Empirical assessment of OHDSI case-based methods and case-population for the identification of drug–related outcome in the French nationwide healthcare database (SNDS)

Nicolas Thurin[1, 2], Régis Lassalle[1], Patrick Blin[1], Marine Pénichon[1], Martijn Schuemie[3], Joshua J Gagne[4], Jeremy A. Rassen[5], Jacques Benichou[6, 7], Alain Weill[8], Cécile Droz-Perroteau[1], Nicholas Moore[1, 2]

[1]Bordeaux PharmacoEpi, INSERM CIC1401, Université de Bordeaux, Bordeaux, France; [2]INSERM U1219, France; [3]OHDSI, New York, NY, USA; [4]Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA; [5]Aetion, Inc., New York, NY, USA; [6]CHU de Rouen, Rouen, France; [7]INSERM U1181, University of Versailles, St-Quentin-en-Yvelines, France; [8]Caisse Nationale de l'Assurance Maladie, Paris, France

Background

- **SNDS** is the French Nationwide Healthcare System Database
  - covering 66.6 million persons (99% of the French population),
  - including individual pseudonymised information on
    - Sociodemographics
    - Drug dispensings
    - Hospital discharge diagnoses
    - Procedures
    - Deaths
    - Costs, *etc.*
- SNDS extractions available after approval for public health purposes, including drug-related risk identification
- Risk identification performances depends on
  - the method,
  - the method settings
  - the environment = **the database**
- ➔ **Tools need to be tested and assessed in *real life* to ensure the generation of meaningful point estimates.**

Objectives

- To evaluate and compare the performances in the SNDS of
  - Self-controlled case series (SCCS)
  - Case control (CC)
  - Case-population (CP)
- For the identification of
  - ALI (Acute liver injury)
  - AKI (Acute kidney injury)
  - MI (Myocardial infarction)
  - UGIB (Upper gasrtointestinal bleeding)

Methods

- **Construction of a reference set adapted to the French Market**
  - 273 drug-outcome pairs from OMOP and EU-ADR reference sets
    - 4 health outcomes of interest (ALI, AKI, MI and UGIB)
    - 139 drug controls supposed associated (*positive*) or not (*negative*) with the outcome
    - Restricted to the pairs with the minimum detectable relative risk (MDRR) <1.30
- **SNDS data extractions** based on cases (2009-2014 period)
  - ALI, non sampled
  - AKI, sampled at 1/3
  - MI, sampled at 1/20
  - UGIB, sampled at 1/10
- **Detection of the drug-outcome pairs** *via* the 3 designs and different settings
  - **96** SCCS variants
  - **20** CC variants
  - **80** CP variants
  - ➔ Generation of one point estimate by pair and variant, in total, 26 068.
- **Performance assessment** of the design variants based on
  - Discriminating power: area under the receiving operator curve (AUC)
  - Accuracy: mean square error (MSE), coverage probability

Results

- Over 6 years, the number of cases extracted from the SNDS ranges from 5 152 (ALI) to 304 369 (MI) [Table 1]

