An automated system combining safety signal detection and prioritization adapted to healthcare databases

Mickael ARNAUD

27 February 2018

Disclosure

- I conducted this work during my PhD for which I was granted by the French Ministry for Higher Education and Research
- This work was also part of the Bordeaux DRUGS-SAFE platform funded by the French Medicines Agency (ANSM)
  - ANSM had no role concerning collection, analysis, and interpretation of the data; writing study reports and scientific articles
  - The results publication represents the views of the authors and does not necessarily represent the opinion of ANSM
- I have no conflict of interest directly relevant to the content presented

Introduction

- Actual post-marketing drug safety monitoring systems are based on spontaneous reporting data
  - They are limited in the identification of adverse events not evocative of a drug causation
    - E.g. myocardial infarction (rofecoxib, rosiglitazone)
- New tools using data from healthcare databases have recently been developed
  - Focused essentially on signal detection or signal validation
  - Signal prioritization have been neglected

Introduction

- Signal prioritization is the step following the detection
  - Detection needs to be sensitive to avoid missing a real drug safety issue
  - Detection leads thus to identify thousands of safety signals

Plausible and unknown	Plausible but already known	Implausible
+	++	+++

- **Signal prioritization is thus crucial to help stakeholders handle these thousands of detected safety signals to make the relevant regulatory actions**

Objectives

- To develop and to assess an automated system combining safety signal detection and prioritization
  - Adapted to healthcare databases
  - Adapted for the surveillance of drugs used in chronic diseases
- Plan
  - Study #1 – Development of the system: detection
  - Study #2 – Development of the system: prioritization
  - Study #3 – Assessment of the performance of the system
  - Study #4 – Concordance of prioritization: system *vs.* stakeholders

Study #1

Development of the system: detection

Study #1 – Introduction

- Background
  - Numerous methods for safety signal detection developed
  - Works conducted notably by International initiatives (ex. OMOP)
    - Systematic investigation of the best configuration
    - Assessment on empirical and simulated data
  - No consensus about the method to use

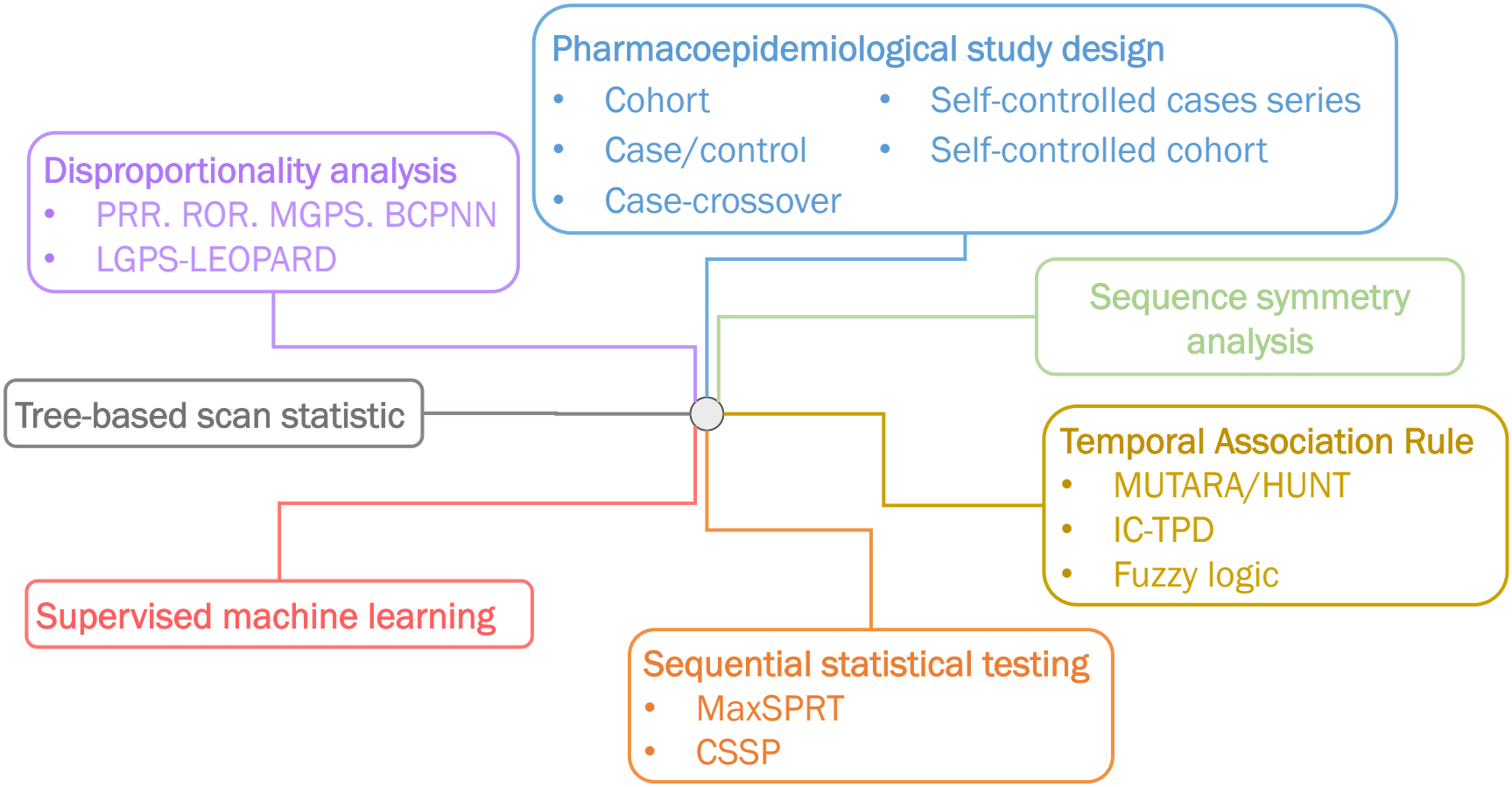
- Objective
  - To identify the most appropriate method for the safety signal detection adapted to healthcare databases

