Adverse events with use of antiepileptic drugs: a prescription and event symmetry analysis. Pharmacoepidemiol Drug Saf. 2009;18:483–91.
40. Wahab IA, Pratt NL, Wiese MD, Kalisch LM, Roughead EE. The validity of sequence symmetry analysis (SSA) for adverse drug reaction signal detection. Pharmacoepidemiol Drug Saf. 2013;22:496–502.
41. Arnaud M, Bégaud B, Thurin N, Moore N, Pariente A, Salvo F. Methods for safety signal detection in healthcare databases: a literature review. Expert Opin Drug Saf. 2017;16:721–32.
42. Petri H, De Vet HCW, Naus J, Urquhart J. Prescription sequence analysis: a new and fast method for assessing certain adverse reactions of prescription drugs in large populations. Stat Med. 1988;7:1171–5.
43. Hallas J. Evidence of depression provoked by cardiovascular medication: a prescription sequence symmetry analysis. Epidemiology. 1996;7:478–84.
44. Coloma PM, Trifirò G, Schuemie MJ, Gini R, Herings R, Hippisley-Cox J, et al. Electronic healthcare databases for active drug safety surveillance: is there enough leverage? Pharmacoepidemiol Drug Saf. 2012;21:611–21.
45. Hanczar B, Hua J, Sima C, Weinstein J, Bittner M, Dougherty ER. Small-sample precision of ROC-related estimates. Bioinformatics. 2010;26:822–30.
46. Lerch M, Nowicki P, Manlik K, Wirsching G. Statistical signal detection as a routine pharmacovigilance practice: effects of periodicity and resignalling criteria on quality and workload. Drug Saf. 2015;38:1219–31.
47. Avillach P, Coloma PM, Gini R, Schuemie M, Mougin F, Dufour J-C, et al. Harmonization process for the identification of medical events in eight European healthcare databases: the experience from the EU-ADR project. J Am Med Inform Assoc. 2012;20:184–92.
48. Cutrona SL, Toh S, Iyer A, Foy S, Cavagnaro E, Forrow S, et al. Design for validation of acute myocardial infarction cases in Mini-Sentinel. Pharmacoepidemiol Drug Saf. 2012;21:274–81.
49. Trifiro G, Pariente A, Coloma PM, Kors JA, Polimeni G, Miremont-Salamé G, et al. Data mining on electronic health record databases for signal detection in pharmacovigilance: which events to monitor? Pharmacoepidemiol Drug Saf. 2009;18:1176–84.
50. Arnaud M, Bezin J, Bégaud B, Pariente A, Salvo F. Trends in the incidence of use of noninsulin glucose-lowering drugs between 2006 and 2013 in France. Fundam Clin Pharmacol. 2017;31(6):663–75.
51. Hauben M, Aronson JK. Defining "signal" and its subtypes in pharmacovigilance based on a systematic review of previous definitions. Drug Saf. 2009;32:99–110.