2017 PISA

Pathobiology for Investigators, Students and Academicians

Lunch & Learn: Science, Statistics and Getting it Right

A Workshop Sponsored by
ASIP Committee for Career Development & Diversity
and ASIP Education Committee

Dan A. Milner, Jr.
American Society for Clinical Pathology, Chicago, Il

September 26, 2017 • 12:30 PM – 1:25 PM

Vignette 1
Vignette 2

Vignettes Edited by Mark E. Sobel, MD, PhD, ASIP, Rockville, MD

ASIP

American Society for Investigative Pathology

American Society for Investigative Pathology
Investigating the Pathogenesis of Disease

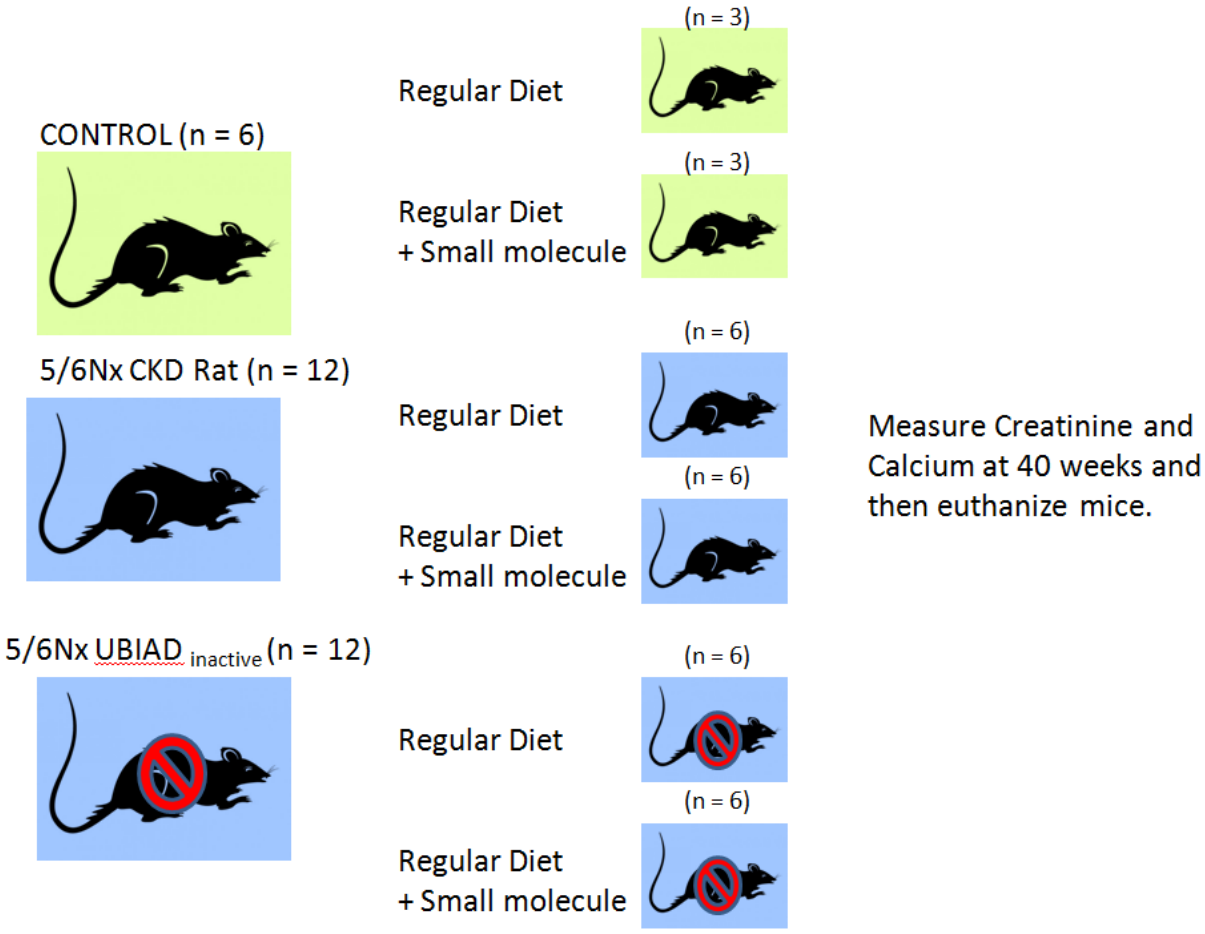
LUNCH & LEARN: SCIENCE, STATISTICS, AND GETTING IT RIGHT

PISA 2017 WORKSHOP

Vignette 1

A small molecule screening program using human endothelial cell lines that develop calcification in the presence of calcification-inducing media identifies a molecule that drastically reduces the process. Interestingly, this molecule has extremely high oral bioavailability and is eliminated by direct excretion in the urine. The working biological hypothesis is that the molecule enhances the activity of UBIAD1 (an intracellular cholesterol regulator). Your laboratory has a