

## Harmonization of Outcome Extraction for Ischemic and Hemorrhagic Stroke across Data Sources in the SAFEGUARD Project



Niklas Schmedt<sup>1</sup>, Miguel Gil<sup>2</sup>, Elisa Martin<sup>2</sup>, Gema Requena<sup>3</sup>, Irene Bezemer<sup>4</sup>, Giorgia De Berardis<sup>5</sup>, Corine de Vries<sup>6</sup>, Ingrid Leal<sup>7</sup>, Gwen Masclee<sup>7</sup>, Giampiero Mazzaglia<sup>8</sup>, Peter Rijnbeek<sup>7</sup>, Silvana Romio<sup>9</sup>, Cormac Sammon<sup>6</sup>, Lorenza Scotti<sup>9</sup>, John Seeger<sup>10</sup>, Mark Smits<sup>11</sup>, Tania Schink<sup>1</sup>, Cristina Varas-Lorenzo<sup>12</sup>, Miriam Sturkenboom<sup>7</sup>, Edeltraut Garbe<sup>1</sup>

<sup>1</sup> Leibniz Institute for Prevention Research and Epidemiology – BIPS GmbH, Bremen, Germany; <sup>2</sup> Spanish Agency for Drugs and Medical Devices, Madrid, Spain; <sup>3</sup> Pharmacology Unit, Department of Biomedical Sciences II, University of Alcalá, Madrid, Spain; <sup>4</sup> PHARMO Institute, Utrecht, Netherlands; <sup>5</sup> Consorzio Mario Negri Sud, Santa Maria Imbaro, Italy; <sup>6</sup> University of Bath, Bath, United Kingdom; <sup>7</sup> Erasmus University Medical Center, Rotterdam, Netherlands; <sup>8</sup> Fondazione Scientifica SIMG-ONLUS, Firenze, Italy; <sup>9</sup> University Milano-Bicocca, Milan, Italy; <sup>10</sup> The Brigham and Women's Hospital, Harvard Medical School, Boston, United States; <sup>11</sup> VU University Medical Center, Amsterdam, Netherlands; <sup>12</sup> RTI Health Solutions, Barcelona, Spain

### Background

In the EU-funded project "Safety evaluation of adverse drug reactions in diabetes" (SAFEGUARD), information of several healthcare databases (DBs) from Europe and the USA will be analyzed to investigate the safety of glucose lowering medications (excl. insulin) in patients with Type 2 Diabetes mellitus. An important step prior to initiating multi database observational safety studies is the harmonization of the definitions and the extraction of diagnosis codes for outcomes in all DBs to avoid inconsistencies.

### Objectives

To harmonize the definition of ischemic (IS) and hemorrhagic stroke (HS) and to assess the comparability of their incidence rates (IRs) in the different DBs involved in the SAFEGUARD project.

### Methods

- Eight population-based cohorts were extracted without any exclusion criteria from eight European DBs and one cohort was extracted among subjects aged 65 years and older from the USA: Italy (HSEARCH/Lombardy/Puglia), UK (CPRD), Spain (BIFAP), Netherlands (IPCI, PHARMO), Germany (GePaRD), USA (Medicare).
- The study period comprised the years from 1999 to 2012 depending on data availability of each DB.
- Based on a clinical definition, codes for IS and HS events in different coding systems were harmonized: ICD-9-CM (Lombardy, Puglia, PHARMO, HSEARCH and Medicare), ICD-10-GM (GePaRD), READ (CPRD) and ICPC (IPCI and BIFAP). Three DBs (IPCI/BIFAP/HSD) also used free-text algorithms.
- Data was extracted locally from each DB and processed using standardized software (Jerboa) to obtain IRs (stratified by age and sex) and standardized IRs (SIR) (WHO population) per 100,000 person years (PY) for each DB.

