SIRE 2.0

Susceptibility, Infectivity and Recoverability Estimation

NOTE: This manual is currently under development and not all SIRE 2.0 features have been added yet.

Christopher M. Pooley[1,2*], Stephen C. Bishop[†], Andrea B. Doeschl-Wilson[1&], and Glenn Marion[2&]

[1] The Roslin Institute, The University of Edinburgh, Midlothian, EH25 9RG, UK.

[2] Biomathematics and Statistics Scotland, James Clerk Maxwell Building, The King's Buildings, Peter Guthrie Tait Road, Edinburgh, EH9 3FD, UK

[†] Deceased

[&] These authors contributed equally to this work.

Table of Contents

1 Introduction

Three key epidemiological host traits affect infectious disease spread: susceptibility (propensity to acquire infection), infectivity (propensity to transmit infection to others, once infected) and recoverability (propensity to recover quickly). SIRE is a desktop application for estimating factors affecting these traits from individual-based data.

SIRE takes as input any combination of information about infection times, recovery times, disease status measurements, disease diagnostic test results, genotypes of SNPs or any other fixed effects, details of which individuals belong to which contact groups and any prior specifications. The output from SIRE consists of posterior trace plots for model parameters θ, distributions, visualisation of infection and recovery times ξ, dynamic population estimates and summary statistics (means and 95% credible intervals) as well as MCMC diagnostic statistics.

1.1 Downloading

SIRE is freely available to download from [www.mkodb.roslin.ed.ac.uk/EAT/SIRE.html.](www.mkodb.roslin.ed.ac.uk/EAT/SIRE.html)

Depending on your platform, the following instructions should be followed:

- **Windows** – The file SIRE_v1.0_windows.zip is first downloaded and upzipped. SIRE is then simply run by clicking on the SIRE.exe icon.
- **Linux** – The file SIRE_v1.0_linux.tar.gz is first downloaded. This can be extracted by using the terminal command "**tar -zxvf SIRE_v1.0_linux.tar.gz". The code can then simply be executed using ./SIRE.**
- **Macintosh** – The file SIRE_v1.0_Mac.zip is first downloaded and upzipped. SIRE is then simply run by clicking on the SIRE.app icon.

