Scoring

- Open one of the provided zip folders with 24 randomly assorted photos
  - The order of these photos correspond with the order of the provided histology
- Examine one picture at a time and give each penis a score, record the score in a spreadsheet (e.g., excel)
- In preparation for the second scoring attempt re-randomize the set of photos
  - Keep track of which new numbers correspond with the old numbers
- Within a week, but at least 24 hours later) rescore the newly randomized set of photos in a different sheet of the workbook
- Unrandomize the second scoring attempt so that data corresponds to histology data
  - Order the scores by the old identifying numbers
- Compile the two scoring attempts into one spreadsheet
- Copy and paste the histology ratio column provided in the sheet below into the spreadsheet
  - 1.1, 1.2, 1.3 ... corresponds to training set one picture 1, 2, and 3 etc.
  - 2.1, 2.2, 2.3 ... corresponds to training set two picture 1, 2, and 3 etc

Analysis

- Run analyses in statistical program of preference (R, SPSS, etc)
  - Pearson’s Correlation on the two scores. (Precision)
    - $r > .8$ (High correlation)
  - Paired T-test comparing two scoring attempts. (Bias)
    - $P$-value $>.05$ (No difference)
  - $R^2$ of histology ratio averaged across given score from second or last attempt. (Accuracy)
    - $R^2 > .95$ (high explanatory value)
- If all thresholds are met, then scoring is considered both precise and accurate, and genitalia abnormality severity can be scored in experimental animals
- If failed restart process with second scoring set

***If training attempts exceed two, the two sets can be combined and 24 random pictures can be picked from the 48