AADOM PRE-SESSION

David P. Yens, Ph.D.
Research Director, NYCOMEC
Research Director
Touro –COM Middletown
David.yens@touro.edu

49 Shades of Research

RESEARCH DESIGNS:
VALIDITY OF RESEARCH SUBJECTS

APRIL 21, 2015
...well, the good news is that we'll both make it into the Guinness Book of World Records!
Idea Development – Pair and Share, Part 2

What research has been done in this area? What does it say?

What variables would you use or measure?
   Independent variable(s)
   Dependent variable(s)
   Moderator or control variable(s)

How would these variables be defined and measured?
   Independent variable(s)
   Dependent variable(s)
   Moderator or control variable(s)

What research design would you use?
   Who will be the subjects?

What statistical analyses might be used?

What outcomes would you predict?
The Spectrum of Biostatistics

Research Designs

Explanatory
- Experimental
  - Clinical Trials
  - Controlled Trials
- Observational
  - Cohort Studies
  - Follow-up
  - Case Control
  - Cross-Sectional

Descriptive
- Case Reports
  - Mostly narrative
  - May contain frequencies
- Population Studies
- Prevalence
- Incidence

Epidemiology
- Screening Tests
- Sensitivity
  - Specificity
  - Positive Predictability
  - Negative Predictability

Samples
- Sample Estimation
**Research hypothesis**

* What outcome do you expect?
  * Trained physicians will be perceived as providing more compassionate care than those not trained based on survey results

**Null Hypotheses - what is tested**

* The outcome if there is no difference
  * Trained physicians will be perceived as providing the same level of compassionate care as those not trained based on survey results
* IDENTIFY AND DEFINE VARIABLES

* In what variable do you want to make a difference?
  * DEPENDENT VARIABLE

* What variables will you manipulate?
  * INDEPENDENT VARIABLE

* NOTE: Variables must be defined precisely to avoid confusion
EXPLANATORY DESIGN TYPES

- Experimental Design
  - Crossover designs
  - Double-blind
  - Placebo control

- Cohort Study / Follow-Up Design

- Historical Cohort / Retrospective Follow-up

- Case Control

- Cross-Sectional
Experimental / Interventional / Clinical Studies

* These are studies where an experimental intervention is introduced to a group of subjects to determine the efficacy of certain procedures or treatments.

* This type of study is also called a trial, and can be either controlled or uncontrolled. Due to the nature of this study design, trials are always prospective.
CONTROLLED CLINICAL TRIALS

* At least two groups are compared to determine the efficacy of an intervention.
* The experimental / treated group receives the intervention while the control / placebo group does not.
* Randomized Clinical Trial: Subjects are randomly assigned to either the experimental group or the control group. Apart from the intervention, these two groups are treated equally. Thus, any differences between the two groups can be attributed to the intervention and not other factors.
* This method provides the strongest evidence for determining cause and provides the best evidence that the result was due to the intervention.
Figure 2-1. Experimental or controlled trial design.
Study Objective: The purpose of this study was to evaluate the efficacy of osteopathic manipulative treatment (OMT) as administered in the emergency department (ED) for the treatment of patients with acute ankle injuries.

Methods: Patients aged 18 years and older with unilateral ankle sprains were randomly assigned either to an OMT study group or a control group. Independent outcome variables included edema, range of motion (ROM), and pain. Both groups received the current standard of care for ankle sprains and were instructed to return for a followup examination. Patients in the OMT study group also received one session of OMT from an osteopathic physician.
* Osteopathic Manipulative Treatment in the Emergency Department for Patients With Acute Ankle Injuries
Anita W. Eisenhart, DO; Theodore J. Gaeta, DO, MPH; David P. Yens, PhD

**Results:** Patients in the OMT study group had a statistically significant \((F = 5.92, P = .02)\) improvement in edema and pain and a trend toward increased ROM immediately following intervention with OMT. Although at follow-up both study groups demonstrated significant improvement, patients in the OMT study group had a statistically Significant improvement in ROM when compared with patients in the control group.

**Conclusions:** Data clearly demonstrate that a single session of OMT in the ED can have a significant effect in the management of acute ankle injuries.
* WHAT WOULD BE A REASONABLE RESEARCH HYPOTHESIS FOR THIS STUDY?

