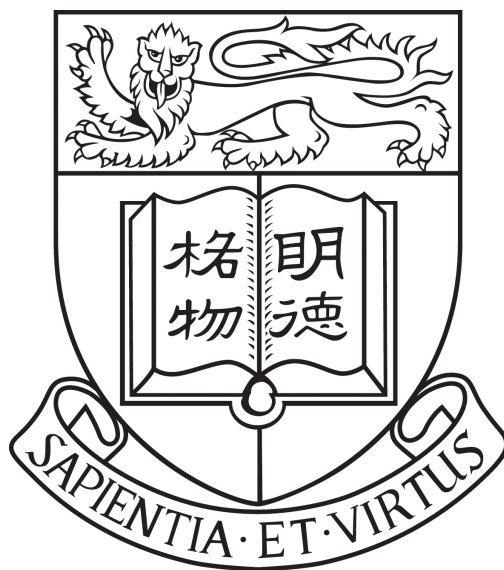


Heritability Estimation and Risk Prediction in Schizophrenia

Choi Shing Wan

A thesis submitted in partial fulfillment of the requirements for
the Degree of Doctor of Philosophy



Department of Psychiatry
University of Hong Kong
Hong Kong
September 1, 2015

Declaration

Acknowledgements

Abbreviations

SCZ Schizophrenia. 13

Contents

| | |
|---------------------------------------------------------------------------------|------------|
| Declaration | i |
| Acknowledgments | iii |
| Abbreviations | v |
| Contents | vii |
| Introduction | 1 |
| 1 Literature Review | 5 |
| 1.1 Twin Studies - Delineating Genetic and Environmental Contribution | 5 |
| 1.2 Searching for Genetic Variants | 5 |
| 1.2.1 Role of Common Variants | 5 |
| 1.2.2 Role of Rare Variants | 5 |
| 1.3 Summary | 6 |
| 2 Heritability Estimation | 7 |
| 2.1 Introduction | 7 |
| 2.2 Methodology | 7 |
| 2.3 Implementation | 7 |
| 3 Quantitative Trait | 9 |
| 3.1 Introduction | 9 |
| 3.2 Methodology | 9 |
| 3.2.1 Simulation | 9 |
| 3.3 Result | 9 |
| 3.4 Extreme Phenotype Selection | 9 |
| 3.4.1 Methodology | 9 |
| 3.4.2 Result | 9 |
| 4 Case Control | 11 |
| 4.1 Introduction | 11 |
| 4.2 Methodology | 11 |
| 4.2.1 Simulation | 11 |
| 4.3 Result | 11 |
| 5 Heritability of Schizophrenia | 13 |
| 5.1 Introduction | 13 |
| 5.2 Heritability Estimation | 13 |
| 5.3 Brain development and Schizophrenia | 13 |

| | | |
|----------|-------------------------------------------------------|-----------|
| 5.3.1 | Methodology | 13 |
| 5.3.2 | Result | 13 |
| 5.4 | Heritability of Schizophrenia drug response | 13 |
| 5.4.1 | Methodology | 13 |
| 5.4.2 | Result | 13 |
| 6 | Risk Prediction | 15 |
| 6.1 | Methodology | 15 |
| 6.1.1 | Simulation | 15 |
| 6.2 | Result | 15 |
| 7 | Conclusion | 17 |

Introduction

Some considerations

1. PRSice requires the phenotype to aid its selection (More information= stronger)
2. It seems like LDSC doesn't necessary perform badly in oligogenic situation. Rather, it is that when the trait is oligogenic, it is more likely for LDSC to behaviour in a strange way.
3. For each condition: extreme phenotype, quantitative trait, case control, we can have a separated review. Discuss on the benefits and challenges of each condition and the method we deal with them. So we can have two chapters (case control, quantitative trait) where extreme phenotype can be a big subsection within quantitative trait.
4. For each chapter, there will be this introduction (review on the method), our methodology (Calculation, implementation and also simulation), result (the simulation result). Then we can have the application (PGC, network)

Chapter 1

Literature Review

1.1 Twin Studies - Delineating Genetic and Environmental Contribution

Should briefly talk about how Twin modeling was used for finding the GE contribution. Should also mention the ACE model. At the end, we can talk about the heritability estimates of SCZ and AD

1.2 Searching for Genetic Variants

1.2.1 Role of Common Variants

Genome Wide Association Study

Should talk about what is GWAS and how it is used. Should also talk about the current GWAS studies in SCZ and AD

1.2.2 Role of Rare Variants

Exome Sequencing

Similar to the GWAS. Talk about the Pros and Cons. Need to briefly mention the Denovo paper and Shaun's paper.

Whole Genome Sequencing

Very very brief description of WGS and the current status.

1.3 Summary

Chapter 2

Heritability Estimation

This chapter should be used in similar way as the general method section in Clara's thesis. Considering that the subsequent chapters all rely on this implementation.

2.1 Introduction

2.2 Methodology

2.3 Implementation

This, I am not sure whether if I should include this section. The sliding window is definitely a challenge that we solved. So is the tSVD which we can some how put that in the calculation part.

Chapter 3

Quantitative Trait

3.1 Introduction

3.2 Methodology

3.2.1 Simulation

3.3 Result

3.4 Extreme Phenotype Selection

3.4.1 Methodology

Simulation

3.4.2 Result

Chapter 4

Case Control

4.1 Introduction

4.2 Methodology

Considering that the implementation of case control is the same as that in quantitative trait, maybe there is no need to repeat it?

4.2.1 Simulation

4.3 Result

Chapter 5

Heritability of Schizophrenia

5.1 Introduction

5.2 Heritability Estimation

This will be a very simple section, focused on how to perform the heritability estimation on Schizophrenia (SCZ). Should also tokenize the heritability into subcategories (e.g. immune, neuron, etc)

5.3 Brain development and Schizophrenia

Here we will perform the WGCNA and brain development network. Seeing how the whether if any brain development network were enriched with SNPs that explain the variance of phenotype

5.3.1 Methodology

5.3.2 Result

5.4 Heritability of Schizophrenia drug response

Here we try to use Beatrice's data and estimate the heritability explained in drug response. Should also repeat the region-wise heritability

5.4.1 Methodology

5.4.2 Result

Chapter 6

Risk Prediction

6.1 Methodology

6.1.1 Simulation

6.2 Result

Chapter 7

Conclusion

Appendix