Use of Stability Selection for signal detection in pharmacovigilance

Ismaïl Ahmed and Pascale Tubert-Bitter

Inserm UMR 1181 « Biostatistics, Biomathematics, Pharmacoepidemiology and Infectious Diseases » (B2PHI)

Institut Pasteur, UMR 1181, B2PHI

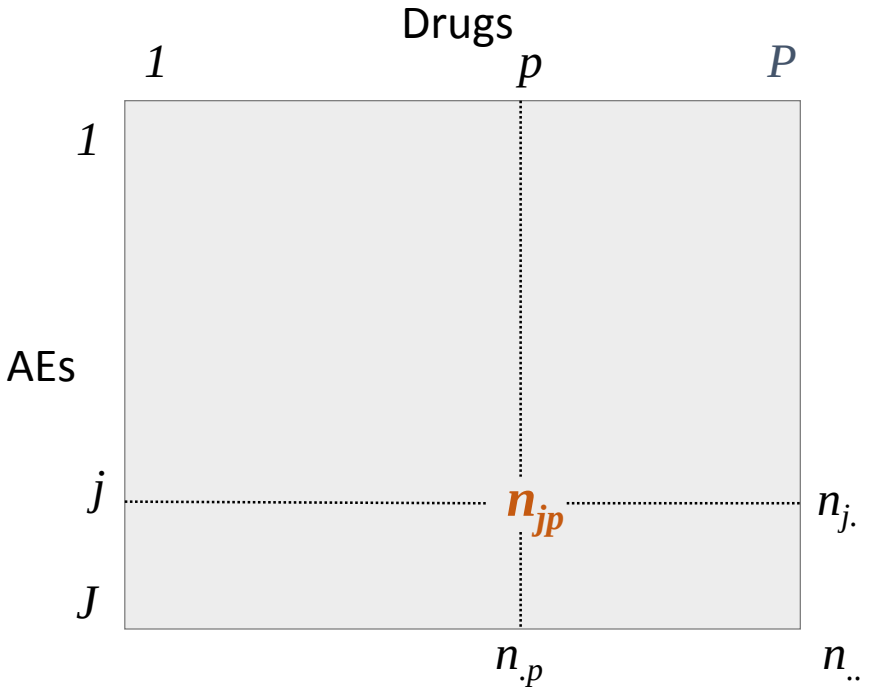
Univ. Versailles St Quentin, UMR 1181, B2PHI

Introduction (1)

- Pharmacovigilance systems
  - Detection of new adverse effects of licensed drugs
  - Based on spontaneous reports (SRs) of possible adverse drug reactions (ADRs)
  - Very large databases

- Automatic signal detection methods
  - Applied to aggregated data

    n_{jp}: Number of SRs involving AE *j* and drugs *p*
  - Called disproportionality methods



