DATA MANAGEMENT PLAN

Types of data:
This project generates no novel data, but will yield R packages to implement analyses or graphics. Public data analyzed in our study will be built into these packages as worked examples. The reproducibility of simulation and analyses in this project will be ensured by making all data and source code associated with them (along with detailed documentation) publicly available under the GNU GPLv3 license.

Policies for data sharing:
We will share all data in public data repositories (e.g. Github) upon submission for publication. We will include R code for any analyses alongside the data tables, to facilitate replication of any particular results. Data analysis pipeline in this project will also be included as tutorials in R packages developed for this project.

Policies for software and data and distribution:
We encourage colleagues to use our software package for their own analyses and publications, either using our example data, or using external data for re-analysis or meta-analysis. We commit to sharing statistical or graphical code in a timely manner upon request. We do not expect co-author status on papers arising from this data/software re-use, except where we become deeply involved in the conceptual design of the derived study, analysis, or writing.

Data backup: All files will be backed up in real-time, both on cloud storage (GitHub) and monthly on hard drives kept in separate buildings from the computers holding data for analysis, to ensure the data cannot be lost or damaged without having a recoverable copy in another location.