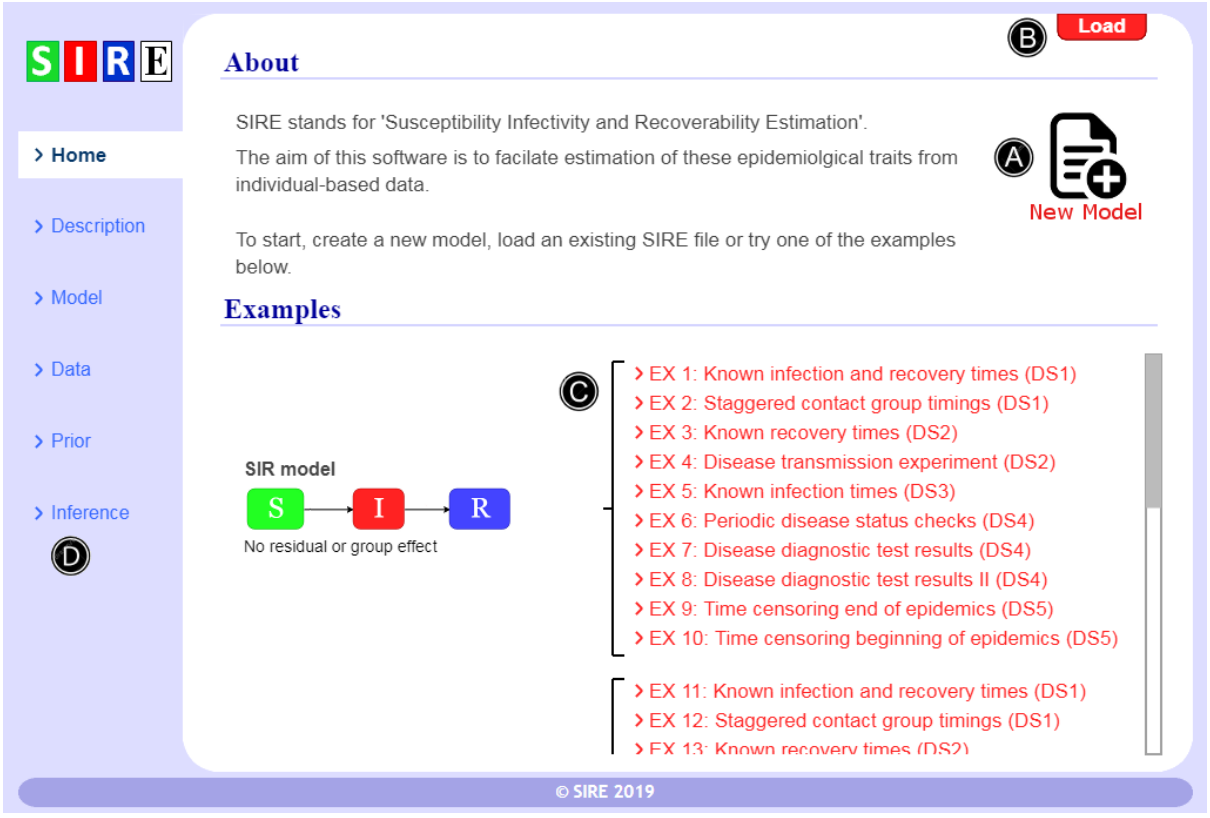


Figure 1 – *The home screen*.

2 The interface

2.1 Getting started

Figure 1 shows the first screen you see when SIRE is loaded. From this three options can be pursued: a new analysis can be started (A), a previous analysis can be loaded (B) or one of the illustrative examples can be investigated at (C). New users are encouraged to try the latter option and spend some minutes exploring the software to get a feel for how it works. The examples (C) are described in detail in section 3 below.

This manual follows the order of the menu items on the main menu (D), which is the order in which analysis would be made