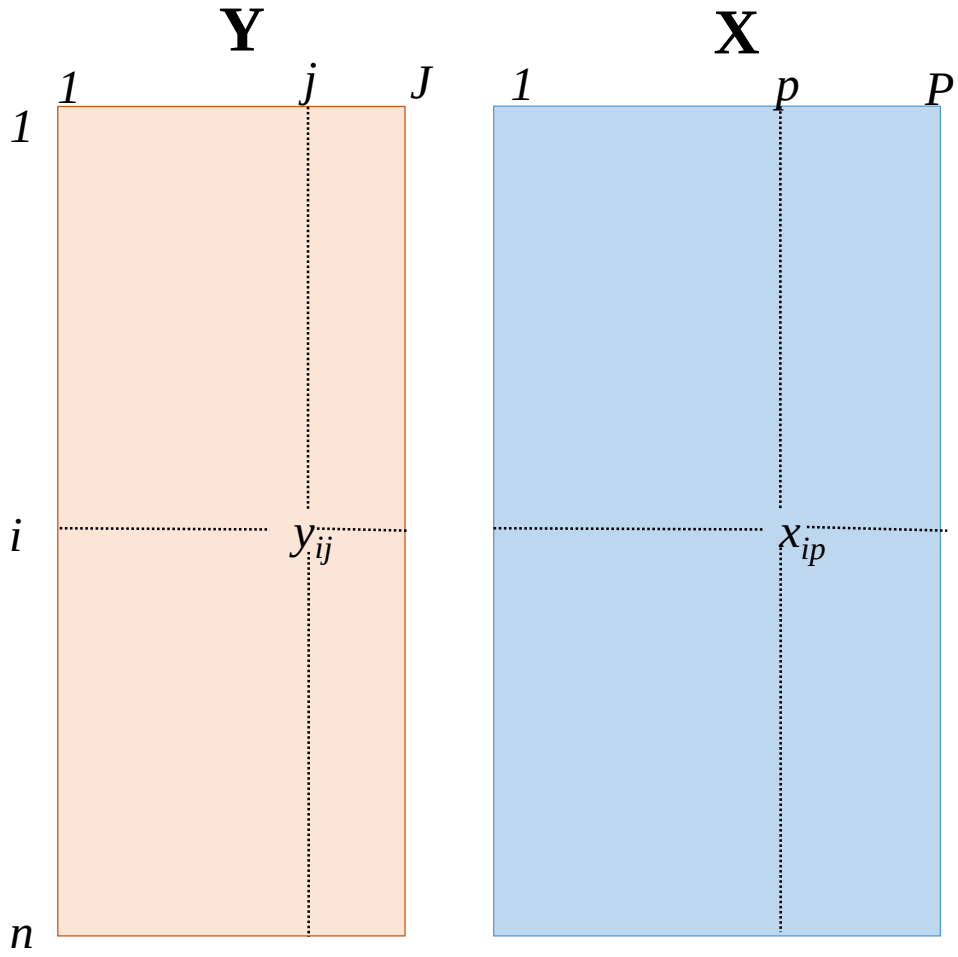
Introduction (2)

- More recently, the idea has been proposed to return to the analysis of individual spontaneous reports (Caster et al. 2010)

- Two matrices
  - Y: matrix of AEs - X: matrix of drugs
  - Y and X are binary
  - Y and X are also sparse

- Use of lasso logistic regression

- Much more computationally intensive
  - One lasso per AE (several thousands)
  - Very large databases



Lasso (Tibshirani 1996)

- Belongs to the family of penalized regressions
- For a given AE

$$\left(\hat{\beta}_0, \ldots, \hat{\beta}_P\right)^{\boldsymbol{\lambda}} = \operatorname{argmax}(\operatorname{logLik}(\beta_0, \ldots, \beta_P) - \boldsymbol{\lambda} \sum_{p=1}^{P} |\beta_p|)$$

- A major difficulty is to choose the constraint $\boldsymbol{\lambda}$
  - When the purpose is prediction, the k-fold cross-validation is a standard choice
  - In a variable selection context, fixing this parameter is much more challenging

Stability Selection (Meinshausen et al. 2010)

- General procedure combining **subsampling** with high dimensional selection algorithm such as the **lasso**

- Algorithm
  - Perform *B* logistic lasso on subsamples of size $\lfloor n/2 \rfloor$
  - For each variable calculate

$$\hat{\pi}_p^{\lambda} = \frac{1}{B} \sum_{b=1}^{B} \mathrm{I}\left\{\hat{\beta}_p^{\lambda,b} > 0\right\}$$

  - $\max(\hat{\pi}_p^{\lambda})$ over the grid of λ values

Objective

Propose an algorithm using the subsampling idea of Stability Selection adapted to the analysis of spontanous reporting data

Stability Selection: an alternative sampling

- Very sparse binary outcomes
- We propose an imbalanced sampling
  - Let's assume there are $\boldsymbol{n_1}$ cases (Set S_1) and $\boldsymbol{n_0}$ observations with no AE (Set S_0)

  1. Draw with replacement $\boldsymbol{n_1}$ observations from $\boldsymbol{S_1}$
  2. Draw without replacement $\boldsymbol{R}$ observations from $\boldsymbol{S_0}$

  In our experiment, we empirically fixed $R = \max(4P, 4n_1)$
- Computational and numerical advantages
  - Running the algorithm on much smaller subsamples
  - Having more 1 helps for the convergence of the logistic lasso

