Chapter 10
Valorizatium chapter

The importance of collaboration: open source data and expertise sharing should be a pillar of future research

The studies included in the present dissertation provided several new approaches to the study of pharmacogenetics in depression and they valorized the importance of data and expertise sharing as a successful model for future research. Pharmacogenetics is a research field mostly represented by scientists with a non-clinical background, such as bioinformatics and statistical genetics. A clinical-oriented point of view is therefore precious in psychiatric pharmacogenetics, since the importance of a good knowledge of psychopharmacology to guide hypothesis formulation and testing. Chapter 2 of this dissertation provides an example of this concept: the knowledge of the pharmacokinetics of antidepressant drugs was the preliminary step for the realization of the study which was born as a spontaneous idea during the discussion between me (a clinical scientist) and a professor in statistical genetics (Cathryn Lewis). Cytochrome P450 (CYP450) genes show a high complexity because of their polymorphic nature and the interpretation of the consequences of the different possible allelic combinations is not straightforward (https://www.pharmvar.org). The direct collaboration with experts in complementary fields allowed the combination of psychopharmacology, genetics and biostatistics in the study, which was the first to demonstrate that CYP2C19 enzymatic level can be accurately estimated using imputed genome-wide data (bioinformatics and genetic relevant finding) and that CYP2C19 poor metabolizers do not probably need any dose reduction when treated with citalopram or escitalopram (clinically relevant finding). This provides an example of complementary expertise sharing that should guide future studies in order to avoid the compartmentalization of science. Genetics and pharmacogenetics should indeed not be the ground of non-clinical scientists only.

Another key point that should be taken into account in research is the possibility of accessing previously collected databases that are often available to qualified researchers and can provide very valuable resources
for primary analyses or replication. We entered in a scientific era more and more often characterized by the sharing of not only results but also raw data. The initiatives of the US NIMH Center for Collaborative Genomic Studies provide a collection of over 150,000 well characterized patient and control samples from a wide-range of mental disorders (https://www.nimhgenetics.org). Genomic and clinical data from NIMH repositories were included in all the studies of this dissertation for replication purpose, for providing complementary evidence or as part of the primary analysis, demonstrating that open access data can provide exceptional extra-value and the importance of not being focused on only in-house data (which is also much more expensive and time consuming). Under the same spirit, data from my ongoing projects will be included in international consortiums, such as the Psychiatric Genetics Consortium (PGC). Future research should not ignore the resources already available, and at the same time sharing our own data should become a routine procedure.

Social and economic considerations
Depressive disorders are the third cause of disability considering all non-communicable diseases (Disease and Injury Incidence and Prevalence Collaboratorus, 2016). The economic burden of major depression increased from $173.2 billion to $210.5 billion in the period 2005-2010 in the US and in Europe it was estimated to be ~92 billion of euro in 2010 (Greenberg et al., 2015; Olesen et al., 2012). Treatment-resistant depression (TRD) was demonstrated to be a relevant factor in contributing to depression-related costs (Mrazek et al., 2014). For this reason, several studies included in the present dissertation assessed the genetic factors involved in TRD risk (chapters 3, 5, and 6).

In this scenario, the identification of biomarkers able to predict antidepressant response and guide treatment prescription is of undoubted value. This dissertation identified genetic markers of antidepressant efficacy and side effects at individual gene level and at pathway level. At individual gene level, there is encouraging evidence that combinatorial approaches including relevant polymorphisms can provide clinically meaningful information to guide antidepressant treatment, as showed for CYP2C19 and
FKBP5 genes in chapter 2 and chapter 3. Pharmacogenetic tests including variants in these genes have preliminary evidence of improving remission rates compared to standard care (Stamm et al., 2016; Bradley et al., 2018). Combinatorial approaches applied at pathway level (i.e. a set of functionally related genes) encompass the exciting opportunity to explain a higher proportion of variance in antidepressant response and reflect the final functional balance of a biological system relevant to antidepressant action. The application of predictive modeling at pathway level was tested in chapter 3 and it was one of the first studies of this kind in psychiatric pharmacogenetics. Many academic centers are planning or implementing start up efforts that aim to integrate genomic information and clinical information from electronic health records to optimize treatments using predictive modeling (Kalinin et al., 2018; Kannry and Williams, 2013). A deep learning approach may represent a game-changing advance compared to currently available pharmacogenetic tests that are based on the results of candidate gene studies (Fabbri et al., 2018).

Starting point for future research
The studies included in this dissertation were the starting point for planning further projects that are now in the implementation phase. Indeed, the study of treatment-resistant depression (TRD) and the idea of expanding genomic coverage to rare exomic variants were considered particularly promising since the paucity of previously published results and the points discussed in the previous section. The studies described in chapters 3, 4, 5 and 6 express the innovativeness and relevance of these concepts for improving the potential impact of pharmacogenetics at clinical level. This idea was shared by stakeholders who gave credit to the potential socio-economic impact of this line of research and funded new projects that are going to extend the studies included in this dissertation (“Fondazione Umberto Veronesi” and European Union (Marie Skłodowska-Curie fellowship)). Moving the focus of analysis from individual variants to pathways was another key point in chapters 3-7 of this dissertation and it is also the main methodological frame guiding the genetic analysis of the new projects. One important advance of the new analyses will be the computing of polygenic scores at both common and rare variant level, and their calculation using all available
variants in the genome/exome but also variants in specific pathways only. Indeed part of the so-called missing heritability may be due to the effect of rare (and possibly unknown) variants, as suggested in chapter 4. Ideally, prediction algorithms should include comprehensive genetic information (common and rare variants), but also clinical-demographic variables, a point that will be addressed in ongoing projects. Indeed the combination of these predictors is expected to fit the hypothesized mechanisms underling TRD and maximize the chances to develop valid and reproducible tests suitable for clinical application.

References