Study #1 – Methods

- Literature review
  - Search strategy
    - PubMed and Scopus
    - Select articles related to International initiatives (e.g. OMOP)
    - Systematic screening of references
      - « snowballing » and « reverse snowballing » approaches
- Selection of the most appropriate method
  - Pre-selection based on statistical performance (AUC)
  - Final selection based on pragmatic criteria
    - Dedicated for screening any drug/AE pair
    - Understanding of its principle
    - Providing of a risk estimation

Study #1 – Results

- 15 methods classified in 7 groups



Study #1 – Results

PRR. ROR. MGPS. BCPNN	AUC = 0.63 / 0.53 / 0.60 / 0.55
LGPS-LEOPARD	AUC = 0.58 / 0.59
Cohort	AUC = 0.68 / 0.54 / 0.61
Case/control	AUC = 0.61 / 0.59 / 0.61
Case crossover	AUC = 0.61
Self-controlled case series	AUC = 0.57 / 0.71 / 0.67
Self-controlled cohort	AUC = 0.53 / 0.81 / 0.77
Sequence symmetry analysis	Se = 0.67, Sp = 0.93, VPP = 0.77, VPN = 0.87
MUTARA/HUNT	AUC = 0.60
IC-TPD	AUC = 0.65 / 0.75 / 0.67 / 0.57
Fuzzy logic	–
MasXPRT	AUC = 0.23
CSSP	AUC = 0.38
Supervised machine learning	AUC = 0.81 / 0.86
Tree-based scan statistic	–

Study #1 – Results

PRR. ROR. MGPS. BCPNN	AUC = 0.63 / 0.53 / 0.60 / 0.55
LGPS-LEOPARD	AUC = 0.58 / 0.59
Cohort	AUC = 0.68 / 0.54 / 0.61
Case/control	AUC = 0.61 / 0.59 / 0.61
Case crossover	AUC = 0.61
Self-controlled case series	**AUC = 0.57 / 0.71 / 0.67**
Self-controlled cohort	**AUC = 0.53 / 0.81 / 0.77**
Sequence symmetry analysis	**Se = 0.67, Sp = 0.93, VPP = 0.77, VPN = 0.87**
MUTARA/HUNT	AUC = 0.60
IC-TPD	AUC = 0.65 / 0.75 / 0.67 / 0.57
Fuzzy logic	–
MasXPRT	AUC = 0.23
CSSP	AUC = 0.38
Supervised machine learning	**AUC = 0.81 / 0.86**
Tree-based scan statistic	–

Study #1 – Results

Pragmatic criteria for distinguishing the methods for signal detection	Self-controlled designs*	Sequence symmetry analysis	Supervised machine learning
1. Dedicated for screening any drug/AE pair	–	+	+
2. Understanding of its principle	+	+	–
3. Providing of a risk estimation	+	+	–

**Include self-controlled case series and self-controlled cohort*

Study #1 – Results

Pragmatic criteria for distinguishing the methods for signal detection	Self-controlled designs*	Sequence symmetry analysis	Supervised machine learning
1. Dedicated for screening any drug/AE pair	–	+	+
2. Understanding of its principle	+	+	–
3. Providing of a risk estimation	+	+	–