Stability Selection: Variable selection criterion

- For one subsample *b* calculate

$$\hat{\pi}_p^b = \frac{1}{\#H} \sum_{\eta \in H} \mathrm{I}\{\hat{\beta}_p^{\eta,b} > 0\}$$

  - η : number of regression parameters in the models
    H: Models with 1 to 50 parameters

- For each drug, we obtain an empirical distribution of $\hat{\pi}_p$ from the *B* subsamples
- Choose a quantile q_α for these empirical distributions
- Select a drug if $q_\alpha > 0$
- Simulations to help choosing q_α

Simulations (1)

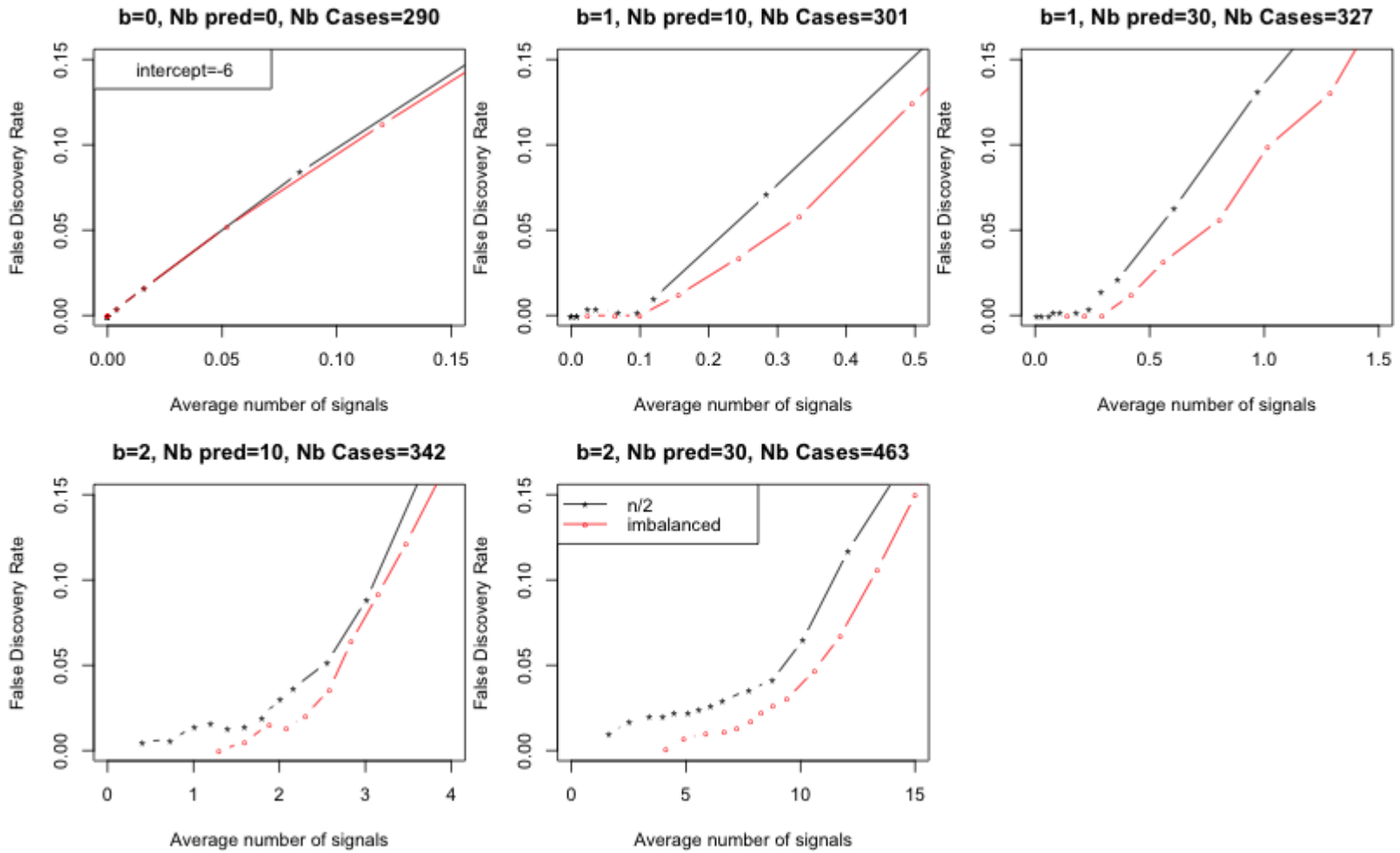
The purpose of the simulation study was twofold