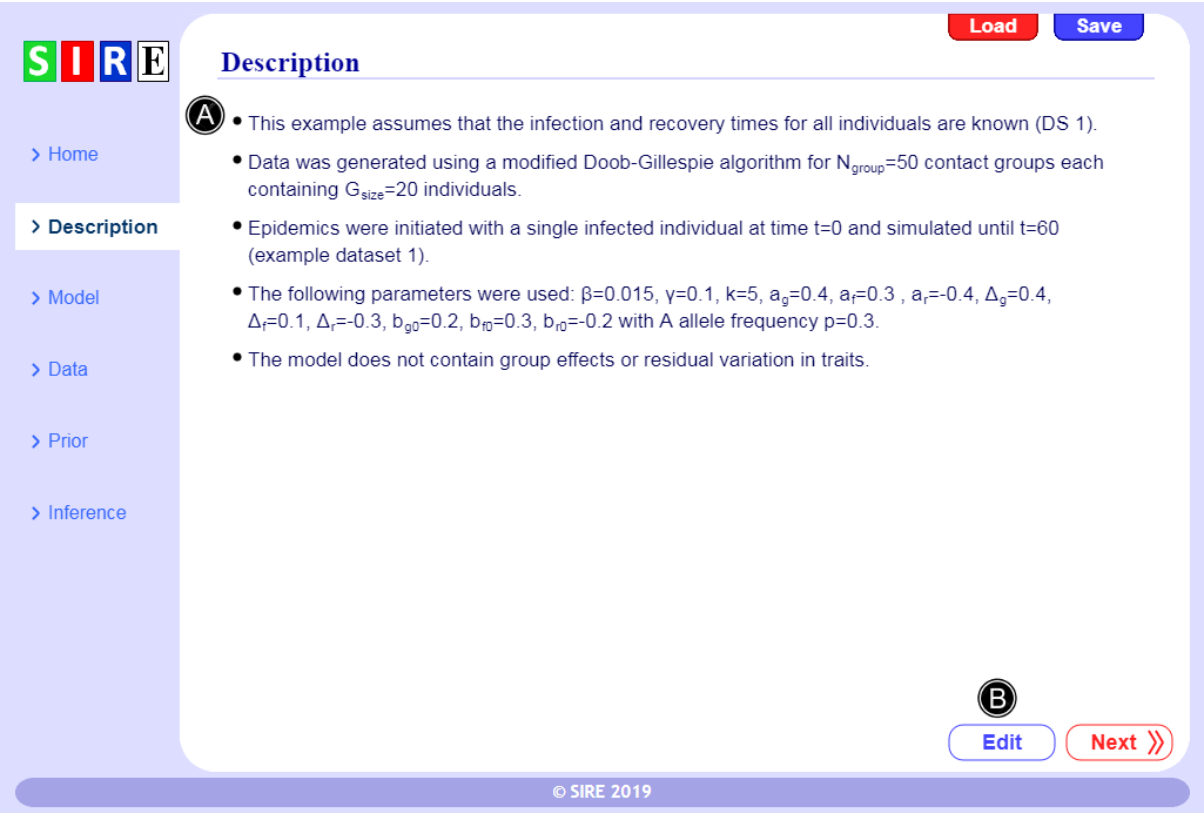


Figure 2 – Data and analysis *description*.

2.2 Analysis description

As shown in Fig. 2, SIRE allows users to provide a brief description of the data and analysis (A). This is not only useful to keep track for personal use, but also makes it easier for other to understand what has been done.

The description can simply be edited by clicking on (B). Note, bullet points are automatically generated for each carriage return.

