Facilitating Genetic Risk Assessment of Mental Disorders

Diana Martínez-Minguet
dmarin@vrain.upv.es
Valencian Research Institute for Artificial Intelligence,
Universitat Politècnica de València
Valencia, Spain

Mireia Costa
micossan@vrain.upv.es
Valencian Research Institute for Artificial Intelligence,
Universitat Politècnica de València
Valencia, Spain

Alberto García Simón
algarsi@vrain.upv.es
Valencian Research Institute for Artificial Intelligence,
Universitat Politècnica de València
Valencia, Spain

Óscar Pastor
opastor@dsic.upv.es
Valencian Research Institute for Artificial Intelligence,
Universitat Politècnica de València
Valencia, Spain

ABSTRACT
Polygenic Risk Scores (PRSs) can contribute to genetic risk assessment of common diseases. In this work we aim to improve the translation of PRS research into a real-world clinical setting. Barriers are heterogeneity and complexity characterizing data and tools in PRS research. To overcome them, we will characterize the framework for homogenization and understandability purposes, and provide a method to assist in the genetic risk analysis process, enhancing its application in medical genetics. Well-known issues in this domain regarding gender inequalities are the under-representation of women in medical research and the poor reporting of sex and gender in medical studies. We will carefully report gender and sex specifications contributing to responsible computing for gender equality.

1 EXTENDED ABSTRACT
Computer science is in widespread use in a variety of fields and has contributed to important advances in the field of medical genetics. In this domain, research on the genetic etiology of common diseases is of great interest since they pose the greatest burden on health care. While some well-known diseases are monogenic disorders, meaning they have a clear genetic cause attributed to a single change or defect in a gene, called a genetic variant, the majority of common diseases are polygenic disorders, with numerous genes and variants contributing to risk of developing a disease [11]. An example of a polygenic disorder is diabetes, a common disease that tends to run in families, demonstrating a strong genetic component, and for which many genetic variants have been associated [5].

These genetic risk variants have only a modest impact on overall risk and are neither necessary nor sufficient to cause disease, in fact, they are common among the population. It is the interplay of numerous risk variants which gives a genetic predisposition to a given polygenic disorder, hence each individual can be associated a genetic risk to a disease. The measure of susceptibility is called a Polygenic Risk Score (PRS), and it is obtained from the aggregation of the contribution of each risk genetic variant found in the individual [6]. Genetic risk analyses involving PRSs can play a part in clinical decision-making assistance, aid in treatment choices, and contribute to risk stratification of the population [14].

In order to perform a genetic risk analysis, a PRS model providing risk variants and their contribution is required. However, PRS models are dispersed in the scientific literature and there is an evident lack of homogenization in their application and reporting [2, 12]. These factors hamper data management and understandability of the domain. In addition, PRS models are still under development and refinement, and many factors have an influence in model performance, hence it is unclear which model performs best. As a result, medical geneticists and clinicians often do not know how to prioritize which PRS model to use for the analysis [10]. Existing approaches mitigate some of the barriers, but comprehensive domain characterization is still lacking. Moreover, open source pipelines automatize PRS model construction and genetic risk analysis development, but do not address model prioritization and often require specialized knowledge or computational skills, which continues to impede the effective application of PRS research outside the research setting.

The goal of this project is to facilitate the translation of PRS research to a clinical setting. To do so, we will characterize the PRS models framework for homogenization and understandability of the domain, and will design a method to operationalize the prioritization of PRS models.
In this research we aim to design a solution that will mitigate the specific barriers above mentioned. A suitable methodology for addressing this type of research is Design Science, proposed by Wieringa [13], which consists of the design and investigation of artifacts in a context, in order to improve some aspect of the context. In this case, the artifacts are the above mentioned, and the context is the medical genetics field applied to common polygenic disorders.

Within the Design Science methodology a Design Cycle is defined, which has three stages: Problem Investigation, Problem Treatment and Treatment Validation. The main objectives of the research are aligned with the three stages (see Figure 1), which are the following:

- (i) Study of the state of the art and existing barriers in translation of PRS research into the clinical context;
- (ii) Design artifacts to enhance understanding of the domain and assist in the PRS analysis application;
- and (iii) Validate the designed artifacts and analyze contributions of this work.

[Figure 1: Design cycle of the research.]

An interesting use case is the Mental Disorders (MD) group, for which utilization of new technology and routine clinical data for research, including genetic medicine, has fallen behind other sectors of health care [9]. Regarding genetic factors affecting MD, despite the growing evidence on the associated genetic causes [1], MD diagnosis are still made solely on the basis of behavioural examinations, without taking into account a possible genetic contribution. However, it is expected that clinical genetic analyses will eventually become part of the diagnostic process of MD [3]. Given the high polygenicity of MD, genetic risk analyses can contribute to the advancements needed to improve the lives of people with mental health disorders.

This research has the potential to contribute to the progress of society within the medical genetics field, narrowing the gap between genetic research in common polygenic disorders and the clinical domain. Computer science tools applied in PRS research will allow for data management and knowledge generation, facilitating the transfer of research advances to clinical practice.

Regarding gender equality issues, medical research represents yet another area where women are in disadvantage. There is an under-representation of women in medical studies, which has direct implications in women’s health since clinical guidelines are developed from these studies and generalized to the general population [8, 15]. This kind of problematic is found within MD research, for instance, even though depression is more common among women, most widely used assays have been optimized in man [4, 7]. Another related issue is the poor reporting and inadequate analysis of sex and gender in medical research studies [8]. PRS models are generated from population-based studies, thus gender and sex specification of data samples will be carefully reported, in order to prevent the perpetuation of the aforementioned malpractices.

ACKNOWLEDGMENTS

This work was supported by the Generalitat Valenciana through the CoMoDID project (CIPROM/2021/023) and the predoctoral grant (ACIF/2021/117).

REFERENCES