



Master project 2021-2022

Personal Information

| | |
|--------------------|---|
| Supervisor | Alfonso Valencia, Davide Cirillo & Jon Sánchez |
| Email | alfonso.valencia@bsc.es; davide.cirillo@bsc.es; jon.sanchez@bsc.es |
| Institution | Barcelona Supercomputing Center |
| Website | http://life.bsc.es/compbio |
| Group | Computational Biology |

Project

Computational systems biology

Project Title:

Generation of age- and gender-dependent temporal disease trajectories from the Catalan population

Keywords:

Comorbidity, Electronic Health Records, personalized medicine

Summary:

The candidate will join Alfonso Valencia's lab (Computational Biology) within the Life Sciences Department at the Barcelona Supercomputing Center. The project will focus on the extraction of age- and gender-dependent temporal disease trajectories from primary care data from the Catalan population. Comorbidity is defined as a higher-than-expected risk of developing a secondary condition in patients already suffering from previous disorders. In particular, aging plays a major role in the emergence of comorbid chronic diseases. Indeed, as a consequence of ageing, chronic diseases tend to accumulate, affecting life quality, increasing disability and hindering the choice of the proper treatment [1]. Historically, comorbidity relations have been analysed by conducting case-control studies on an assembled cohort followed for a period of time (variable depending on the relationships to be analysed). The creation and accumulation of Electronic Health Records, EHRs (longitudinal collections of electronic health information including disease diagnoses and laboratory test results among others), enables the systematic analysis of comorbidity relations. Tens of studies have analysed comorbidity relations between diseases using EHRs. Hidalgo et al. generated a static view of disease comorbidity relationships analysing insurance claims from a population of elderly citizens in the USA [2]. They observed different disease associations depending on patients' age and ethnicity and identified that more connected diseases tend to be associated with an increased mortality. Jensen et al. went one-step further and, using the EHR registry of the population of Denmark, composed by more than 6.2 million patients, studied the temporal association between diseases [3]. They identified disease pairs that tend to co-occur more than the expected by chance presenting a significant temporal association (disease A precedes disease B more times than the other way around). By concatenating significantly associated disease diagnoses, they identified sequences of diseases followed by hundreds of patients, known as disease trajectories. By means of these analyses, the authors evaluated the risk of developing secondary diseases depending on the sequence of previous diagnoses, identifying key disorders that are central to disease progression, such as chronic obstructive pulmonary disease. Following Jensen's study, Westergaard et al. analysed the effect of patients' gender in the risk of developing secondary conditions, generating gender-specific temporal disease trajectories [4]. Since diseases' age of onset varies in each disease, we expect that comorbidity relations and associated trajectories could be dependent on patients' age. As an example, young type 2 diabetes patients tend to develop depression and schizophrenia as comorbid conditions, while older ones tend to suffer from cardiovascular diseases [5]. To date, no study has evaluated on a systemic way the effect of ageing on temporal disease trajectories. The Information System for Research in Primary Care (SIDAP) contains disease diagnoses, coded using the International Code of Diseases, 10th version, conducted in primary care centres covering the 75% of the Catalan population (more than 5.5 million patients) from 2008 to 2018. This project aims at identifying gender- and age-dependent comorbidity relationships and disease trajectories in the Catalan population. Making use of the SIDAP data, the student will first generate temporal disease trajectories, following a case-control approach, that will be compared to the ones generated by Soren Brunak's lab [3,4] in the Denmark population. Secondly, the role of gender and ageing on comorbidity relations will be evaluated stratifying patients based on their gender and age and generating gender- and age-dependent temporal disease trajectories.

References:

1. Barnett K. et al. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. Lancet. 380: 37-43. (2012) 2. Hidalgo, C. A. et al. A Dynamic Network Approach for the Study of Human Phenotypes. PLoS Comput. Biol. 5, e1000353 (2009) 3. Jensen, A. B. et al. Temporal disease trajectories condensed from population-wide registry data covering 6.2 million patients. Nat. Commun. 5, 4022 (2014) 4. Westergaard D. et al. Population-wide analysis of differences in disease progression patterns in men and women. Nat. Commun. 10, 666 (2019) 5. Klimek P et al. Quantification of Diabetes Comorbidity Risks across Life Using Nation-Wide Big Claims Data. PLoS Comput. Biol. 11(4): e1004125 (2015).

Expected skills::

Background in biomedicine // Basic computational skills (R, python and bash programming) // Ability to analyse and interpret large datasets // Ability to access and evaluate scientific literature

Possibility of funding::

Yes

Possible continuity with PhD: :

To be discussed

Comments:

Please send any future communications to supervisors and Alba Jené.