* A. Standard Treatment is better than OMT
* B. There is no difference between the treatments
* C. The OMT treatment is better than the standard treatment
* D. None of the above
* Remember that precision is important:
* In what characteristics are things better or worse?
  * Range of motion
  * Pain
  * Edema
* OMT plus standard treatment of a sprained ankle will result in an increase in range of motion, decreased pain, and decreased edema within one hour when compared with standard treatment.
Remember that the null hypothesis is what we actually test with statistics.

It should be the opposite of the research hypothesis.

One hour after treatment of a sprained ankle there will be no difference in range of motion, pain, and edema due to OMT treatment compared with standard treatment.
Now, what are the variables?

**DEPENDENT VARIABLE(S)?**

- Range of motion
- Pain
- Edema

**INDEPENDENT VARIABLE?**

- Treatment (OMT plus standard treatment vs standard treatment alone)
*MODERATOR OR CONTROL VARIABLES?
Bielec, G., et.al. Effectiveness of basic life support instruction in physical education students-a pilot study. Teaching and learning in medicine, 26(3), 252-257.

Second-year physical education students were randomly assigned to three groups to learn basic life support (BLS) knowledge and skills: computer-assisted presentations, practicing the BLS algorithm in pairs, and learning on their own. A 10-question multiple-choice test and performance of BLS on a mannequin were used to assess learning. No significant differences were found.
The purpose of this study was to determine whether an abridged mindfulness-based stress reduction (MBSR) intervention can improve measures of wellness in a randomized sample of 1st-year medical students. **Methods:** Fifty-eight participants were randomized to control or 8-week MBSR intervention and then invited to participate in the study. All participants were assessed using the Perceived Stress Scale (PSS), the Resilience Scale (RS), and Self-Compassion Scale (SCS) at 3 separate time points: baseline, at the conclusion of the study intervention (8 weeks), and at 6 months after the conclusion of the intervention. The intervention consisted of 75 minutes of weekly class time, suggested meditation at home, and a half-day retreat in the last week. **Results:** The intervention group achieved significant increase on SCS scores both at the conclusion of the study (0.58, *p* = .002), 95% confidence interval (CI) [0.23, 0.92], and at 6 months (0.56, *p* = .001), 95% CI [0.25, 0.87]. PSS scores achieved significant reduction at the conclusion of the study (3.63, *p* = .03), 95% CI [0.37, 6.89], but not at 6 months poststudy (2.91, *p* = .08), 95% CI [-0.37, 6.19]. The study did not demonstrate a difference in RS after the intervention, though RS was significantly correlated with both SCS and PSS. **Conclusions:** An abridged MBSR intervention improves perceived stress and self-compassion in 1st-year medical students and may be a valuable curricular tool to enhance wellness and professional development.
*Let’s look at some other designs*
Observational

* Cross-sectional
* Case Control
* Cohort Studies
* Follow-up
These studies analyze data collected on a group of subjects at a given point in time rather than over an extended period of time.

Designed to determine what is happening presently, not what happened in the past.

Frequently used for surveys.
Measure/Classify and Compare

Begin

Study Population

Free of Disease/Outcome

Have Disease/Outcome

Risk/Factor +
Risk/Factor -
Risk/Factor +
Risk/Factor -
Context: The academic credential awarded to osteopathic physicians is the doctor of osteopathy or doctor of osteopathic medicine (DO) degree. Public recognition of the degree has been disappointingly low, however, leading some members of the profession to argue for a change in the degree’s name and formal designation.

Objectives: To investigate antecedents to the desire among osteopathic medical students to change vs retain the DO degree designation and maintain “the DO difference.”

Methods: A self-administered cross-sectional 38-item electronic survey was distributed to 480 students at an osteopathic medical school in the Midwestern United States. The instrument included knowledge-based items about osteopathic principles and practice (OPP) as well as items designed to assess attitudes, subjective norms, perceived behavioral control, and intention to support a proposed degree change.
Results: An overall response rate of 45% was achieved (n=214).

Structural equation modeling revealed that low levels of OPP knowledge were associated with positive attitudes and subjective norms favoring a degree change with the reverse true for opposing students. Knowledge did not influence perceived behavioral control. Attitudes were the best predictor of intention to vote with 85% variance predicted in our models; perceived behavioral control was the best predictor of intention to debate with approximately 38% variance observed.