1. Compare the proposed sampling strategy with $\lfloor n/2 \rfloor$
2. Help us deciding which quantile to choose for the drug selection

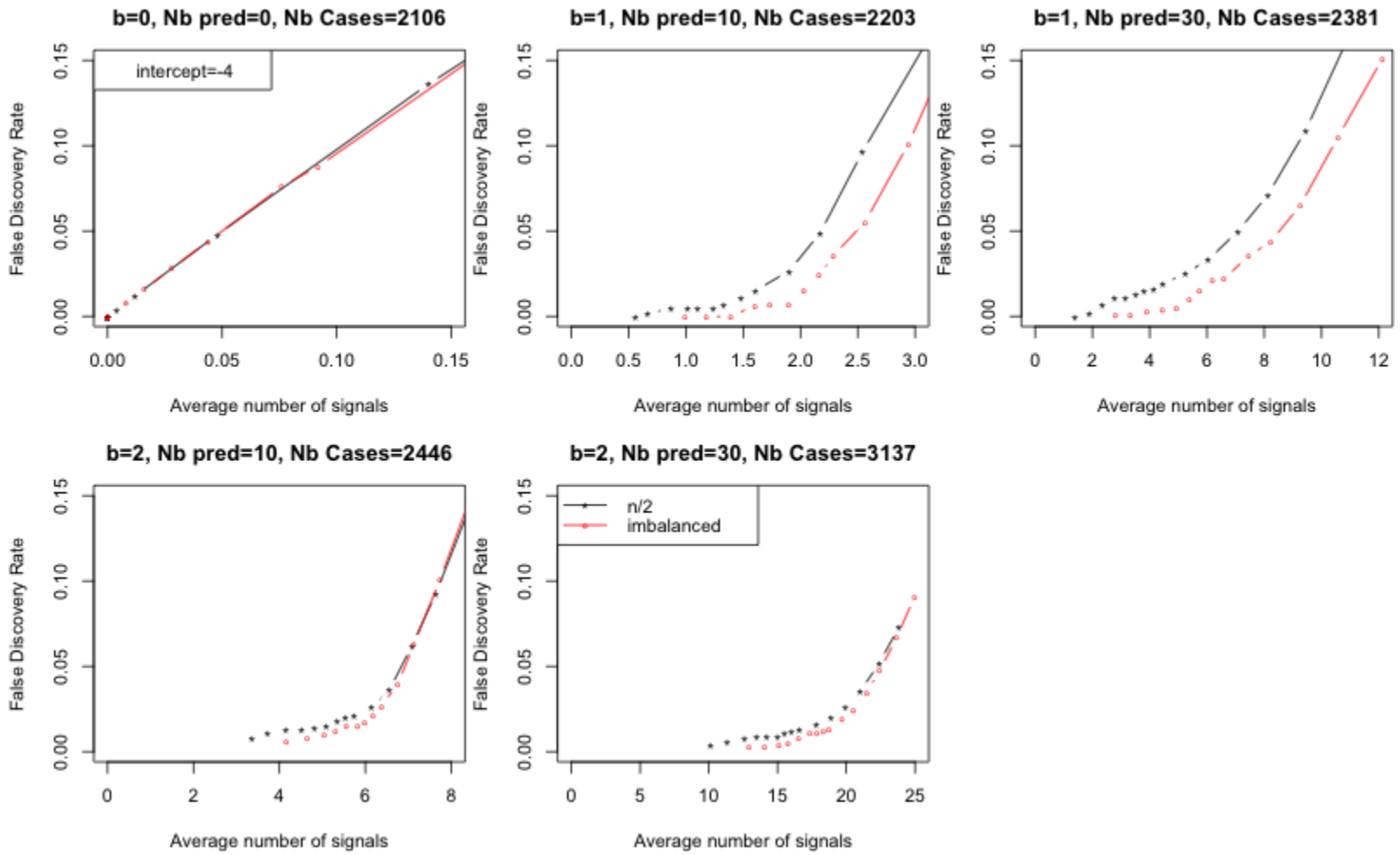
Simulations (2)