**Include self-controlled case series and self-controlled cohort*

Study #1 – Conclusion

- The sequence symmetry analysis (SSA)
  - Considers, for each drug/event pair and for each person,
    - 1st drug dispensing and 1st event occurrence
    - Only if they occur after a given run-in period (e.g. 12 months)
  - Computes the ratio of the number of persons, observed during the study period, experiencing
    - Sequence Drug→Event *vs.* sequence Event→Drug
    - Considering a given time period (e.g. 12 months)
  - Adjusts this ratio on trends of drug use and outcome occurrence
    - Adjusted sequence ratio (ASR)
    - Interpretation: equivalent to an incidence rate ratio

Study #1 – Conclusion

- Strengths of the SSA
  - Principle easy to understand and provides risk estimates
  - Dedicated for screening any drug/event pair in longitudinal data
  - Controlling for numerous confounding factors
    - Time-constant confounding factors
    - Biases related to the trends of drug use or event occurrence
- Limitations of the SSA
  - Sensitive to protopathic and indication biases
    - Concerned every method for signal detection in healthcare data
    - Biases to control +++ in the signal prioritization process
  - Sensitive to events affecting the probability to receive the drug
    - But, this does not impact the sensitivity of detection

Study #2

Development of the system: prioritization

Study #2 – Prioritization

- Background
  - Complex and multifactorial process
    - Clinical
    - Epidemiological
    - Pharmacological
    - Regulatory
  - Essential for handling the thousands of the detected signals
  - Pharmacovigilance systems have implemented
    - Concepts or frameworks for standardizing the prioritization process
    - Automated algorithms for the signal prioritization
      - Adapted to spontaneous reporting data
- Objective
  - To develop an automated algorithm for the signal prioritization

Study #2 – Methods

- Literature review
  - Search strategy
    - PubMed and Scopus
    - Articles related to “signal” AND “prioritization”, “filter”, or “triage”
    - Systematic screening of references
      - « snowballing » and « reverse snowballing » approaches
- Selection of the strategies
  - Essential criteria to use
    - Criteria adaptable to healthcare data
    - Supplementary criteria dedicated to healthcare data
- Selection of (semi-) automated algorithms
  - Combination of the criteria
  - Presentation of the results

Study #2 – Results

- 14 strategies for signal prioritization identified
- Inspired from the 'SNIP' concept of Waller *et al.* (1999)*
  - **S**trength of the signal
  - **N**ovelty of the signal
  - clinical **Importance** (or **I**mpact) of the signal
  - **P**otential for preventive measures
- Consensus on
  - « Strength », « Novelty », « Impact » are essentials
  - « Prevention » abandoned or replaced by other regulatory considerations

**Waller PC, et al. Pharmacoepidemiol Drug Saf 1999; 8: 535–552.*

Study #2 – Results

- Main criteria retrieved

Strength of the signal	Novelty of the signal	Impact of the signal	Others criteria
Measure of association	Unknown signal	Seriousness of cases reported	Measures for prevention
Biological plausibility	Recent drug	Number of cases reported	Therapeutic alternative
Number of sources indicating the signal	Increasing of the risk estimate	Severity of the adverse event	Interest of stakeholders for the signal
Information in favor of a causal link *(e.g. positive rechallenge)*	Increasing of reporting rate	Frail populations *(e.g. children, pregnant women)*	Risk perception in the population
		Prevalence of drug use	
		Estimated number of cases in excess	

Study #2 – Results

- Main criteria retrieved
  - Adaptable to healthcare data

Strength of the signal	Novelty of the signal	Impact of the signal	Others criteria
Measure of association	Unknown signal	Seriousness of cases reported	Measures for prevention
Biological plausibility	Recent drug	Number of cases reported	Therapeutic alternative
Number of sources indicating the signal	Increasing of the risk estimate	Severity of the adverse event	Interest of stakeholders for the signal
Information in favor of a causal link *(e.g. positive rechallenge)*	Increasing of reporting rate	Frail populations *(e.g. children, pregnant women)*	Risk perception in the population
		Prevalence of drug use	
		Estimated number of cases in excess	

Study #2 – Results

- Criteria specific to healthcare data
  - Criteria related to the use of drugs
    - Incidence of use
    - Trends of incidence of use

Other

  - Criteria related to the limitations of the signal detection
    - Event not related to drug indications (control for protopathic bias)
  - Criteria related to the risk estimates
    - Lower limit of the 95% confidence interval (95%CI)
    - Precision of the risk estimate

Strength

  - Criteria related to economical aspect
    - Cost of the event for the insurances