Conclusions: As a result of diminished use of palpation and osteopathic manipulative treatment—two historic markers of professional identity among osteopathic physicians—the DO degree designation as an indicator of difference has received increasing scrutiny. Improved student awareness of OPP is essential to maintaining the DO difference in clinical practice and with regard to the DO degree designation.
The purpose of this study was to investigate differences in medical student and faculty attitudes regarding preclinical classroom attendance and the impact of nonattendance on educators and the learning environment. **Methods:** Using Internet-based surveys, we assessed attitudes about preclinical classroom attendance among medical students and teaching faculty at Washington University School of Medicine. Our primary hypothesis was that students would be less likely than faculty to place societal value on attendance and relate it to professionalism. **Results:** Both groups recognized a negative impact of poor attendance on faculty enthusiasm for teaching (students 83%, faculty 75%), but faculty were significantly more likely to endorse a negative impact on effectiveness of lectures (75% vs. 42%, \( p < .0001 \)) and small-groups (92% vs. 76%, \( p < .0001 \)) and a relationship between attendance and professionalism (88% vs. 68%, \( p < .0001 \)). Students were significantly more likely to support free choice among learning opportunities (90% vs. 41%, \( p < .0001 \)) including regularly missing class for research and community service activities (70% vs. 14%, \( p < .0001 \)) and to consider lecture videos an adequate substitute for attendance (70% vs. 15%, \( p < .0001 \)). Free-text responses suggested that students tended to view class-going primarily as a tool for learning factual material, whereas many faculty viewed it as serving important functions in the professional socialization process. **Conclusions:** In this single-center cohort, medical student and teaching faculty attitudes differed regarding the importance of classroom attendance and its relationship to professionalism, findings that were at least partially explained by differing expectations of the purpose of the preclinical classroom experience.
Follow-up Study Design

BEGIN

Study Population

- Free of Disease
  - Risk Factor (+)
  - Have Outcome Already (exclude)
  - Risk/Factor (-)

Measure/Classify

- Measure Outcome/Compare
  - Disease/Outcome (+)
  - Disease/Outcome (-)
  - Disease/Outcome (+)
  - Disease/Outcome (-)
Abstract

Current data suggest that increases in hemoglobin may decrease nitric oxide and adversely affect vascular function. In the preclinical setting, these changes could precipitate the development of heart failure (HF). We hypothesized that higher hematocrit (HCT) would be associated with an increased incidence of new-onset HF in the community. We evaluated 3,523 participants (59% women) from the Framingham Heart Study who were 50 to 65 years old and free of HF. Participants were followed prospectively until an HF event, death, or the end of 20 years of follow up. HCT was subdivided into 4 gender-specific categories (women: HCT 36.0 to 40.0, 40.1 to 42.0, 42.1 to 45.0, >45.0; men: 39.0 to 44.0, 44.1 to 45.0, 45.1 to 49.0, >49.0).
Gender-pooled multivariable Cox proportional hazards models were used to estimate the association of HCT with incident HF, adjusting for clinical risk factors. During the follow-up period (61,417 person-years), 217 participants developed HF (100 events in women). There was a linear increase in risk of HF across the 4 HCT categories (p for trend = 0.002). Hazards ratios for HF in the low-normal, normal, and high HCT categories were 1.27 (95% confidence interval 0.82 to 1.97), 1.47 (1.01 to 2.15), and 1.78 (1.15 to 2.75), respectively, compared to the lowest HCT category (p for trend <0.0001). Adjustment for interim development of other cardiovascular diseases and restriction of the sample to nonsmokers did not alter the results. In conclusion, higher levels of HCT, even within the normal range, were associated with an increased risk of developing HF in this long-term follow-up study.
* The experimental and control groups are matched as closely as possible with regard to age, sex, geographic location, or any other factor that may have influenced the variable being examined.
Figure 2-2. Case-control design.
The authors undertook a case control study to determine whether hospitalized patients with pneumonia had reflex points in the anterior chest wall as described by Frank Chapman, DO, specifically those classified as relating to the lung. Sixty-nine hospitalized patients were enrolled in the study. Patients with an admitting diagnosis of pneumonia were compared to those without pneumonia as their admitting diagnosis. All patients were examined to determine if Chapman reflex points for the lungs were present. The study controlled for potential confounding diagnoses by excluding patients with lung pathology other than pneumonia. Results demonstrated a statistically significant relationship between the presence of Chapman reflex points and pneumonia in hospitalized patients.
Abstract