- AE generated according to a logistic regression model
  - $y_i \sim Bernouilli\ (\alpha_i)$
  - $\alpha_i = 1/(1 + \exp(\beta_0 + \boldsymbol{\beta}\mathbf{x_i}))$
- The X matrix is that of the French data (period 1995-2002)
  - 1111 drugs and 117160 observations
- The model depends on three parameters
  - The intercept β_0 (control the number of cases): -8, -6, -4
  - The number of true predictors: 0, 10 or 30 (the true predictors are randomly chosen for each dataset)
  - The value of the regression parameters for the true predictors $\boldsymbol{\beta}$: 1 or 2
- 250 datasets for each configuration

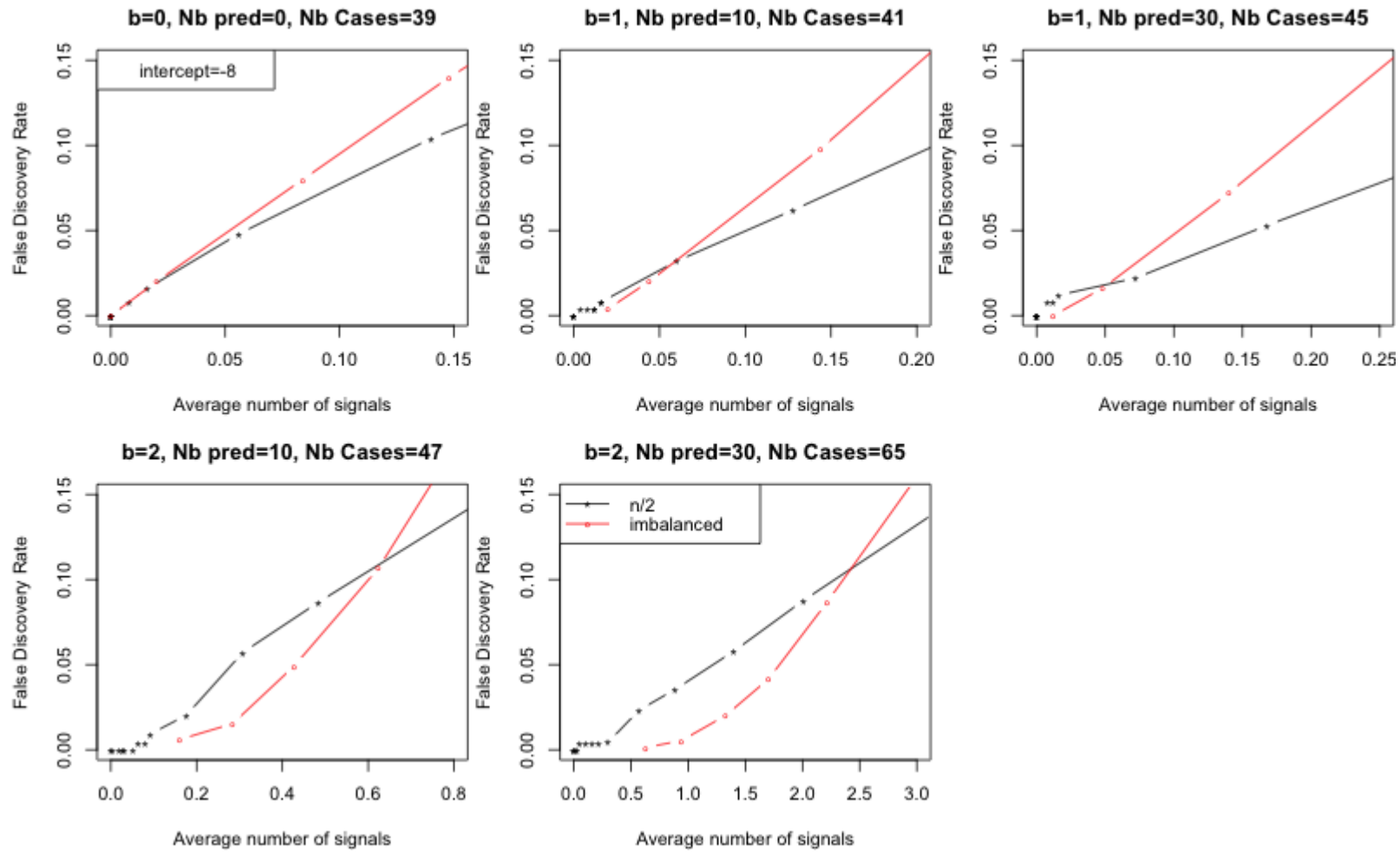
Simulation results (1): Common AEs - β_0=-6



