IMPaC-TB: Integrated analysis and dynamical systems modeling of experimental TB immunology data

Brooke Anderson PhD, Mike Lyons PhD, Amy Fox, Burton Karger

2019-12-05

Contents

This is a laboratory guide book for the CSU IMPAC-TB experiments.

Chapter 1

Introduction

The overall objective of the data analysis and mathematical modeling component for the CSU mouse immunology experimental studies is to develop an iterative framework to identify key biological components of immunity and to quantify their relationships to one another in both data-driven and mechanistic models for the purpose of evidence-based decision-making for tuberculosis (TB) vaccine development. The data sharing plan (DSP) will include the experimental data, the quantitative analysis framework, and an application programming interface (API) to the results. As detailed in our respective biosketches, we have established expertise in all critical areas of this data analysis project.

The host immune response to TB vaccination and infection is complex and involves interactions between large networks of molecular and cellular constituents that vary in time and location within the host. The experimental data will be generated from a wide range of measurements and across multiple scales; including measures of disease pathology, cellular and chemical measurements for cell type and cytokine concentrations, and intracellular measurements involving RNA expression and proteomics. The conceptual basis for our proposed data analysis and modeling framework is described in a summary of the recent National Institute of Allergy and Infectious Diseases (NIAID) workshop, 'Complex Systems Science, Modeling and Immunity' [1]. The major components of this framework are illustrated in Figure 1, where our approach will integrate experimental data with data-driven modeling to identify significant correlations and possible causal structures among the data elements, and with mechanistic modeling of cell-mediated immunity that translates biologically-based hypotheses into a dynamical system of time-dependent mathematical equations that can be used to simulate and test these hypotheses and to inform the design of subsequent experiments.

The proposed work plan begins with the collection and organization of quantitative and qualitative CSU generated experimental data that will then be used for

data-driven and mechanistic modeling, with the analysis results and software modeling tools being made available through a web-based API. The milestones of this project are: (1) establishing protocols and standardized documentation for data collection and pre-processing from each CSU experimental type, (2) construction of the relational database (RDB) for CSU-generated experimental data, (3) collection of qualitative data describing key immune features as input for mechanistic modeling, (4) development of single-type data-driven analysis tools for each separate experimental system, (5) development of integrated data-driven analysis tools for the combined experimental data, (6) development of a dynamical systems model of cell-mediated immunity based on qualitative analysis results, (7) development of parameter estimation and model calibration procedures for the dynamical systems model, (8) development of software tools that provide for a start-to-finish process framework, and (9) development of an API for public access to relevant data and results. For each milestone, the gates for Go/No Go decisions will be based on the positive reproducibility of the major results by each of the individual CSU investigators. This approach will ensure the integration and quality control of each component within the entire framework.

Immunology data is collected from each CSU mouse experiment as both quantitative measurements and qualitative data that includes hypotheses regarding key biological constituents in the context of TB vaccine development. An RDB will be developed to provide access and queries to all combined data sets. Data-driven modeling will proceed directly from the quantitative data while mechanistic modeling will begin with the qualitative data, with the two modeling approaches increasingly informing each other as analyses proceed. The data-driven integrated data analysis will include visualization and statistical analyses and will also inform parameter estimation for the dynamical systems model. Software tools will be developed for all quantitative data and results, including user testing. A tailored user interface will provide access to all data and analysis results.

You can label chapter and section titles using `{#label}` after them, e.g., we can reference Chapter 1. If you do not manually label them, there will be automatic labels anyway, e.g., Chapter **??**.

Figures and tables with captions will be placed in `figure` and `table` environments, respectively.

```
par(mar = c(4, 4, .1, .1))
plot(pressure, type = 'b', pch = 19)
```

Reference a figure by its code chunk label with the `fig:` prefix, e.g., see Figure 1.1. Similarly, you can reference tables generated from `knitr::kable()`, e.g., see Table 1.1.

```
knitr::kable(
  head(iris, 20), caption = 'Here is a nice table!',
```

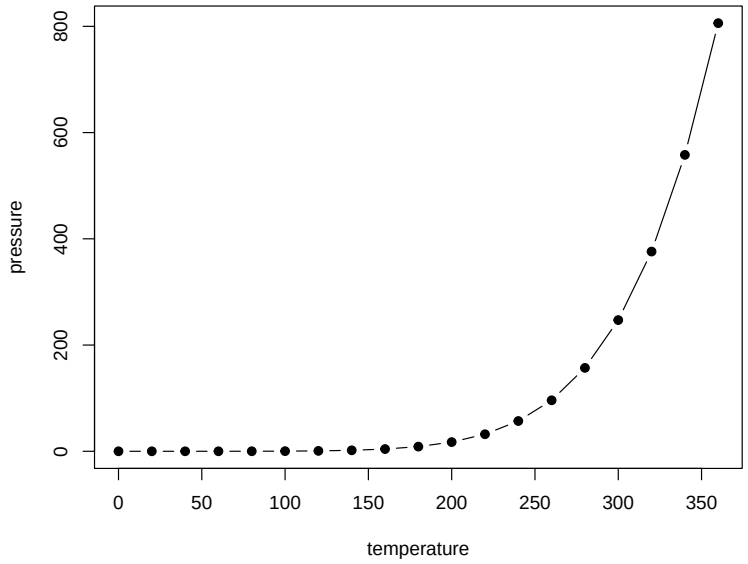

Figure 1.1: Here is a nice figure!

```
  booktabs = TRUE
)
```

You can write citations, too. For example, we are using the **bookdown** package (Xie, 2019) in this sample book, which was built on top of R Markdown and **knitr** (Xie, 2015).

Table 1.1: Here is a nice table!

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa
4.6	3.4	1.4	0.3	setosa
5.0	3.4	1.5	0.2	setosa
4.4	2.9	1.4	0.2	setosa
4.9	3.1	1.5	0.1	setosa
5.4	3.7	1.5	0.2	setosa
4.8	3.4	1.6	0.2	setosa
4.8	3.0	1.4	0.1	setosa
4.3	3.0	1.1	0.1	setosa
5.8	4.0	1.2	0.2	setosa
5.7	4.4	1.5	0.4	setosa
5.4	3.9	1.3	0.4	setosa
5.1	3.5	1.4	0.3	setosa
5.7	3.8	1.7	0.3	setosa
5.1	3.8	1.5	0.3	setosa

Chapter 2

Day -247

The purpose of ths timepoint is to identify the cage numbers with the treatments and create a short cage id.

Note: in future experiments, we will tag the mice at this timepoint.

The data template files can be found here

Chapter 3

Day -242

The purpose of this timepoint is to shave the mice before vaccination and to find the iniital total weight of the mice in the cage before vaccination.

The data template files can be found here

Chapter 4

Day -240

The purpose of this timepoint is to perform the first round of vaccinations.

The data template files can be found here

Chapter 5

Day -233

The purpose of this timepoint is to get the toal cage weight post-vaccination.

The data template files can be found here

Bibliography

Xie, Y. (2015). *Dynamic Documents with R and knitr*. Chapman and Hall/CRC, Boca Raton, Florida, 2nd edition. ISBN 978-1498716963.

Xie, Y. (2019). *bookdown: Authoring Books and Technical Documents with R Markdown*. R package version 0.16.