Rainbow project 12: Phenotype 2.0 - proof of principle for large scale phenotyping: major depression

“We want to lay the foundation for harmonized phenotypic assessments in ongoing studies”

In October 2012, a stir was created at the World Congress of Psychiatric Genetics in Hamburg, when the findings of a GWA-study for schizophrenia were presented: more than one hundred loci had been found. The study confirmed what many researchers already believed, but had no evidence for: that major psychiatric disorders are subject to, and caused by genes. It triggered BBMRI-NL’s Rainbow project 12, which will use existing biobank collections to set up a GWA-study into Major Depressive Disorder (MDD). Principal investigators are Professor Dorret Boomsma (VU University Amsterdam, Biological Psychology) and Professor Brenda Penninx (VUmc Psychiatry).

The idea is ingenious. Professor Boomsma explains: “The problem with research on psychiatric disorders is sample size. When you want to set up a relevant study for a disorder as common as MDD (13% of men and 24% of women in the Netherlands suffer from MDD at least once in their lifetime), you need thousands upon thousands of participants, who have been carefully phenotyped and for whom DNA and GWAS data are available. Moreover, at the moment, cohorts of subjects suffering from MDD are often compared on their genetic profiles to unscreened controls. This is a problem, as depression is so prevalent, and many controls thus may be cases. So we said: why don’t we use the cohorts that are already there and contain DNA and GWAS data, and ask the participants of those studies to take part in a phenotypical assessment, to ascertain whether or not they have ever suffered from MDD, or are still suffering from it?”

An email was sent out to all biobanks
participating in BBMRI-NL, asking if they would be interested in participating in an MDD study. “As participating meant that they would have to approach their own participants, gather the information themselves—in a way that makes it possible to link the phenotypical assessment to the correct DNA and GWAS data, without revealing to any third party the identity of the participant—perform the meta-analyses and then provide us with the aggregated data, we knew we were asking for quite an effort on their part”, says Professor Penninx.

Assessment tool

“But at the same time, we knew many biobanks would be interested. Firstly, because the data forthcoming from the phenotypical assessment remains in their possession. And secondly, because the tool we are going to build for the phenotypical assessment is not exclusively for psychiatric testing, it can be used for any phenotypical assessment. So participating in our study also means preparing your biobank for future phenotypical assessments, for instance yearly. The advantage of having such a tool readymade, and using the same tool as other biobanks, is evident.”

Professor Boomsma adds: “Developing this assessment tool will help to strengthen the framework of data and knowledge already there. It will enable closer collaboration between biobanks and studies, and all new data can be gathered in a uniform way. And the phenotypical data need not be exclusively linked to DNA and GWAS data, it can also be tied to biological data, metabolomics for instance, such as is now being gathered and analysed in Rainbow projects 3 and 4. So you see, the ‘web’ becomes ever more closely knit.”

Up until now, twenty-two biobanks have expressed their interest in participation in the MDD assessment, among which LifeLines (n=165,000), the Netherlands Twin Register (n=22,000), the Rotterdam Study (n=14,926), EPIC-NL (n=40,011), Leiden Longevity (n=3,359), and HEBON (n=27,000). In total, 397,905 samples are stored in the participating biobanks. “Allowing for the fact that not in all cases GWAS data has been gathered, we still are very hopeful that this amount of samples will lend us the scope to perform the meta-analyses we want”, says Professor Boomsma. “We are very lucky to be able to perform imputation of the data through reference sets of the Genome of the Netherlands project, BBMRI-NL’s first Rainbow project.”

Beautiful data

Professor Penninx continues: “At any rate, the phenotypical assessment will give the biobanks—and us—beautiful, clean data, not only on the occurrence (and recurrence) of MDD in their participants, but also on which type of MDD; there are several subtypes, and until now, differentiation between those has never been properly performed. Our assessment will provide us with that information. And, of course the assessment will not only provide us with MDD cases, but also with screened controls.”

Although some preliminary work has been ongoing for over half a year now, the official start data for the Rainbow project is 1 July, with a time scale of 36 months and a budget of ca € 795K. There are of course several challenges to be faced during the project. The development of the online tool is the first one, explains Professor Penninx: “The tool itself will be based on the Composite International Diagnostic Interview short form, or CIDI-sf. There have already been experiments with such an online tool in Australia and the USA, so that looks promising. The challenge lies in the encryption. How do you provide a key that at the same time guarantees anonymity to the participant, but can be traced back to the individual DNA and GWAS data? It is a puzzle, but one we will be able to solve.”

“The biggest challenge is to persuade as many people as possible to participate”, concludes Professor Boomsma. “For participants, the possibility of a better treatment for one of the most prevalent diseases will be a factor, but also the degree of anonymity with which they can participate. For the biobanks, we feel that the extra work incurred is well worth the trouble: the biobanks gain an online phenotypical assessment tool, plus they co-operate in gathering data for a project that will prepare the Netherlands for valuable future research and international collaboration.”