Impact

Study #2 – Results

- 8 (semi-) automated algorithms for signal prioritization

	Impact Analysis*	RPPS*	UMC triage*	VigiRank *	PS-SP	Lab MADA	Thai MADA*	EU-ADR
Criteria based on								
Strength of the signal	x	x	x	x	x	x	x	x
Novelty of the signal			x	x		x	x	x
Impact of the signal	x	x	x	x		x	x	x
Other		x				x		
Criteria processing								
Categorization	x	x	x	x		x	x	x
Normalization					x			
Output presentation								
Grouping by level of priority	x	x	x			x		
Ranking by decreasing order of value of priority				x	x		x	x

**developed by regulatory authorities*

Study #2 – Results

- 8 (semi-) automated algorithms for signal prioritization

	Impact Analysis*	RPPS*	UMC triage*	VigiRank *	PS-SP	Lab MADA	Thai MADA*	EU-ADR
Criteria based on								
Strength of the signal	x	x	x	x	x	x	x	x
Novelty of the signal			x	x		x	x	x
Impact of the signal	x	x	x	x		x	x	x
Other		x				x		
Criteria processing								
Categorization	x	x	x	x		x	x	x
Normalization					x			
Output presentation								
Grouping by level of priority	x	x	x			x		
Ranking by decreasing order of value of priority				x	x		x	x

**developed by regulatory authorities*

Study #2 – Conclusion

- Criteria to consider
  - Strength of the signal
  - Novelty of the signal
  - Impact of the signal
  - Other : patterns of drug use
- Categorization of the criteria
  - Allows to combine criteria of different nature
  - Weighting of the criteria
- Signals grouped by levels of priority
  - Appear more consistent with stakeholders expectations

Study #2 – Conclusion

- Longitudinal–SNIP (L–SNIP) algorithm
  - 14 criteria: categorized and weighted from 1 to 4
  - L–SNIP score: weighted sum of the criteria
  - Signal prioritization results
    - "prioritized" if L–SNIP score in top 10%
    - "not prioritized" otherwise

Strength of the signal		Novelty of the signal		Impact of the signal		Patterns of drug use	
Risk estimate	3	Signal not mention in SPC	4	Potential number of attributable cases	2	Event not related to drug indications	4
Lower limit of the 95%CI	4	Drug seniority	2	Cost of hosp. for the event	2	Drug use in vulnerable pop. (1): children	3
Precision of risk estimate	2	Increasing in risk over time	1			Drug use in vulnerable pop. (1): childbearing women	3
						Prevalence of drug use	2
						Incidence of drug use	2
						Increasing of the incidence of drug use over time	1

Study #3

Assessment of the performance of the developed system

Study #3 – Performance

- Objective
  - To assess the performance of the system combining safety signal detection and prioritization from healthcare databases
- Pilot study applied on Type 2 diabetes
  - Frequent chronic disease in the population
    - New risk identified = major impact in terms of public health
  - Treatments have greatly changed for a decade
    - New drugs marketed in 2008
    - Withdrawal or restriction of use of glitazones for safety reasons
      - Bladder cancer, heart failure

Study #3 – Methods

- Data source
  - *Echantillon Généraliste des Bénéficiaires* (EGB) claims database
    - 1/97th sample of the population covered by the French national health insurance system
    - Representative in terms of age, sex, geographic location, and care consumption
  - EGB includes comprehensive and anonymous data
    - Outpatient drug dispensing (coded with ATC classification)
    - Hospitalization diagnoses (coded with ICD-10 classification)
- Study population
  - Persons included in EGB at least 1 year between 2005 and 2015

Study #3 – Methods

- Drug exposure definition
  - Dispensing as surrogate for drug exposure
  - Noninsulin glucose-lowering drugs (NIGLDs) identified at ATC level 5
  - Selection of the 1st dispensing if it occurred ≥1 year of follow-up
- Event definition
  - Hospitalization diagnoses as surrogates for adverse events
  - Selection of ICD-10 codes corresponding to MedDRA® important medical events (IME)
    - Alignment of ICD-10 codes with those included in the MedDRA® IME terms list using the Unified Medical Language System tool
  - Selection of the 1st occurrence if it occurred ≥1 year of follow-up