working model of the 5/6Nx rat chronic kidney disease system and a collaborator happens to have a CRISPR/CAS9 tool to replace the UBIAD1 gene with an inactive form of the protein. In your system, the 5/6Nx rats develop chronic kidney disease including vascular calcifications and you monitor the disease using a peripheral blood measure of creatinine prior to euthanasia. In your collaborator's system, serum calcium is elevated in the UBIAD1 inactive form in normal rats.

You design an experiment to test the small molecule in your system as follows: At the end of 40 weeks, you measure the creatinine and calcium (see graph on the next page) of all of the rats and then sacrifice them. Using histology (H&E and Von Kassa stain) along with ImageJ (a free software program that allows you to do image based analyses, such as count cells or parse out a specific feature (nuclei, cytoplasm, etc) – download from http://imagej.nih.gov/ij/). You quantify the amount of calcification in the kidneys and the heart (see graph on the next page).

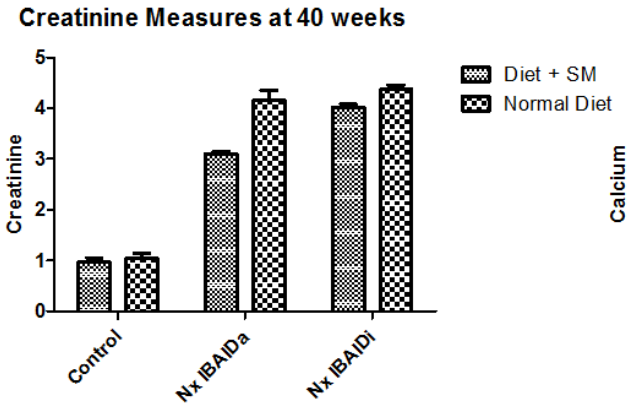


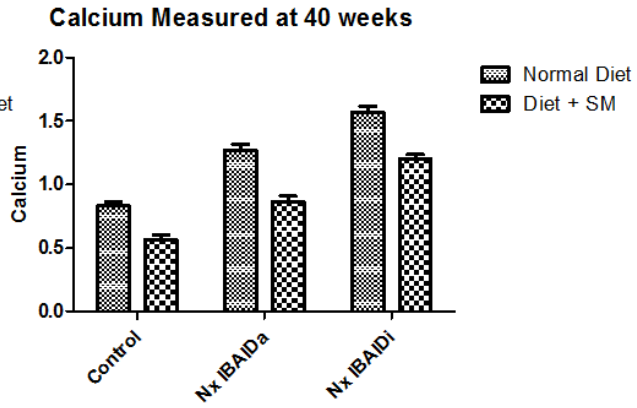
All rats are fed a 2% Ca, 1% P diet and followed for 40 weeks