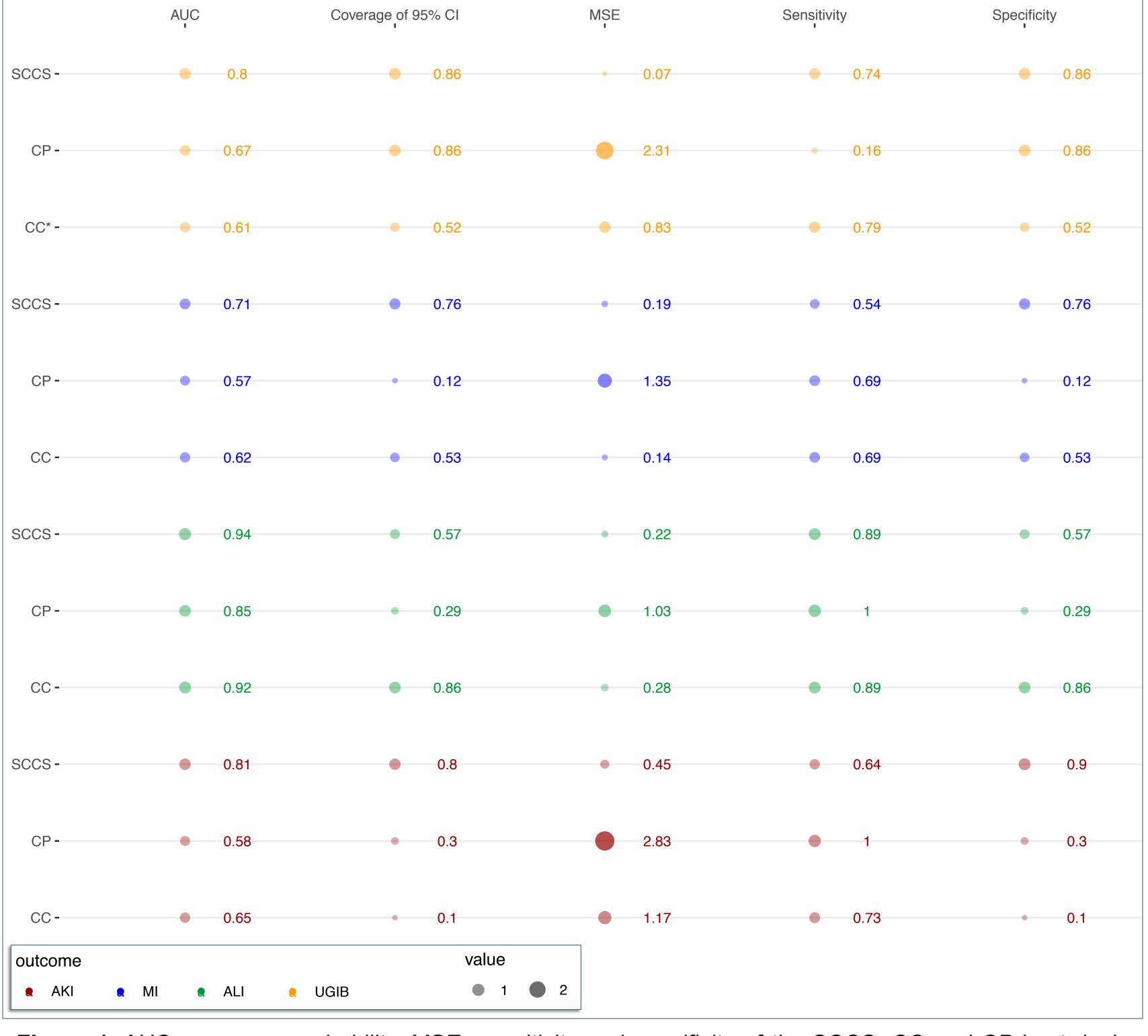**Table 1**. Outcomes and number of patients extracted from SNDS by health outcome of interest

	ALI	MI	AKI	UGIB
Outcomes	5 225	354 109	12 633	156 057
Patients	5 152	304 369	12 317	139 172

- Considering the raw data extraction of the SNDS, the number of drug-outcome pairs with a MDRR<1.30 ranged from 25 for ALI to 61 to MI [Table2]
- Sampling data extraction was necessary to reduce execution time. The result is a decrease of the number of detectable drug-outcome pairs, especially the number of negative controls

Table 2. Number of positive and negative controls by health outcome of interest available in the French market and detectable in the SNDS

Drug controls		French market Reference set	Number of detectable controls[1]			
			SNDS extraction raw	1/3rd sample	1/10th sample	1/20th sample
ALI	+	58	18			
	-	23	7			
MI	+	28	25			26
	-	42	36			20
AKI	+	22	17	11		
	-	36	13	10		
UGIB	+	22	22		19	
	-	42	36		22	

[1] Drug controls with MDRR<1.30

Figure 1. AUC, coverage probability, MSE, sensitivity and specificity of the SCCS, CC and CP best design variants across the four health outcome of interest.

- SCCS variants achieved the best performance across all outcome definitions with AUC ≥0.9 for ALI, ≥0.8 for KI and UGIB and ≥0.7 for MI [Figure 1]
- CC achieved higher AUC than CP for ALI (≥0.89 *vs.* ≥0.85), KI (≥0.62 *vs.* ≥0.58) and MI (≥0.62 *vs.* ≥0.57); and lower for UGIB (≥0.60 *vs.* ≥0.67)
- The best predictive accuracy was observed with SCCS for UGIB (MSE=0.07) and MI (MSE=0.19)
- MSE was lower for the most discriminating CC variants (0.23≤MSE≤1.34) as compared to the most discriminating CP variants (1.07≤MSE≤3)

Conclusions

- **OHDSI Methods Library can be implemented in the SNDS framework**
- **First overview of CP performances: CP seems to be less accurate and discriminant for alert generation than SCCS and CC**
- **SCCS achieves better performances across all outcomes with**
  - **high discriminative ability**
  - **high predictive accuracy**
- **Designs and settings tested can be used as reference methods for the identification of drug-related outcome in the SNDS**