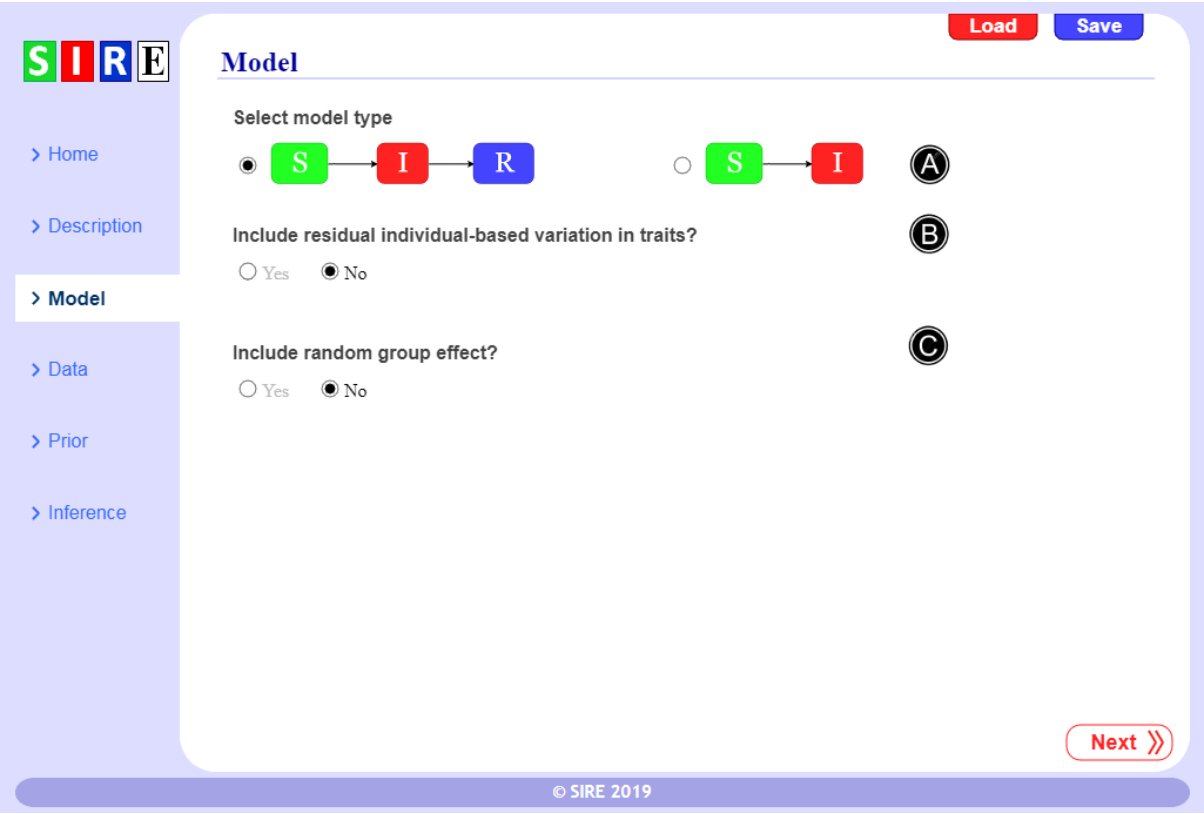


Figure 3 – *Selecting model options.*

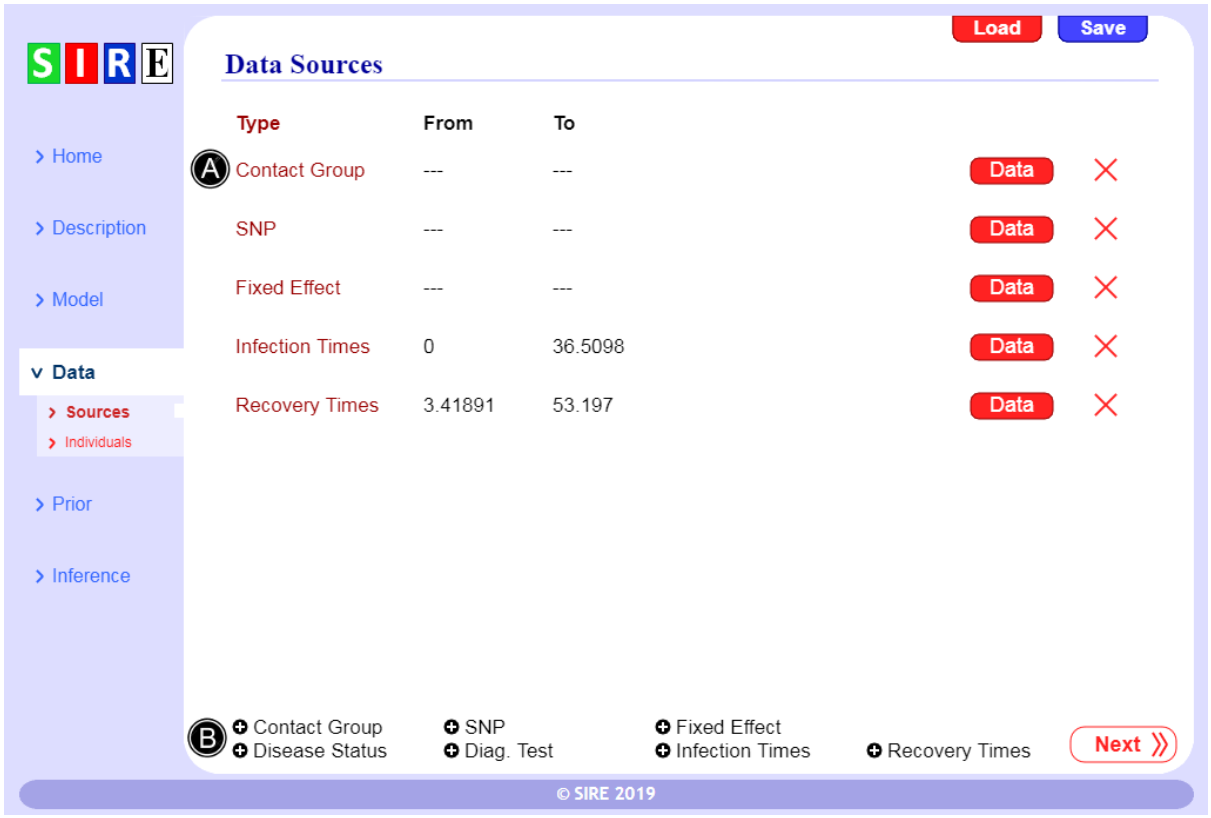
2.3 The model

Here we specify the model used for analysis, which is dependent on the nature of the disease, the conditions under which individuals are kept and prior assumptions regarding individual based variation in traits.

