DRUGS SAFE

Drugs Systematized Assessment in real-liFe Environment

ansm Agence nationale de sécurité du médicament et des produits de santé

# *Biostatistician / Machine Learning Data Analyst/ Data Scientist (PhD, engineer or Master 2 level)*

The SISTM (*Statistics In System biology and Translational Medicine*) team common to the Bordeaux Research Centre for Population Health INSERM U1219 www.bordeaux-population-health.center and to the INRIA Bordeaux - Sud-Ouest Research Centre www.inria.fr/en/teams/sistm offers a **1 year position** of Biostatistician/ Machine Learning Data Analyst / Data Scientist.

***Background:***

French health insurance databases (SNIIRAM) cover the entire French population. These databases include demographic (age, gender, city of residence), and out-hospital reimbursement (drug dispensing and long-term diseases). The permanent, representative cross-sectional sample of health insurance beneficiaries created from the SNIIRAM database (EGB) monitors beneficiaries' health care consumption over a period of 20 years. It contains anonymous sociodemographic and medical characteristics and records of health care reimbursements. The national healthcare databases can be used to conduct longitudinal studies (where measures are repeated over time) as they permit tracing back patients' care paths and use of care in both hospital and office-based care environments and to calculate individual expenditures. Recently, a growing attention has moved toward the exploitation of administrative databases to conduct epidemiological and drug safety research (Bezin et al. The national healthcare system claims databases in France, SNIIRAM and EGB: Powerful tools for pharmacoepidemiology. *Pharmacoepidemiol Drug Saf*. 2017).

***Mission:***

The goal is to adapt statistical machine learning methods for Big Data analysis to the longitudinal nature of exposures found in administrative health databases (e.g. reimbursed prescription drugs) to accurately predict the health event of interest (e.g. injury).

Use of high performance computing is a key point.

The current work program of the SISTM team includes the development of specific algorithms (e.g. Lasso and Random Forest methods) for that purpose.

***Team:***

This project will be conducted in the INRIA/INSERM SISTM team under supervision of Marta Avalos and in close collaboration with the INSERM teams of *Pharmaco-epidemiology and Population Impact of Drugs* and *Injury epidemiology, transport, occupation* in Bordeaux as part of DRUGS-SAFE platform funded by ANSM.

université de BORDEAUX

Instituts thématiques Inserm Institut national de la santé et de la recherche médicale

Inria INVENTEURS DU MONDE NUMÉRIQUE

DRUGS SAFE

Drugs Systematized Assessment in real-liFe Environment

ansm Agence nationale de sécurité du médicament et des produits de santé

***Profile:***

Preferably PhD in Biostatistics, Statistics, Machine Learning, or Computational Science and Engineering.

Applications with Engineer or Master 2 degree in these areas will also be considered but not as a priority.

A strong background in the development of statistical, machine learning or datamining algorithms for solving important large and complex problems.

Demonstrated conceptual and practical knowledge and skills operating with large datasets, particularly focused on developing fast and efficient algorithms that can potentially be parallelized on high performance computing environments using R or another language (Fortran, Matlab, Python, C/C++).

The ability to work effectively as part of a multi-disciplinary research team, plus the motivation and discipline to carry out autonomous research.

Fluency in French and/or English.

***Start of the contract:***

As soon as possible from November 1st 2017.

No later than January 1st, 2018.

***Salary:***

According to diploma and experience, between 1600€ (engineer or master level, with less than 3 years of experience) and 2085€ (PhD level, with less than 3 years of experience) net monthly.

***Applications:***

Send CV, references and letter of motivation (in French or English) by email to Marta AVALOS, Email : marta.avalos-fernandez@inria.fr

email subject: « DRUGS-SAFE position »

SISTM INRIA / INSERM U1219 - University of Bordeaux

146 Rue Leo Saignat, 33076 Bordeaux Cedex, France