In 1951 the British Medical Association forwarded to all British doctors a questionnaire about their smoking habits, and 34,440 men replied. With few exceptions, all men who replied in 1951 have been followed for 20 years. The certified causes of all 100,720 deaths and subsequent changes in smoking habits were recorded. The ratio of the death rate among cigarette smokers to that among lifelong non-smokers of comparable age was, for men under 70 years, about 2:1, while for men over 70 years it was about 1.5:1. These ratios suggest that between a half and a third of all cigarette smokers will die because of their smoking, if the excess death rates are actually caused by smoking. Smoking caused death chiefly by heart disease among middle-aged men (and, with a less extreme relative risk, among old men,) lung cancer, chronic obstructive lung disease, and various vascular diseases. The distinctive features of this study were the completeness of follow-up, the accuracy of death certification, and the fact that the study population as a whole reduced its cigarette consumption substantially during the period of observation. As a result lung cancer grew relatively less common as the study progressed, but other cancers did not, thus illustrating in an unusual way the causal nature of the association between smoking and lung cancer.
In this study, investigators examined more than 465,000 babies born over 13 years to mothers who were patients in the Kaiser Permanente health care system in northern California. The study found that birth defects occurred at the same rate among all women with high blood pressure, regardless of whether they took angiotensin-converting enzyme (ACE) inhibitors, other drugs used to treat high blood pressure, or no blood pressure drugs during the first trimester of pregnancy. This finding suggests that the underlying high blood pressure itself may increase the risk of birth defects, rather than ACE inhibitors. ACE inhibitors had been linked to birth defects in a previous study published in 2006 that examined nearly 30,000 births over 15 years to mothers enrolled in the State of Tennessee Medicaid system.
Case Study: A simple descriptive report of interesting characteristics of one unique subject. This type of study takes place over a relatively short period of time, is conducted by a doctor or scientist, and is considered to be anecdotal evidence.

Case Series: A type of case study that reports on at least five cases. This type of study is useful for indicating that a trend may be present and deserves more formal study.
* In God we trust.
* All others must bring data.
WE NEED SUBJECTS (PATIENTS)

- Nature of Subjects
- Number of Subjects
- Effect Size
- Selection Procedures
- Sampling and Randomization
* **Probability Sampling**
  * Simple Random
  * Stratified Random
  * Cluster

* **Nonprobability Sampling**
  * Consecutive
  * Convenience
  * Judgmental
NUMBER OF SUBJECTS

* Paper, Pencil, & Tables
* Computer Programs
The EFFECT SIZE is like a z-score; it tells you how large the difference is in sd units. It’s used in computing the sample size for an experimental study, but it is also used to compare the effects of multiple studies — it permits a basis of comparison.

Different computations are used for each statistical method.
Sample size computation differs for each statistical analysis

Easiest way - use G-Power (free from the Internet)

A general power analysis program, free for Mac and PC. Written by E. Erdfelder, F. Faul and A. Buchner, at the University of Trier.

wwwpsychouni-duesseldorfdetaapprojectsgpower
INTERNAL AND EXTERNAL VALIDITIES
What is validity?

* How well the measurement represents the phenomena of interest

* Internal validity - Did, in fact, the experimental treatment make a difference in this specific experimental instance? The basic minimum without which any experiment is uninterpretable.

* External validity - the extent to which the results are generalizable or applicable to a particular target population
Internal Validity
- History
- Maturation
- Testing
- Instrumentation
- Statistical Regression

Selection
- Experimental mortality
- Stability
- Expectancy
- Interactive Combinations
Internal Validity (continued)

- Compensatory equalization
- Compensatory rivalry
- Resentful demoralization
- Diffusion/initiation of treatments
* Construct Validity

* Inadequate preoperational definitions
* Hypothesis guessing
* Evaluation apprehension

* Experimental expectancies
* Interaction of different treatments
* Interaction of testing & treatment
Experimental Validity