### Results

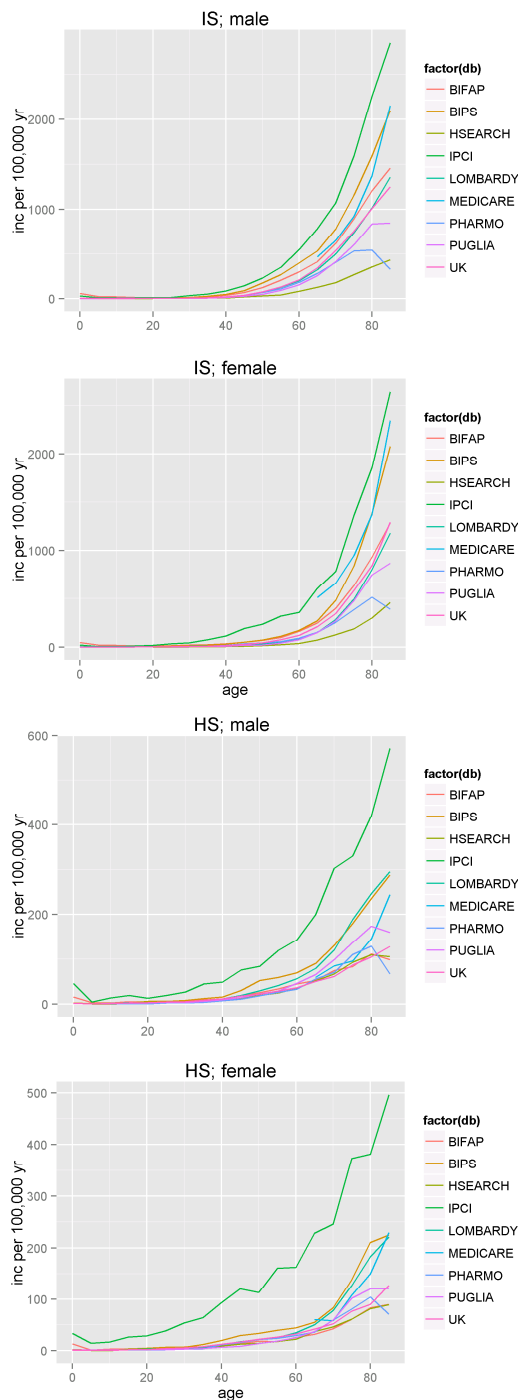
- From a source population of almost 42 million subjects, we detected 360,546 incident IS events and 81,700 incident HS events during more than 267 million PYs.
- Age- and sex-specific IRs revealed similar patterns for both IS and HS (**Figure 1**).
- Compared to other DBs, IRs were considerably higher in IPCI for both IS and HS and lower in HSEARCH for IS (**Figure 1**).
- SIRs for IS and HS per 100,000 PYs ranged from 23.9 to 174.2 and 11.4 to 71.1, respectively, for European DBs.

### Conclusions

Outcome harmonization of IS and HS is crucial to detect inconsistencies between DBs and between outcome definitions across different coding systems prior to initiating multi database observational safety studies. Observed variations among DBs might be explained by different background incidences, different characteristics of source populations, use of different coding systems and different types of DBs (electronic medical record DBs, record-linkage systems and administrative DBs). Further validation efforts will be applied to verify the accuracy of algorithms used to identify IS and HS events.

### Contact Information

**Niklas Schmedt**  
Department of Clinical Epidemiology  
Leibniz Institute for Prevention Research and Epidemiology – BIPS GmbH  
Achterstraße 30, 28359 Bremen, Germany  
E-mail: schmedt@bips.uni-bremen.de



**Figure 1: Incidence Rates of IS and HS Diagnoses stratified by Age and Sex**

### Conflict of Interest

E. Garbe is the head of a department and T. Schink and N. Schmedt are working for a department that occasionally performs studies for pharmaceutical companies such as Bayer, Celgene, GlaxoSmithKline, Mundipharma, Novartis, Sanofi-Aventis, Sanofi Pasteur MDS, and STADA. E. Garbe has been a consultant to Bayer-Schering, Nycomed, Teva, GlaxoSmithKline, Schwabe and Novartis. J. Seeger is a consultant to Optum Insight and WHISCON. C. Varas-Lorenzo as an employee of RTI Health Solutions participates in project advisory boards funded by pharmaceutical companies. M. Sturkenboom is the head of an academic unit that conducts research for pharmaceutical companies: Pfizer, Lilly, AstraZeneca, Boehringer Ingelheim.