For the SIR model individuals are classified as being either susceptible to infection (S), infected and infectious (I), or recovered/removed/dead (R). The time-dependent force of infection for a susceptible individual *j* (*i.e.* the probability per unit time of becoming infected) is given by $\lambda_j(t)$. For those individuals which do become infected, the distribution in the duration of the disease is assumed to be gamma distributed with individual-based mean w_m and shape parameter *k*.

In some circumstances individuals do not recover from disease (*e.g.* bovine tuberculosis), so the recovery dynamics become redundant. This possibility can be selected by coosing the SI model in Fig. 3(A).

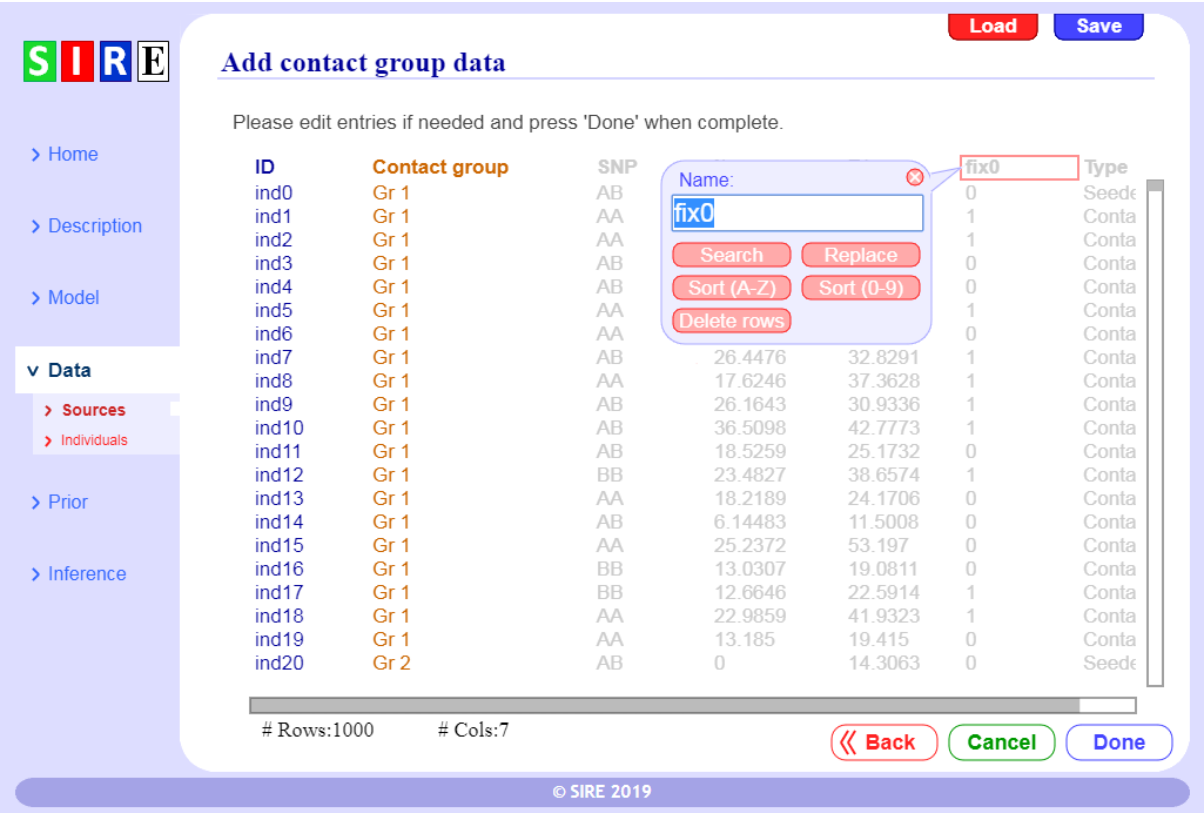
In most standard analyses individual-based variation is ignored. However in reality it may play an important role is determing disease dynamic behaviour. For this reason SIRE allow either option to be selected


