Working course outline:

1. Introduction
   1. Getting people set up in Rstudio (30 min), including downloading the course files
   2. Conceptual introduction to programming (within Rstudio)
   3. Sequencing operations
   4. Reproducibility in data analysis – pull down sample Rmarkdown, then modify code
2. End to end on clean data set – method validation (Dan and Nik)
   1. Importing data
      1. Readr – csv and tsv
      2. Readxl – multi-tabbed spreadsheet
   2. Data structure – simple data frames/tibbles
   3. Plotting with ggplot2
   4. Exercises
      1. Histogram of values – normally distributed?, reference ranges
      2. Method comparison – correlation
      3. Precision – QC data
      4. Co-linearity (same data set), linearity – need additional calibration data
3. Exploratory data analysis
   1. Diving into a new data set (Patrick)
      1. Loading data & understanding distributions
         1. Understanding variable types after import
         2. Cleaning up variable names
         3. Utility of factors
         4. Using mutate
         5. Skimr as a handy tool
      2. Tabulating and visualizing data
         1. Filter to select rows
         2. Base table function for quick tabulations
         3. More advanced tables (tabyl from janitor package)
         4. Visualizing distributions (bar charts and histograms)
         5. Manipulating dates and visualizing time-course data
   2. Beyond basic summaries (Joe)
      1. Working with timestamps
      2. Summarizing data at a patient level
      3. Merging data sets

Action items, 7/17:

All

* Standardize formatting throughout lessons
  + Use bold **Start exercise** and **End exercise** to indicate exercises
  + Standardize ggplot formatting
  + Setting up environment – code chunks
* Reorder lessons – 01, 02a, 03a, 02b, 02c, 03b
* Push up to master

Patrick

* Find other contact at AACC to communicate to course attendees
* Update readme to upload to presentation

Action items, 7/5:

Deadline: preliminary draft of course materials completed by 7/18, will provide a link to course attendees to access pdf’s for printing

All

* Review and modify survey for course attendees to complete prior to attending
* Plan to create two documents for every lesson – start with answers document and when final post version without answers

Patrick

* Reach out to AACC to get attendee list
* Push exploratory data analysis changes to master for Joe to work from

Action items, 6/21:

Patrick

* Share copy of old survey to modify for AACC
* Add piping explanation to intro section
* Add updated data set to repo (master)
* Add course outline/action items to repo

Nik

* Review method validation & exercises and select sections to work on

All

* Add package installs within lessons (for non-tidyverse packages)

Action items, 6/7:

* Joe to add R Markdown document creation to intro
* Patrick to build out more data exploration content
* Patrick to identify exploratory analysis pieces for Joe to work on
* Dan to add data loading content & method validation
* Everyone to build out 30-45 min chunks of content

Action items, 5/10:

* Joe to start writing up intro materials
* Dan to start writing solutions for method validation/verification data
* Patrick to finish cleaning data set & get sign off from compliance to use data for teaching purposes
* Patrick to set up Github site for course
* Will need to reschedule meeting the week of 5/21 – late afternoon on Wednesday?

Action items, 4/26:

* Dan to look at existing method validation materials and share
* Patrick to review existing orders data sets for good candidates for exploratory data analysis
* Set up recurring meeting schedule for every 2 weeks