Simulation results (2): Very common AEs - β_0=-4



Simulation results (3): Rare AEs - β_0=-8



Simulation results (4): choice of a quantile

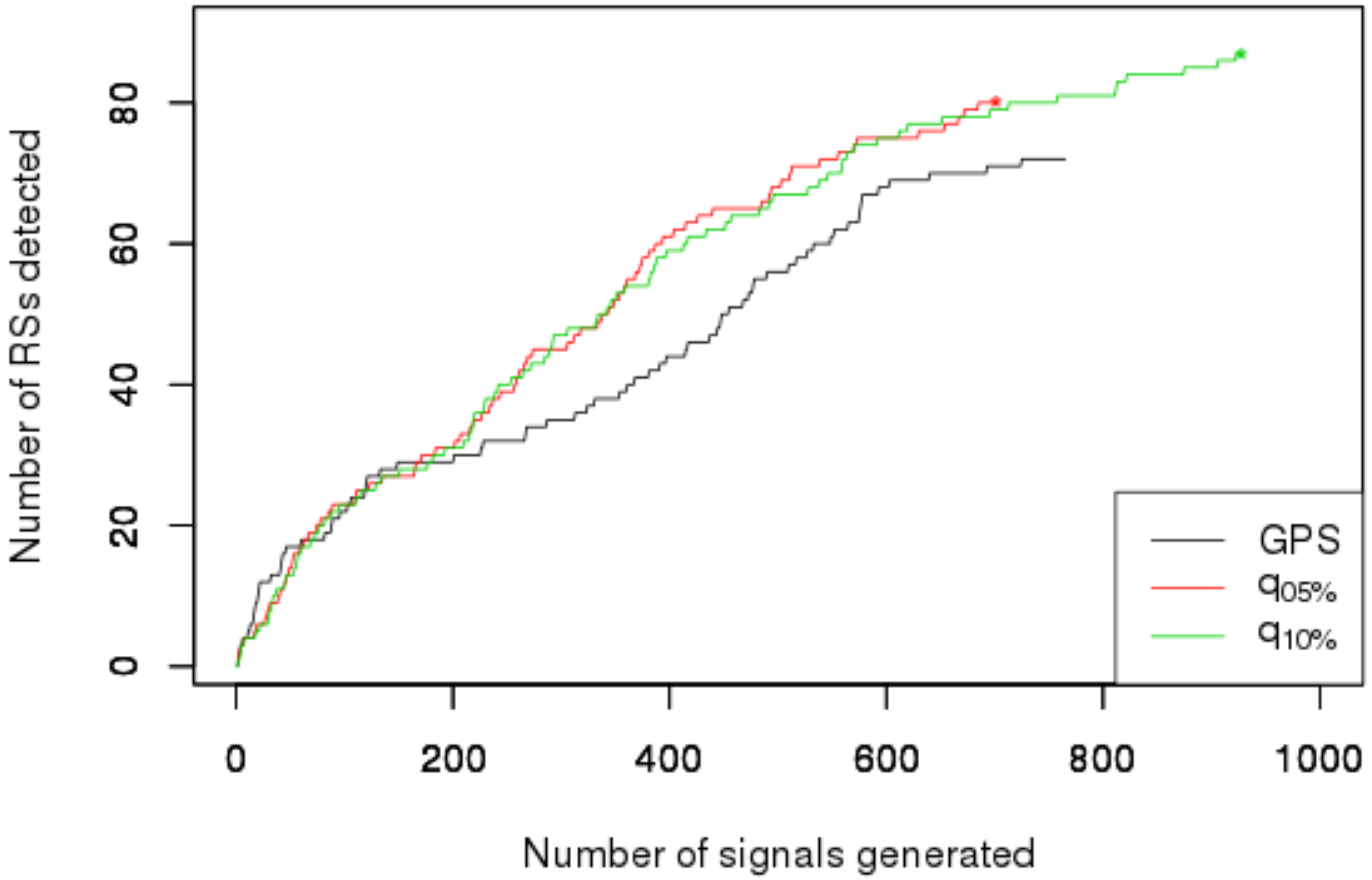
- FDR according to several quantiles
- For a given quantile the FDR decreases when
  - The AE is common
  - The number of true predictors increases
  - The strength of the association increases
- $q_{05\%}$keeps the FDR lower than 10%
- For more common AEs $q_{10\%}$ seems to be sensible choice