SIRE
Load
Save
Data Sources
Home
Description
Model
Data
Sources
Individuals
Prior
Inference
Type
From
To
A
Contact Group

Data
SNP

Data
Fixed Effect

Data
Infection Times
0
36.5098
Data
Recovery Times
3.41891
53.197
Data
B
Contact Group
Disease Status
SNP
Diag. Test
Fixed Effect
Infection Times
Recovery Times
Next
© SIRE 2019



The downloaded folder contains the subdirectory “Datasets”.

3.1

Parameter	Description
β	Population average contact rate.
γ	Population average recovery rate.
k	Shape parameter that characterises the gamma distributed infection duration.
λ_j	Force of infection (probability per unit time for individual j to become infected).
w_j	Mean recovery time for individual j.
g_j, f_j, r_j	Fractional deviation in susceptibility, infectivity and recoverability of individual j.
$g_j^{SNP}, f_j^{SNP}, r_j^{SNP}$	SNP-based contribution to g_j, f_j, r_j.
a_g, a_f, a_r	SNP effects, *i.e.* change in g_j, f_j, r_j coming from an A allele compared to a B allele.
$\Delta_g, \Delta_f, \Delta_r$	Scaled dominance factors (1 when A is completely dominant over B).
$\boldsymbol{X}$	The design matrix for fixed effects.
$\boldsymbol{b_f}, \boldsymbol{b_f}, \boldsymbol{b_r}$	Vectors of fixed effects for the three traits.
$\boldsymbol{\varepsilon_g}, \boldsymbol{\varepsilon_f}, \boldsymbol{\varepsilon_r}$	Residual contributions to $\boldsymbol{g}, \boldsymbol{f}, \boldsymbol{r}$ (coming from sources other than the SNP).
$\boldsymbol{\Sigma}$	Covariance matrix of residual contributions.
G_z	Group effects (accounts for differences in transmission rates in different groups).
σ_G	Standard deviation in group effects.
θ	Model parameters.
ξ	Event data (infection and recovery times).

H_{seed}, H_{cont}	Proportion of homozygotes (*i.e. AA* or *BB*) in the seeders and contacts.
χ_{seed}, χ_{cont}	Homozygote balance (*i.e.* the proportion of *AA* individuals minus the proportion of *BB* individuals) in the seeders and contacts.
N_{group}	Number of contact groups.
N_{seed}	Number of seeders (initially infected individuals) in each contact group.
N_{cont}	Number of contacts (initially susceptible individuals) in each contact group.
G_{size}	Total number of individuals per group $G_{size} = N_{seed} + N_{cont}$.
N_{total}	Total number of individuals $N_{total} = N_{group} \times G_{size}$ in the experiment.
N_I	Total number of observed infection events in the experiment.
ϕ	The average fraction of contacts that become infected.
h	The proportion of infections that occur during initialisation of contact groups.
$\boldsymbol{M}$	Observed Fisher information matrix.

Table 1. A description of key quantities used in the paper.

1.3 Loading and Saving

SIRE permits users to load and save analysis in a special ".sire" format which save the data along with everything required for analysis. This is useful not only because it allows for analysis options This

Exporting

Posterior distribution graphs can be exported from SIRE and also files containing posterior samples of θ and ξ for further analysis using other tools. The user guide for SIRE is available in the electronic supplementary material and on the website.