Study #3 – Methods

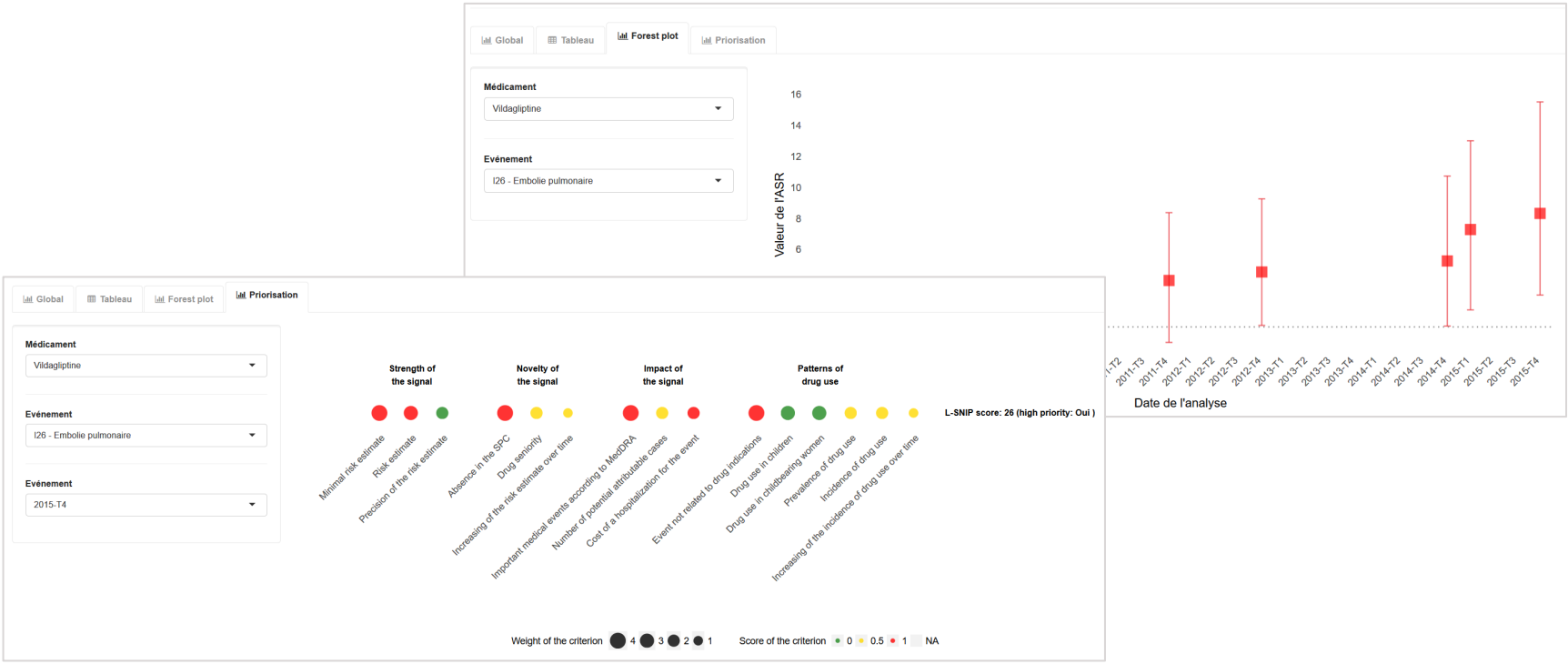
- Signal detection
  - Sequence Symmetry Analysis
  - Quarter analyses between 2008-2015 for each NIGLD/IME pair
    - ≥1 exposed case observed during the quarter of analysis
    - ≥3 exposed cases observed in the population
  - 95%CI computed using the bootstrap method (500 replications)
  - Signal detected if the lower limit of the 95%CI > 1
- Signal prioritization
  - L–SNIP algorithm
    - Signal prioritized if L–SNIP score in the top 10%
    - Signal not considered as priority otherwise