**Statistical Conclusion Validity**

- Low statistical power
- Violated assumptions of statistical tests
- Fishing & error rate
- Reliability of measures
- Reliability of treatment implementation
- Random irrelevancies in setting
- Random heterogeneity of subjects
External Validity

* Interaction of selection and treatment
* Interaction of setting and treatment
* Interaction of history and treatment
<table>
<thead>
<tr>
<th>Pitfall</th>
<th>Explanation of the problems and examples</th>
<th>Recommended solutions</th>
<th>Good example from the literature</th>
</tr>
</thead>
</table>
| Using an inappropriate control condition   | When you compare an experimental condition with a control condition, then you can attribute differences in outcome to differences between the conditions. If these conditions vary on too many elements, it is impossible to attribute outcomes to a specific element.  
  - If you compare Web-based learning with lectures, which differ in many aspects (e.g., learning pace, interaction with peers and teachers), you won’t know which aspect(s) of these learning modes caused differences in outcomes. | - Identify the crucial element of your intervention.  
  - Make the experimental and control conditions as similar as possible, except for the crucial element. | Issa et al<sup>3</sup> compared a lecture that was designed according to multimedia principles with a lecture that was not designed according to these principles, but similar in every other aspect. |
| Failing to align your outcome measures to your research questions | Outcome measurements should reflect the dependent variable(s) stated in your research question(s). If your outcome measures do not match your theory, your results do not answer your research questions.  
  - If you expect that students learn communication skills better when they have contact with real instead of simulated patients, you should measure communication skills rather than knowledge or perceptions about communication. | - When designing a study, first clarify expected effects.  
  - Next, define how you can observe these effects.  
  - Then decide which instruments measure these effects. | Cook et al<sup>4</sup> operationalized their dependent variable (learning outcomes) with two test types: a post-test after each module and a cumulative test. |
| Ignoring possible reactive effects of a pretest | A pretest could provide information on baseline differences between participants. However, a pretest can cause participants to acquire relevant information. Therefore, the pretest can reinforce your intervention or have a direct effect on the dependent variable(s) that you measure with the posttest.  
  - If you ignore effects of a pretest that assesses prior knowledge, you won’t know whether your results can be attributed solely to your intervention. | In a nonrandomized design:  
  - Let students do an irrelevant task between the pretest and intervention.  
  - Use existing data (e.g., grades) as a pretest.  
  In a randomized design:  
  - Don’t use a pretest. | Hatala et al<sup>5</sup> randomly allocated students to one of two instructional approaches and didn’t use a pretest to investigate the superiority of one of these approaches. |
## Academic Medicine, continued

### Not taking time-on-task into account

It is likely that increased time spent on learning tasks yields increased learning outcomes. If you do not take this into account, it is impossible to attribute your outcomes solely to the variables you measured because they might be explained by differences in time-on-task as well.

- If you compare Web-based learning with lectures, the *time-on-task* is the actual time spent on the study activities.

### Confusing ecological and external validity

Ecologically valid experiments do not necessarily have high external validity. Ecological validity is the extent to which your study approximates the real world. It often introduces elements (e.g., teacher characteristics, motivation) that mask or change effects, which, in turn, may compromise the external validity or generalizability of your study.

- If you investigate effects of an individual assignment in a classroom setting, student interaction can influence the effect and thus compromise external validity.

### Design conditions so that participants spend the same amount of time on the task.

- Control for time-on-task in statistical analyses if there are differences between conditions.

### In Mamede et al. the time participants were allowed to spend on each study case was the same for all conditions.

### Focus on external validity instead of ecological validity.

- Achieve high external validity by conducting a well-controlled experiment that is repeated in different settings and populations.

### Marquard et al. investigated patient identification errors, controlling for the number and type of errors identified, during medication administration.
MORE QUESTIONS
What threats to validity might you encounter in your study?

<table>
<thead>
<tr>
<th>THREAT</th>
<th>HOW WOULD YOU HANDLE THE THREAT?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Also published as an independent monograph.
Selected References

• Stephen B. Hulley, et.al. Designing clinical Research, 4th Ed. Lippincott Williams & Wilkins. Considered the best reference for designing research.


"Could you try to write more clearly? Last time, the girl at the pharmacy slapped me!"
THANK YOU!
ARE THERE ANY QUESTIONS?