intercept	beta	Nb pred	Nb Cases	**q 0.01**	**q 0.05**	**q 0.1**	**q 0.15**
-8	0	0	39	0.000	0.080	0.300	0.776
-8	1	10	41	0.004	0.098	0.286	0.733
-8	1	30	45	0.000	0.072	0.279	0.669
-8	2	10	47	0.006	0.049	0.217	0.508
-8	2	30	65	0.001	0.020	0.087	0.224
-6	0	0	290	0.000	0.000	0.016	0.112
-6	1	10	301	0.000	0.000	0.034	0.125
-6	1	30	327	0.000	0.000	0.031	0.099
-6	2	10	342	0.000	0.016	0.020	0.064
-6	2	30	463	0.001	0.010	0.013	0.022
-4	0	0	2106	0.000	0.000	0.016	0.044
-4	1	10	2203	0.000	0.000	0.007	0.016
-4	1	30	2381	0.001	0.003	0.005	0.015
-4	2	10	2446	0.006	0.010	0.015	0.017
-4	2	30	3137	0.003	0.004	0.008	0.012

Empirical evaluation

- French data from the period 1995-2002
- Evaluation based on a set of 181 reference signals
  - Alerts launched by an expert committee from the French drug safety agency
  - 68 different AEs
- Comparison with a disproportionality method: Gamma Poisson Shrinker (Dumouchel 1999)

Results of the empirical evaluation



Conclusion and perspectives

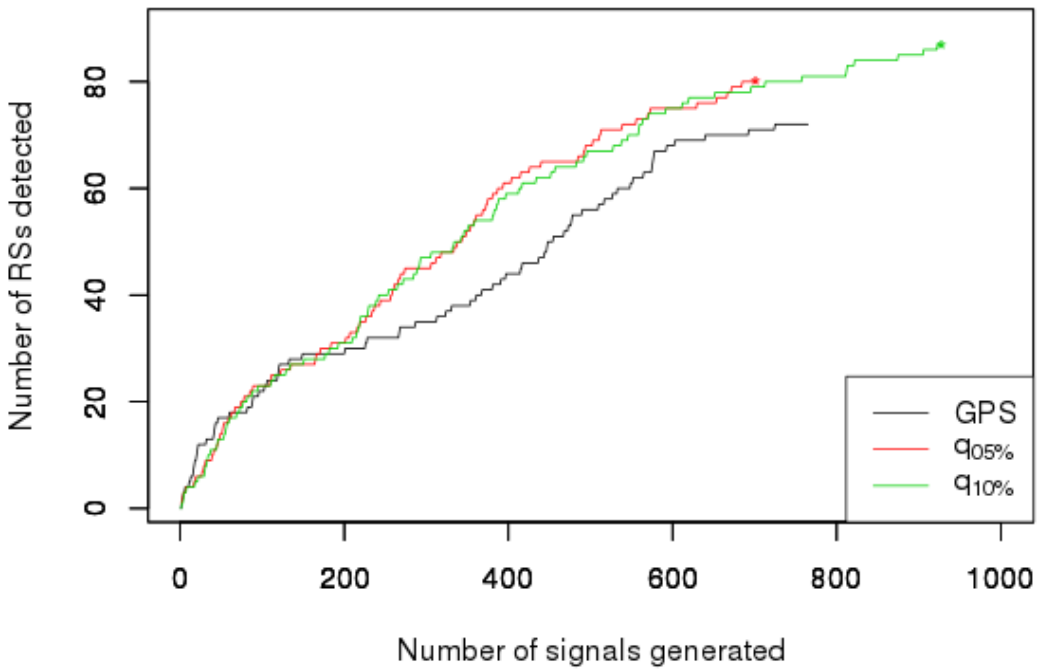
- We have proposed an extension of Stability Selection adapted to the analysis of pharmacovigilance data.
  - More powerful than the $\lfloor n/2 \rfloor$ sampling in most situations
  - Could be suited to other types of data with sparse outcomes
  - Performed better than GPS on an empirical study
  - Faster than the $\lfloor n/2 \rfloor$ sampling