Study #3 – Methods

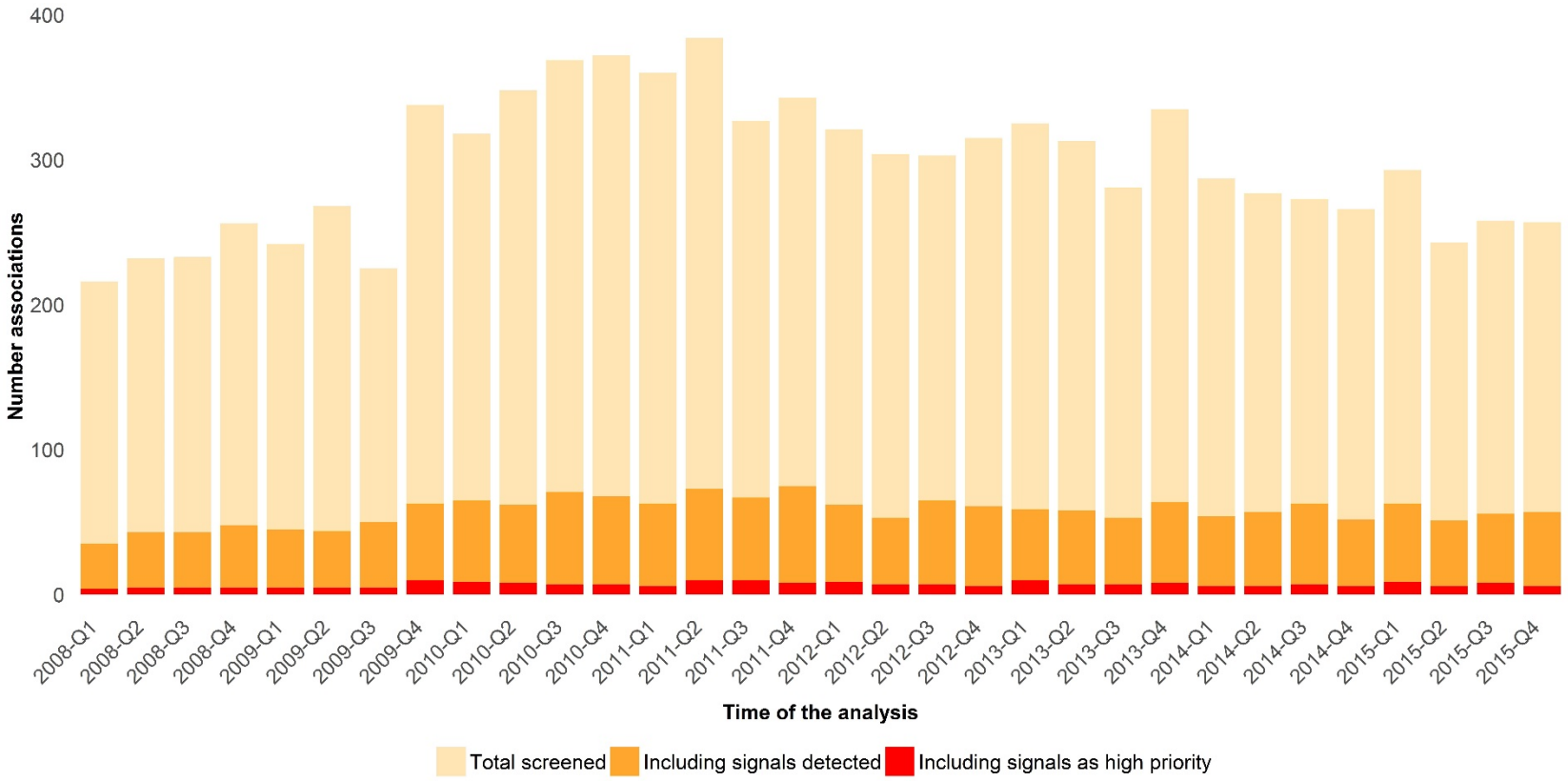
- Reference dataset
  - NIGLD/IME pairs with sufficient power to detect a relative risk of 2 based on the drug and event prevalence in the EGB
    - Positive controls : associations listed in SPCs
    - Negative controls
      - Random selection among all the other associations
      - Ratio of 3 negative controls for 1 positive control
- Performance assessment
  - Se, Sp, PPV, and NPV
  - Performance for signal prioritization
    - Selection of positive controls including in SPCs after 2008
    - Ability of the system to identify them before the mention in SPCs

Study #3 – Methods

- Analysis of prioritized signals
  - Development of a R Shiny App to analyze the detected and/or prioritized signals



Study #3 – Results



Study #3 – Results

- Performance for signal detection
  - Reference set: 15 positive controls and 45 negative controls

