SMS5279

Vincent van Hoef

Dec 01, 2020

TABLE OF CONTENTS

PROJECT DESCRIPTION

1.1 Support request

Children ages 8-12 years spend an average of 4-6 hours a day watching or using screens. However, especially in children and adolescents, consistent evidence point to the negative effects of increased screen time on several aspects such as the development of physical abilities, socio-emotional development, sleep, obesity, depression, anxiety, stress regulation, and mental imagery.

Based on empirical findings, Lillard et al.describe different theories that account for long-term media influences on executive functions. One points to the fact that media time is time away from other activities that may foster executive functions. Another highlights the interferences of attentional processes due to rapid scene changes and high levels of sensory stimulation in media.

However, previous evidence most frequently based on correlational approaches or retrospective comparisons, which are biased by numerous confounding variables and errors in estimating screen time. Therefore, causal relations between screen time and executive functions are limited.

To solve this issue, here in this project we will combine the approaches of Mendelian Randomization and Polygenic Scores based on the genotype data and screen time use of a sample of children. The sample comes from the ABCD study and consists of N=11,875 individuals aged 9/10 years old at baseline. A wide range of measurements was collected for each individual. In addition to demographic and cognitive variables, we also obtained the whole-genome genotyping data.

Genotyping was performed using the Smokescreen array35, consisting of 646,247 genetic variants. Before variant imputation, quality controls (QC) on the genotyping were performed to ensure each genetic variant has been successfully called in more than 95 percent of the sample and that missingness for each individual was not higher than 20%. After this QC 517,724 SNPs and 10,659 individuals remained.

Now, we need support to perform the SNP imputation of this QCed genotype data using the appropriate reference panel. This will allow us to later calculate polygenic scores in our sample based on the effect sizes of a large GWAS and use these for Mendelian Randomization analyses.

Our main request for NGI is to perform the imputations of the genotype data for us and then perform (if needed) a post imputation QC (to guarantee parameters such as minor allele frequency above 5% and Hardy-Weinberg threshold of 10-6). Additionally, if there is anyone with the expertise, we could also use technical help later with the Mendelian Randomization analyses.

1.2 Requested analyses

The following analyses were agreed upon:

- Preprocessing of data supplied by user
    - Additional QC for HWE, Heterozygosity rate, MAF
    - PCA to check for outliers
    - Splitting into separate chromosomes
    - Pre-phasing using Shapeit
- Imputation (XY not included)
    - Imputation using IMPUTE2 (or similar)
    - Merging of data
    - Post-imputation QC with info score
    - Converting to selected output format
- Results
    - Report including methods and parameters used, code and files

TWO

MATERIAL AND METHODS

- *Quality Control*

This is a overview of the different steps taken to clean up the data and prepare the files for imputation.

2.1 Quality Control

Quality control is an important step prior to imputation.

THREE

RESULTS

- *Quality Control*
- *Imputation*

3.1 Quality Control

3.2 Imputation

FOUR

ACKNOWLEDGEMENTS

If you are presenting the results in a paper, at a workshop or conference, we kindly ask you to acknowledge us.

- **NBIS Staff** are encouraged to be co-authors when this is merited in accordance to the ethical recommendations for authorship, e.g. ICMJE recommendations). If applicable, please include **Vincent van Hoef, National Bioinformatics Infrastructure Sweden, Science for Life Laboratory, Uppsala University**, as co-author. In other cases, NBIS would be grateful if support by us is acknowledged in publications according to this example: "Support by NBIS (National Bioinformatics Infrastructure Sweden) is gratefully acknowledged".

- **UPPMAX** If your project has used HPC resources we kindly asks you to acknowledge UPPMAX and SNIC. If applicable, please add: "The computations were performed on resources provided by SNIC through Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) under Project SNIC XXXX/Y-ZZZ".

- **NGI** In publications based on data from NGI Sweden, the authors must acknowledge SciLifeLab, NGI and UPPMAX: "The authors would like to acknowledge support from Science for Life Laboratory, the National Genomics Infrastructure, NGI, and Uppmax for providing assistance in massive parallel sequencing and computational infrastructure".

FIVE

DATA RESPONSABILITY

- **NBIS & Uppnex** Unfortunately, we do not have resources to keep any files associated with the support request. We suggest that you safely store the results delivered by us. In addition, we ask that you remove the files from UPPMAX/UPPNEX after analysis is completed. The main storage at UPPNEX is optimized for high-speed and parallel access, which makes it expensive and not the right place for long time archiving.

- **Sensitive data** Please note that special considerations may apply to the human-derived sensitive personal data. These should be handled according to specific laws and regulations.

- **Long-term backup** The responsibility for data archiving lies with universities and we recommend asking your local IT for support with long-term data archiving. Also the newly established Data Office at SciLifeLab may be of help to discuss other options.

CLOSING PROCEDURES

You should soon be contacted by one of our managers, Jessica Lindvall (jessica.lindvall@nbis.se) or Henrik Lantz (henrik.lantz@nbis.se), with a request to close down the project in our internal system and for invoicing matters. If we do not hear from you within **30 days** the project will be automatically closed and invoice sent. Again, we would like to remind you about data responsibility and acknowledgements, see Data Responsibility and Acknowledgments sections.

You are naturally more than welcome to come back to us with further data analysis request at any time via NBIS support.

Thank you for using NBIS and all the best for future research.

This report described the analysis of project #5279. It contains several sections.

Project Description: A description of the projects and an overview of the requested and agreed upon analyses.

Material and Methods: Overview of the material and methods used during the project.

Results: Overview of the quality control and imputation results.

Warning: Please do not forget to read the important sections on the *Acknowledgements*, *Data Responsability* and *Closing Procedures*.