- One limit lies in the selection strategy
  - Required simulations to help us deciding which quantile to choose
  - Ideally, it should be based on an estimate of an error criterion such as the FDR

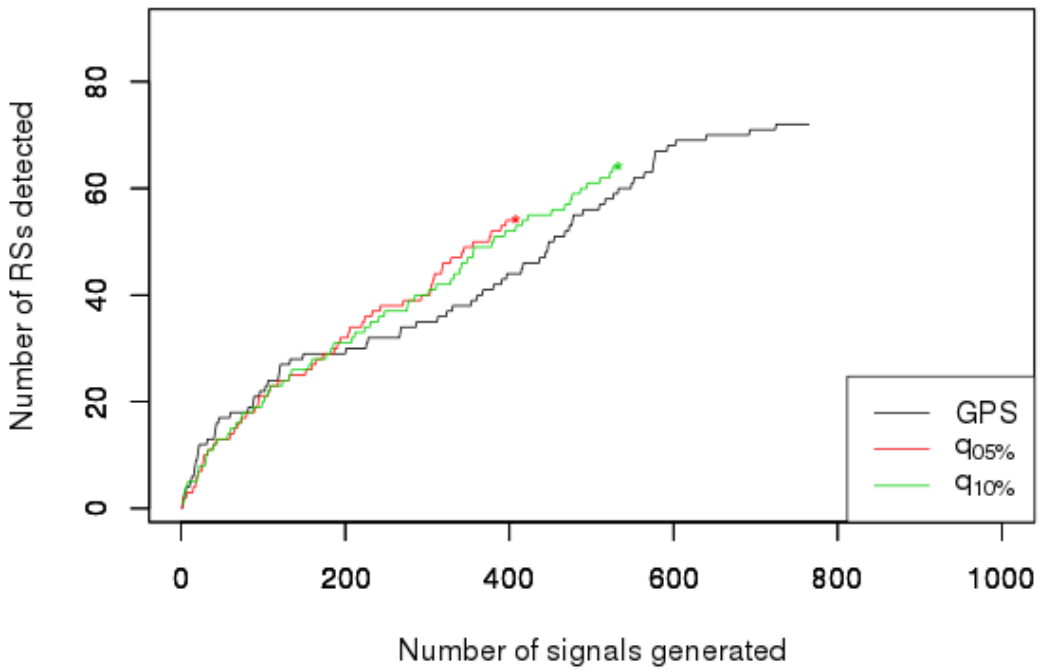
Thank you for your attention

Results of the empirical evaluation

Imbalanced sampling



N/2 sampling



β_0	β	n. true pred.	n. of cases	Imbalanced Sampling	$\lfloor n/2 \rfloor$	Ratio
-8	NA	0	39	3.08	122.63	39.8
-8	1	10	41	4.28	107.89	25.2
-8	1	30	45	4.69	105.95	22.6
-8	2	10	47	4.85	107.11	22.1
-8	2	30	65	4.86	118.25	24.3
-6	NA	0	290	4.75	70.26	14.8
-6	1	10	301	4.74	69.07	14.6
-6	1	30	327	4.76	67.97	14.3
-6	2	10	342	4.81	67.51	14.0
-6	2	30	463	4.84	53.87	11.1
-4	NA	0	2106	4.92	8.92	1.8
-4	1	10	2203	5.04	9.63	1.9
-4	1	30	2381	5.16	9.54	1.8
-4	2	10	2446	5.39	10.91	2.0
-4	2	30	3137	5.92	11.19	1.9