Reproducible Research with *R*, LaTeX, & Sweave

Theresa A Scott, MS[1]★, Frank E Harrell, Jr, PhD[1]★

1. Vanderbilt University School of Medicine, Department of Biostatistics
★Contact author: theresa.scott@vanderbilt.edu

Keywords: reproducible research, literate programming, statistical reports

During this half-day tutorial, we will first introduce the concept and importance of reproducible research. We will then cover how we can use *R*, LaTeX, and Sweave to automatically generate statistical reports to ensure reproducible research. Each software component will be introduced: *R*, the free interactive programming language and environment used to perform the desired statistical analysis (including the generation of graphics); LaTeX, the typesetting system used to produce the written portion of the statistical report; and Sweave, the flexible framework used to embed the *R* code into a LaTeX document, to compile the *R* code, and to insert the desired output into the generated statistical report. The steps to generate a reproducible statistical report from scratch using the three software components will then be presented using a detailed example. The ability to regenerate the report when the data or analysis changes and to automatically update the output will also be demonstrated. We will show more advanced usages and customizations of Sweave for making more beautiful reports and for making research more reproducible, as well as discuss differences between Sweave and knitr. In addition, the tutorial will provide a hands-on session during which a practice report will be generated and modified, useful tips and needed resources/references.