

A Practical Primer in Human Complex Genetics

with a use-case in cardiovascular disease

dr. Sander W. van der Laan  

Version 2.0.0 (2024-03-26)

Contents

1 About this primer	5
2 Some background reading	7
3 Epilogue	9
4 Additional: EIGENSOFT	11
4.1 Install homebrew	11
4.2 Install missing packages	12
4.3 Installing EIGENSOFT	12
5 Things to do	13
5.1 MoSCoW	13
5.2 Book fixes	14
5.3 Useful links (for me mostly)	14
6 Licenses and disclaimers	15
6.1 Copyright	15
6.2 Disclaimer	16
6.3 Images used	16
7 Colofon	17

Chapter 1

About this primer



option 1 - above the header



Figure 1.1: option 3

Ever since the first genome-wide association study (GWAS) on age-related macular degeneration, and the promise of personalized medicine in the wake of the Human Genome Project, large-scale genetic association studies hold significant sway in contemporary health research and drive drug-development pipelines. In the past 2 decades, researchers delved into GWAS, aiming to unveil genetic variations linked to both human traits, such as the color of your eyes, and rare and common complex diseases. These findings serve as crucial keys to unravel the intricate mechanisms underlying diseases, shedding light on

whether the correlations identified in observational studies between risk factors and diseases are truly causal.

These studies have ushered in an exciting era where many researchers thrive on developing new methods and bioinformatic tools to parse ever-growing large datasets collected from large population-based biobanks. However, the analyses of these data are challenging and it can be daunting to see the forest for the trees among the many tools and their various functions. Enter *A Practical Primer in Human Complex Genetics*. This GitBook was originally written back in 2022 for the **Genetic Epidemiology** course organized by the Master Epidemiology of Utrecht University. This practical guide will teach you how to design a GWAS, perform quality control (QC), execute the actual analyses, annotate the GWAS results, and perform further downstream post-GWAS analyses. Throughout the book you'll work with 'dummy', that is fake, data, but in the end, we will use real-world data from the first release of the *Welcome Trust Case-Control Consortium (WTCCC)* focusing on coronary artery disease (CAD).

A major component of modern-day GWAS is genetic imputation, but for practical reasons it is not part of this book. However, I will provide some pointers as to how to go about doing this with minimal coding or scripting experience. Likewise, the course does not cover the aspects of meta-analyses of GWAS, but some excellent resources exist to which I will direct. As this practical primer evolves, these and other topics may find their place in this book. I should also point out that emphasis of this book is on it being a *practical primer*. It is intended to provide some practical guidance to doing GWAS, and while theory is important, I will not cover this. Again, some very useful and excellent work exists to which I will point you, but I really want you to learn - and understand the theory - by *doing*.

So, although originally crafted as a companion for the course, this practical guide stands on its own as a comprehensive resource for diving into all facets of doing a GWAS — save for experimental follow-up, of course □ .

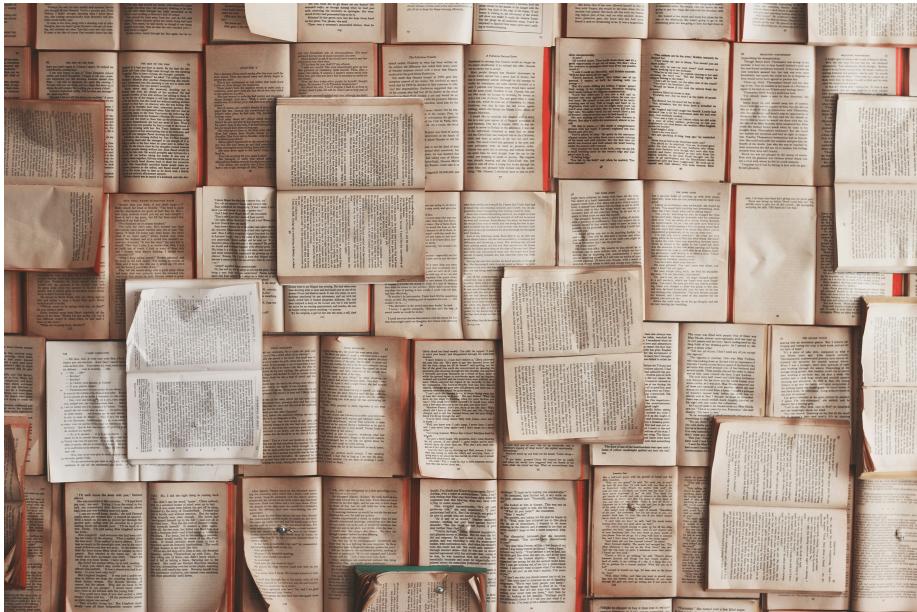
I can imagine this seems overwhelming, but trust me, you'll be okay. Just follow this practical. You'll learn by doing and at the end of the day, you can execute a GWAS independently.

Ready to start?

Your first point of action is to prepare your system for this course in Chapter 2.

Chapter 2

Some background reading



Standing on the shoulders of giants, that's what this book and I do. I want to acknowledge some great work that has helped me tremendously and, really, this book wouldn't exist without this awesome work. So, I do want to give you some background reading. Is it a prerequisite? No, not really. For starters, the course covers most and you'll learn as you go. And if you didn't come here through the course, you'll be fine just the same. That said, it's always good idea to get familiar with these works as you move forward on your path towards your first GWAS - in fact, I had these printed out with markings and writings all over them as I executed my first GWAS, and they've been great as a reference

many times after.

Large parts of this work are based on four awesome Nature Protocols from the Zondervan group at the Wellcome Center Human Genetics.