	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%	

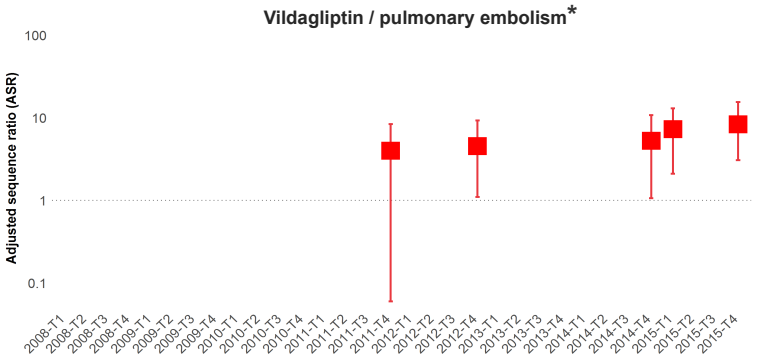
Study #3 – Results

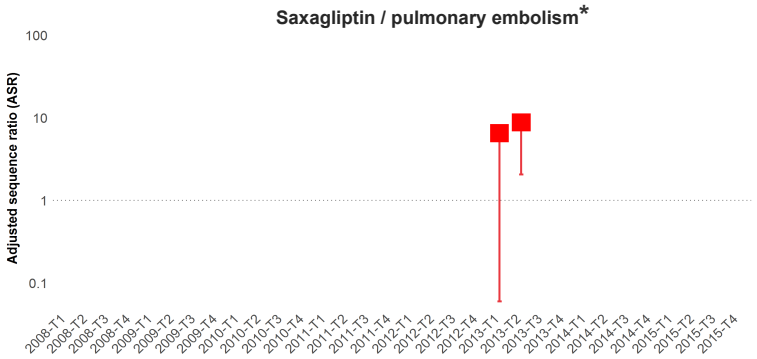
- Performance for signal prioritization
  - Positive controls limited to 3 associations
  - Added value of the prioritization of the signals detected
    - Se, Sp, and NPV similar
    - 7-fold increase of the PPV with the prioritization

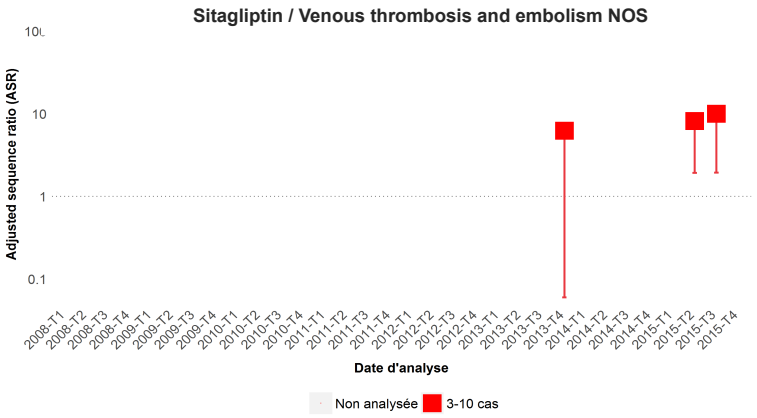
	Se	Sp	PPV	NPV
Detection + L–SNIP	33%	100%	100%	96%
Detection	33%	100%	14%	95%

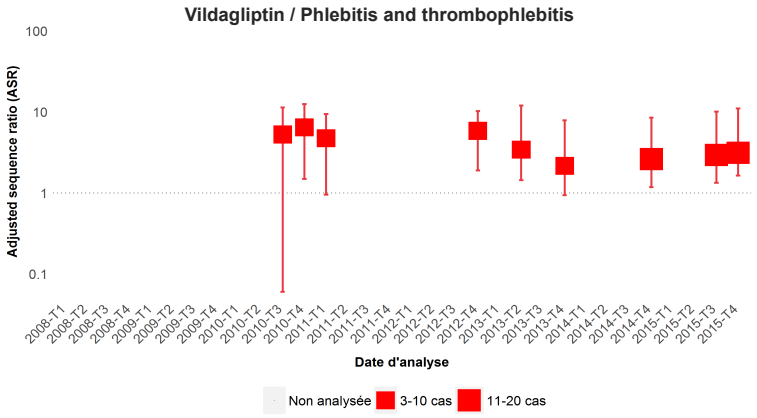
Study #3 – Results

- A relevant safety signal identified with the ***gliptins*** and the risk of ***venous thromboembolic events***









**Association prioritized at least once*

Study #3 – Conclusion

- Performance for signal detection promising
  - Similar to that observed in other studies using SSA
- The use of the L–SNIP algorithm for the signal prioritization
  - Makes the identification of relevant signals easier
  - Performance needs to be confirmed in a larger reference set
    - Based on only 3 positive controls
- The developed system highlighted a new drug safety issue
  - Gliptins and risk of venous thromboembolism
  - Potentially major impact in terms of public health

Study #4:

Concordance of the signal prioritization: system *vs.* stakeholders

Study #4 – Concordance

- Background
  - The developed system needs to match with the stakeholders' point of view
    - In the perspective of a future use for the routine surveillance of the safety of drugs
  - The prioritization is the crucial aspect
  - The L–SNIP algorithm could be subject to discussion
    - Criteria retained, weighting
- Objective
  - To assess the concordance of the signal prioritization from the L–SNIP algorithm with that of the stakeholders

Study #4 – Methods

- Target population
  - Persons with decision-making power
    - Health professionals with expertise mission for the Public Agencies
    - Managers in Public Agencies
    - Managers in pharmaceutical companies
  - Persons with ability to influence the decision-making
    - Managers in patient organizations
    - Journalists

Study #4 – Methods

- Questionnaire-based survey
  - Available online
    - To guarantee anonymity
    - Available during 3 months
    - E-mail reminder every 2 weeks
  - Data collection
    - Social information (position, diploma, *etc.*)
    - Appraisal of
      - The use of an automated prioritization as decision support
      - The criteria proposed for the signal prioritization
    - Signal prioritization exercise considering 10 fictive detected signals

