CSEB Workshop

Atul Sharma MD, MSc, M.Stat, FRCP(C)

June 8, 2016

Propensity Scores: Making Sense of Non-Randomized Observational Data

You will need to install R and RStudio prior to the workshop. For installation details under Windows, see A (very) short introduction to R by Torfs and Brauer. Installation is similar for Mac or Linux.

A number of R packages will be used in this lab. To view the packages in your base installation, run the following command at the R prompt:

```
library()
```

The following add-on packages will be needed: `survival`, `Matching`, `rbounds`, `pROC`, `plotrix`. If not part of your current installation, they can be installed manually from the R prompt via

```
install.packages(c("suvival", "Matching", "rbounds", "pROC", "plotrix"))
```

The CSEB conference website includes a link to the ancillary materials needed for this workshop. You should take a minute to download the *code, dictionary, and data set* before you arrive, as you may not have wireless internet access on the U Manitoba campus without a UMNET ID. In the folder **ComputerLab**, you will find a data set in the form of a comma separated variable or .csv spreadsheet called *RHC.csv*. These data are described in the included reference: *The Effectiveness of Right Heart Catheterization in the Initial Care of Critically Ill Patients* by Connors *et al*, JAMA (1996). A formal data dictionary is included in the file *RCHCodebook.rtf*. Copies are also available at my website by following the CSEB workshop link.

The R code for our hands-on computer lab is included in the file *RCH.R*. To work through the examples, you will need to first set your working directory to the folder **ComputerLab** using the **Session-Set Working Directory** menu in RStudio or manually via the `setwd()` command-line.

As we will see, R is particularly rich in packages for matching, assessing post-match balance, and performing sensitivy analysis (Matching, rbounds). Stata is another good choice (ps2match, rbounds). For those of you new to R, you may wish to consider Dr. Hilbe's advice:

authors of research articles in scientific journals now appear to overwhelmingly employ R for executing and displaying published statistical results. Joseph M. Hilbe, Journal of Statistical Software, Sept 2010

R binaries and add-on packages for Windows, Mac, and Linux are freely available from the Comprehensive R Archive Network or CRAN. There are also many on-line training materials and textbooks for R, including Rob Kabacoff's *excellent* Quick-R site. Specific questions may be addressed through several on-line support forums, such as Sasha Goodman's Rseek or John Baron's R Site Search. CRAN also provides a brief tutorial introduction to R by Venables and Smith, which I included in the **R-resources** folder.

For the record, this is an R Markdown document. R Markdown is a 'literate' formatting syntax designed to encourage reproducible research by embedding executable code chunks in HTML, PDF, or Word documents. Yihui Xie's `knitr` package supports R, SAS, and python code chunks with a variety of output formats.

Please contact me if you have questions at atul.sharma@umanitoba.ca.