1. Zondervan KT *et al.* *Designing candidate gene and genome-wide case-control association studies.* Nat Protoc 2007.
2. Pettersson FH *et al.* *Marker selection for genetic case-control association studies.* Nat Protoc 2009.
3. Anderson CA *et al.* *Data QC in genetic case-control association studies.* Nat Protoc 2010.
4. Clarke GM *et al.* *Basic statistical analysis in genetic case-control studies.* Nat Protoc 2011.

An update on the community standards of QC for GWAS can be found here:

1. Laurie CC *et al.* *Quality control and quality assurance in genotypic data for genome-wide association studies.* Genet Epidemiol 2010.

With respect to imputation and meta-analyses of GWAS you should also get familiar with the following two works:

1. Marchini, J. and Howie, B. *Genotype imputation for genome-wide association studies.* Nat Rev Genet 2010
2. de Bakker PIW *et al.* *Practical aspects of imputation-driven meta-analysis of genome-wide association studies.* Hum Mol Genet 2008.
3. Winkler TW *et al.* *Quality control and conduct of genome-wide association meta-analyses.* Nat Protoc 2014.

Are you ready?

Are you ready? Did you bring coffee and a good dose of energy? Let's start! Your first point of action is to prepare your system for this course in Chapter ??.

Chapter 3

Epilogue

What started as a simple ‘let’s write a practical on how to do a GWAS’, escalated into this GitBook. My second book. My fifth child (as far as I know). I hope you found it to be useful and learned a bit.

That said, much can be improved and so don’t hesitate to contact me.

As with any proper epilogue I should not forget a heartfelt and honest thank you to the readers and users of this work. Gratitude also goes to my dear colleagues Charlotte Onland, Jessica van Setten, and Kristel van Eijk, and former colleague Sara Pulit who asked me back in 2017 to join the course as a lecturer. It has been a pleasure to work with and learn from you, and it has been a fun (and sometimes stressful) experience to teach this course.

Chapter 4

Additional: EIGENSOFT

There used to be a time that the preferred software **EIGENSOFT** for Principal Component Analysis (PCA) was **EIGENSOFT**. For many, it still is. However, **EIGENSOFT** is a bit challenging to make it work to say the least. You need to install some programs, and this is not always straightforward.

So, here's the deal.

I will share the how-to for a macOS environment below in [EIGENSOFT] - this should work in a Linux environment too as macOS is UNIX-based. You can choose to try and make it work on your system (be it UNIX or macOS based) at home (or in the office).

However, I recommend that you use the --pca function which is present in PLINK v1.9 and up. This means you should probably simply skip this section and jump straight to Chapter ??.

4.1 Install homebrew

You need to install brew, the missing package-manager and accompanying packages that Apple didn't provide.

```
/bin/bash -c "$(curl -fsSL https://raw.githubusercontent.com/Homebrew/install/HEAD/install.sh)"
```

Next, check that everything is in order.

```
brew doctor
```

4.2 Install missing packages

Right, now that you've done that, you're ready to install gsl and openblas.

```
brew install gsl  
brew install openblas
```

You may also require llvm for **EIGENSOFT** to work.

```
brew install llvm
```

4.3 Installing EIGENSOFT

I am still sharing the code you'll need - you could try this on your system.

```
mkdir -v $HOME/git  
cd $HOME/git  
git clone https://github.com/DReichLab/EIG.git  
cd EIG/src  
make  
make install
```

Chapter 5

Things to do

I provide a list of things to do, to improve, to alter, to edit or to add. Crazy ideas. Useful tips, tricks, or links.

Obviously this list is not exhaustive nor intended for practical use for anyone else but me.

5.1 MoSCoW

List of to-do's according to MoSCoW: *must have*, *should have*, *could have*, *would have*.

5.1.1 Contents

- ☒ M add in full installation instructions for UBUNTU
 - ⓘ M add in full installation instructions for macOS
 - ⓘ M conditional analysis
 - ⓘ M statistical finemapping
- ☒ M regional association plotting
 - ⓘ M colocalization with formal testing
 - ⓘ S meta-analysis with dummy data
 - S including stratified QQ plots
 - M HPC version with MetaGWASToolKit
 - S stand-alone version with METAL
- ⓘ C GWASToolKit
- ⓘ C PlaqView lookups

5.2 Book fixes

- M overall rendering too slow, paste in images as figure instead of on the fly generating
- M PDF is too large
 - S PDF is not formatted properly (text runs over)
 - C Different font type in PDF
 - S EPUB is not formatted properly (text runs over)
 - S different setup for the chapter Additional chapters (this as a Appendix)
- S fix the way the team is displayed
 - S fix images per header in EPUB
 - S fix book cover
 - <https://www.designhill.com/design-blog/the-perfect-ebook-cover-size-guide-and-publishing-tips/>
 - <https://snappa.com/blog/ebook-cover-size/>
 - https://kdp.amazon.com/en_US/help/topic/G200645690.

5.3 Useful links (for me mostly)

<https://bookdown.org/yihui/rmarkdown-cookbook/unnumbered-sections.html>

<https://bookdown.org/yihui/bookdown/cross-references.html>

https://pandoc.org/MANUAL.html#extension-header_attributes

Add an image above the title:

- <https://stackoverflow.com/questions/62074546/add-image-before-bookdown-title>

But it causes an issue, described here. So you better remove this chunk of code:

```
\'\''{js , echo = FALSE}
title=document.getElementById( 'header ');
title.innerHTML = '
- woman\_working\_on\_code - <https://unsplash.com/photos/woman-wearing-black-t-shirt-holding-white-computer-keyboard-YK0HPwWDJ1I>
- licenses - <https://unsplash.com/photos/book-lot-on-black-wooden-shelf-zeH-IjawHtg>

## **Chapter 7**

## **Colofon**

[Add in text on colofon]