Study #4 – Methods

- Prioritization exercise
  - 10 fictive signals related to long-term treatments
  - Based on some criteria included in the L–SNIP algorithm
    - Risk estimate
    - Prevalence and Incidence of drug use
    - Year of marketing
    - Mean cost of hospitalization for the adverse event
    - Number of potential attributable cases
    - Knowledge of the association
  - Prioritization according 3 levels of priority: *High*, *Moderate*, *Weak*

Study #4 – Methods

- Assessment of concordance of prioritization
  - Prioritization from the surveyed
    - Selection of the modal response among those collected
  - Prioritization from L–SNIP algorithm
    - L–SNIP scores compared to those obtained in Study #3
    - Classification modified to match with that proposed in the exercise
      - *High* if L–SNIP score in Top 1-10%
      - *Moderate* if L–SNIP score in Top 11-50%
      - *Weak* if L–SNIP score in Top 51-100%
  - Concordance measured with Kendall's concordance coefficient τ

Study #4 – Results

- 32 respondents among ~150 persons solicited

		N (%)
Actual or former organization	Hospital	26 (81.3)
	Public Agencies	8 (25.0)
	Pharmaceutical companies	4 (12.5)
	Media	5 (15.6)
	Patients organizations	1 (3.1)
Main diploma	M.D.	14 (43.8)
	Pharm.D.	9 (28.1)
Specific skills	Pharmacology	16 (50.0)
	Epidemiology / pharmacoepidemiology	9 (28.1)

Study #4 – Results

- 20 (62.5%) favorable to an automated signal prioritization
- Appraisal of criteria proposed for signal prioritization

Criteria	No, not at all	Rather no	Rather yes	Yes definitely
Risk estimate	1	0	17	14
Number of potential attributable cases	1	1	18	12
Prevalence of drug use	1	6	14	11
Incidence of drug use	1	4	17	10
Knowledge of the association	1	3	20	8
Year of marketing	5	10	12	5
Mean cost of hospitalization for the event	11	14	7	0

Study #4 – Results

- Concordance of signal prioritization

Fictive case	Drug indication	Adverse event	Prioritization stakeholders			Prioritization L–SNIP
			High	Moderate	Weak	
1	Type 2 diabetes	Cognitive disorders	5	22	5	Weak
2	Type 2 diabetes	Crohn's disease	17	11	4	Moderate
3	Type 2 diabetes	Crohn's disease	17	10	5	Moderate
4	Prevention of VTE	Orthostatic hypotension	4	15	13	Weak
5	Prevention of AMI	Migraine	8	13	11	Moderate
6	Serious sleep disorder	Femoral neck fracture	15	13	4	High
7	Epilepsy	Ventricular tachycardia	14	15	3	Moderate
8	Schizophrenia	Anorexia	9	15	8	Weak
9	Oral contraceptive	Iron deficiency anemia	3	18	11	Moderate
10	Asthma	Sleep disorder	3	15	14	Weak

Study #4 – Results

- Concordance of signal prioritization

Fictive case	Drug indication	Adverse event	Prioritization stakeholders			Prioritization L–SNIP
			High	Moderate	Weak	
1	Type 2 diabetes	Cognitive disorders	5	22	5	Weak
		ase	17	11	4	Moderate
		ase	17	10	5	Moderate
			4	15	13	Weak
5	Prevention of AMI	Migraine	8	13	11	Moderate
6	Serious sleep disorder	Femoral neck fracture	15	13	4	High
7	Epilepsy	Ventricular tachycardia	14	15	3	Moderate
8	Schizophrenia	Anorexia	9	15	8	Weak
9	Oral contraceptive	Iron deficiency anemia	3	18	11	Moderate
10	Asthma	Sleep disorder	3	15	14	Weak

Kendall's concordance coefficient, $\tau = 59\%$

Study #4 – Conclusion

- L–SNIP algorithm could have a role of decision support among stakeholders
  - Signal prioritization globally concordant with that of the stakeholders
  - Criteria included in the L–SNIP algorithm judged favorably
    - Excepted for the cost of hospitalization for the event
      - Other medico-economic criterion to consider?
- Main limitation
  - Small sample with potential non-representativeness

Conclusion & perspectives

Conclusion

- The developed automated system of safety signal detection and prioritization in healthcare databases was based on the best evidence from the scientific literature
- The assessment of the developed system
  - Good performance from reference dataset assessment
  - Able to highlight a relevant safety signal (need to confirm)
  - Prioritization concordant with that of stakeholders
- Main perspectives
  - Improving the system by reducing biases in signal detection
  - Adapting the system for the identification of long-term drug adverse events