**PNR74**
Automated detection of adverse drug events from older inpatients' electronic medical records using structured data mining and natural language processing

**Name and address of responsible researcher at the University Hospital Zurich (USZ)**
Nicola Colic
University Hospital Zurich
Raemistrasse 100
CH-8091 Zurich
079 754 67 61
nicola.colic@usz.ch

**Confirmation of the responsible researchers**
With my signature I confirm that everything stated in this document is correct, and that the study will be conducted accordingly.

**Responsible researcher at USZ: Nicola Colic**

Zurich, January 6th, 2020 _____
Location and date

_____
Signature

**Coinvestigator (external collaborator):**
Pr Chantal Csajka (hereafter "CC"); principal investigator of the study and of the SNF grant, Centre de Recherche et Innovation en Sciences Pharmaceutiques cliniques, University Hospital and University of Lausanne, Lausanne. According to the governance plan submitted to the Ethical Committee of the Canton de Vaud, she is the responsible person who gives authorization to reaearchers to complete certain tasks related to the project. The collaborators in charge are:
Dr. med. Marie-Annick Le Pogam (hereafter "MALP"), MD, MPH, coinvestigator in charge of structured data analysis and supervision of structured data analysis. Centre universitaire de médecine générale et santé publique (Unisanté), Université de Lausanne, Lausanne.
Dr Jérôme Pasquier (hereafter "JP"), PhD, senior statistician in charge of structured data analysis Centre universitaire de médecine générale et santé publique (Unisanté), Université de Lausanne, Lausanne.

Lausanne, January 6th, 2020 _____
Location and date

_____
Signature

**Dr Marie-Annick Le Pogam**

Marie-Annick Le Pogam

Signature numérique de Marie-Annick Le Pogam
DN : cn=Marie-Annick Le Pogam, o=Centre Universitaire de Médecine Générale et Santé Publique (Unisanté), ou=Département Epidémiologie et Systèmes de Santé, email=marie-annick.le-pogam@unisante.ch, c=CH
Date : 2020.02.28 18:11:01 +01'00'

**Dr Jérôme Pasquier**

**Pr Chantal Csajka**

**Background and aim**
This project aims at the automated identification of adverse drug events induced by antithrombotics prescribed for inpatients aged 65 or older.

**Origin of the data**
The data were extracted from the clinical data warehouse of the clinical information system «KISIM» at the USZ. All data were routinely collected as part of usual daily care for the patients.

**Inclusion criteria**
- Patients aged ≥65
- A hospital admission between 01.01.2015-31.12.2016
- At least one antithrombotic drug prescription during each stay
- Hospital stays >24 h.

**Exclusion criteria**
Outpatients and patients aged <65. Patients without any anticoagulants prescribed.

**Patient data from electronic health records**
This project analyzes pseudonymized data (structured electronic health record data) and de-identified data (reports). The Cantonal Ethics Committee in Lausanne (project identifier 2018-00272) approved the study protocol for all involved hospitals (University Hospital Zurich, and hospitals in Lausanne, Geneva, and Baden). The real patient and admission identification numbers were extracted from the clinical data warehouse but were replaced by other identifiers. A key table is stored at the Research Data Service Center (RDSC) that could enable reestablishing the connection between the extracted datasets (pseudo-identifiers) and the patients' electronic health records. **The key table will not be sent to MALP, JP, or CC**.

**Methods**
A variety of computational techniques will be used to automatically identify adverse drug events induced by antithrombotics. Rule-based algorithms and machine learning are some of the techniques used.

**Description of data**
From the patient report database at USZ, approximately 18000 reports were extracted that match the criteria described above. These reports are extracted in JSON format, within which there are free text sections in which doctors note symptoms, consultations, diagnoses, procedures performed and drugs administered. Those free text sections average to 1500 words per report. The entire report, including some structured meta-information, has been deidentified according to HIPAA standard, replacing identifying information such as name, address, hospital unit etc. by placeholders.

**Data protection**
The anonymized data will be transferred to the external collaborators by means of the official USZ encrypted data transfer service "secure file transfer" (transfer.usz.ch). The data will always remain encrypted when stored outside of the hospital, and only the specified external collaborators will ever have access to the data with an appropriate password set by CC. This password will neither be shared within the research group nor with any other person, except for MALP and JP. No third party besides the USZ collaborators and the specified external collaborators, MALP, JP, and CC will ever have access to the data according to this document (any other use of the data requires another agreement). Any hardware used in this research project will remain inaccessible to the public (e.g. will be locked in an office).

The external collaborators shall only use the transferred data for the purpose agreed, i.e. the research project as described herein and is not allowed to transfer the data to a third party. In particular, the data must not be used for commercial purposes and no patents will be based on the data or evolve from the project.

**Ethics and regulatory affairs**
We refer to the approved study protocol (see appendix 1). The Cantonal Ethics Committee in Lausanne (project identifier 2018-00272) approved the study protocol for all involved hospitals (University Hospital Zurich, and hospitals in Lausanne, Geneva, and Baden).

**Reporting obligations**
This study was approved by the Cantonal Ethics Committee in Lausanne (project identifier 2018-00272) for all involved hospitals (University Hospital Zurich, and hospitals in Lausanne, Geneva, and Baden). The project is not submitted to any reporting obligation.