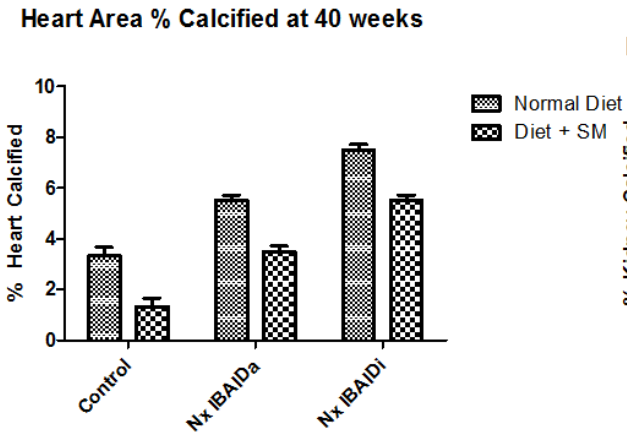
Co-sponsored by the ASIP Committee for Career Development & Diversity and the ASIP Education Committee

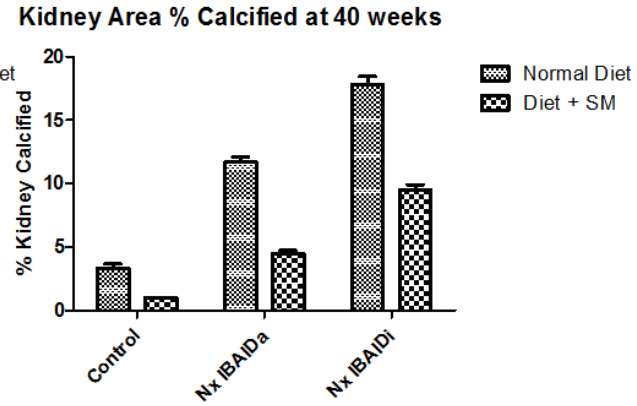
LUNCH & LEARN: SCIENCE, STATISTICS, AND GETTING IT RIGHT

PISA 2017 WORKSHOP









Questions:

1. Are the differences in measurements of creatinine, calcium, and tissue calcification different between the groups? Which ones?
2. When approaching data such as this, a few questions need to be answered prior to beginning any analysis (and should best be thought of before designing the experiment!). These include the following:
    i. What kind of variables do I have?
    ii. What kind of statistical test(s) can I perform?
    iii. What kind of result am I looking for?

Co-sponsored by the ASIP Committee for Career Development & Diversity and the ASIP Education Committee

LUNCH & LEARN: SCIENCE, STATISTICS, AND GETTING IT RIGHT

PISA 2017 WORKSHOP

Vignette 2

A group of 1175 healthy subjects (43% Caucasian, 33% African or African American, 24% Hispanic/Latino) were recruited from college campuses in the Boston area (from among 26 different colleges) and were asked to provide a buccal swab for DNA sequencing along with a detailed questionnaire regarding their family history and medical health as well as a tube of blood for laboratory testing. They also agreed to complete a follow up survey every 5 years for the next 25 years in order to look at new diagnoses and diseases. For each patient, an aliquot of blood as well as the buccal swab were both used to sequence each patient to 40X coverage as well as perform comparative genomic hybridization to a sequenced and assembled reference genome. All genomes were cataloged for mutations included insertions/deletions, single nucleotide polymorphisms, and gene duplication. The survey included questions about all of the following: diabetes, hypertension, malignancy (specifically of breast, lung, colon, prostate, kidney and/or brain), infections (including frequency and specifically for mononucleosis, ear infections, head colds, urinary tract infections, toenail infections, persistent/excessive acne), diet, and exercise habits. All of the subjects were counseled to use a free pedometer (provided by the study team) which was connected to the internet and report their daily activity, which was monitored by the study.

After 10 years (3 total surveys), a manuscript was published by a non-competing group in a mouse model showing that a specific mutation of pyruvate dehydrogenase kinase 4 (PDK4) caused a massive decrease in mouse activity as well as obesity in mice. You propose to look at the pedometer data of the study's subjects' activity to see if there is an association with fewer steps and mutations in PDK4. Your PI, however, thinks that such an association may be polygenetic (or even spurious in the mouse) and the entire genome should be examined in the context of all of the data.

Questions:

1. How would you go about investigating any potential associations in your data set?
2. What statistical considerations are important in thinking about this question?
3. How should the pedometer data be parsed for the analysis?

Sponsored by the ASIP Committee for Career Development & Diversity and the ASIP Education Committee