

CLRES 2040: Measurement in Clinical Research

Instructor: Galen E. Switzer, PhD
Associate Professor, Medicine, Psychiatry, Clinical and Translational Science
502 Iroquois Building
412-624-2530 (w); 412-731-2448 (h)
Email: switzerge@upmc.edu

Course Description:

This course focuses on defining and providing examples of measurement issues in clinical research. Theories of measurement, including classical test theory and generalizability theory will be discussed and will form the basis for examining the psychometric properties of instruments. However, the course will focus primarily on providing students with practical skills that will allow them to locate, select and evaluate instruments for use in clinical research. Key aspects of measure reliability – including inter- and intra-rater, test-retest, parallel/alternate forms and internal consistency – and validity – including content, criterion and construct – will be examined in detail. Additional topics include cultural issues in measurement, survey research methods, overview of health status measures and writing the measurement section of a grant proposal.

Course Objectives:

1. To recognize the universality and complexity of measurement issues.
2. To discuss measurement issues in specific areas of research interest.
3. To apply various methods of testing the psychometric properties of measures used in clinical research.
4. To describe and analyze the reliability and validity evidence for specific instruments
5. To conduct basic psychometric analyses, including internal consistency and factor analysis.

Course Requirements:

Prerequisites:

1. Basic research methods course; basic computer
2. Experience with a statistical program (e.g. STATA, SPSS or SAS)

Course Location:

Room 305 A/B Parkvale
Room 222 Parkvale for computer sessions

Course Credits and Contact Hours

1 credit; 16 contact hours; 1 session/week for 8 weeks (~2 hours each session)

Session Dates and Times

Session 1:	July 8, 2009	1:00-3:00
Session 2:	July 15, 2009	1:00-3:00
Session 3:	July 22, 2009	1:00-3:00
Session 4:	July 29, 2009	1:00-3:00
Session 5:	August 5, 2009	1:00-3:00
Session 6:	August 12, 2009	1:00-3:00
Session 7:	August 19, 2009	1:00-3:00
Session 8:	August 26, 2009	1:00-3:00

Grading:

Letter grade based on assignments and examination:

1. Summary of measurement issue in content area of interest (10%)
2. Full reliability and validity report on a single measure (25%)
3. Grant proposal descriptions of three measures (25%)
4. Final exam (40%)

Textbook:

1. Carmines, EG and Zeller RA. (1979). *Reliability and Validity Assessment*.
2. DeVellis, Robert F. (2003). *Scale Development: Theory and Applications*. Thousand Oaks, CA: Sage Publications.
3. Additional reading materials are provided online at the CourseWeb site:
<http://courseweb.pitt.edu>

Teaching Methodology:

Each in-class session will be a combination of didactic instruction, small group activities, individual exercises and/or computer based experience. Seminar participation, written assignments and a final exam are required work for this course.

Course Policies:

Academic Integrity:

Students in this course will be expected to comply with the University of Pittsburgh's Policy on Academic Integrity. Any student suspected of violating this obligation for any reason during the semester will be required to participate in the procedural process, initiated at the instructor level, as outlined in the University Guidelines on Academic Integrity. This may include, but is not limited to, the confiscation of the examination of any individual suspected of violating University Policy. Furthermore, no student may bring any unauthorized materials to an exam, including dictionaries and programmable calculators.

Disabilities:

If you have a disability that requires special testing accommodations or other classroom modifications, you need to notify both the instructor and the Disability Resources and Services office no later than the second week of the term. You may be asked to provide documentation of your disability to determine the appropriateness of accommodations. To notify Disability Resources and Services call 412-648-7890 (Voice or TTD) to schedule an appointment. The office is located in 216 William Pitt Union.

Copyright Notice:

These materials may be protected by copyright. United States copyright law, 17 USC section 101, et seq., in addition to University policy and procedures, prohibit unauthorized duplication or retransmission of course materials. See Library of Congress Copyright Office and the University Copyright Policy.

CourseWeb Accessibility:

CourseWeb is ADA Compliant and has fully implemented the final accessibility standards for electronic and information technology covered by Section 508 of the Rehabilitation Act

Amendments of 1998. Please note that, due to the flexibility provided in the product, it is possible for some material to inadvertently fall outside of these guidelines.

Session 1 (July 8th)

Focus:

Session 1 provides an overview of key steps in the measurement process, including building a conceptual model, operationalizing key variables, selecting measures and gathering data. This overview will provide the foundation for the remainder of the course.

Learning Objectives:

1. Introduction to the course
2. Variables and level of measurement
3. Overview of course content

Assignment:

1. Identify a current measurement issue in the published literature and write a short summary (~ 1 page) of the issue that includes something about the history of the construct to be measured and the current state of measurement of the construct. **Due July 15th, 2009 in class.**

Session 2 (July 15th)

Focus:

In Session 2 we will explore two concepts related to reliability -- consistency and absence of error. The concept of reliability focuses on the stability or repeatability of results. Questions related to reliability are "If blood pressure data are collected repeatedly from the same patient, will the same results be obtained?" or "If blood pressure is obtained by different medical staff, will different staff obtain similar readings?" In particular, this session will focus on inter-rater and intra-rater consistency.

Learning Objectives:

1. Define reliability and its relationship with measurement error.
2. Distinguish between inter- and intra-rater reliability.
3. Statistically assess reliability in a number of different ways.
4. Determine the appropriate measure of reliability of measurement of the key variables.

Readings:

1. DeVellis, pp. 1-13
2. Topf, M. (1986). Three Estimates of Inter-rater Reliability for Nominal Data. *Nursing Research*, 35(4), 253-255.
3. Roberts, I., Moran, K. (1995). Inter-rater Reliability in the Medical Diagnosis of Child Sexual Abuse. *Journal of Pediatrics and Child Health*, 31, 290-291.
4. Kroboth, FJ., Hanusa, BH., Hyg, MS., Parker S., Coulehan, JL., Kapoor WN., et al. (1992). The Inter-rater Reliability and Internal Consistency of a Clinical Evaluation Exercise. *Journal of General Internal Medicine*, 7, 174-179.
5. Viera, AJ, Garrett JM. (2005) Understanding Interobserver Agreement: The Kappa Statistic. *Family Medicine*; 37(5):360-3.

Due:

1. Copy of your measurement summary.

Session 3 (July 22rd)

Focus:

Session 3 continues to focus on reliability across time, across measurement versions and across indicators or measurement items. This session includes test-retest reliability, parallel/alternate form reliability and internal consistency.

Learning Objectives

1. Describe the full range of reliability evidence for a measure.
2. Identify potential sources of low and/or inadequate reliability
3. Compute and interpret consistency and coefficient using statistical software.

Activities:

1. Computer activity.

Readings:

1. Carmines and Zellar, pp 9-16, 29-51
2. DeVellis, pp 27-47
3. John OP., Bennet-Martinez, V. (2000). Measurement: Reliability, Construct Validation and Scale Construction. In (Ed.) H.T. Reiss & C.M. Judd, *Handbook of Research Methods in Social and Personal Psychology*. (pp 399-350). New York, NY: Cambridge University Press.

Session 4 (July 29th)

Focus:

In Session 4, we focus on validity, or the idea that our operationalized measure is an accurate reflection of the conceptual variable that we intended to assess. Reliability (discussed in Session 3) is a necessary, but not sufficient prerequisite for measurement validity. In other words, a measure can be reliable without being valid, but not vice versa. Validity is a complex concept and researchers take a wide variety of approaches to its assessment. Three general types of validity evidence are discussed in this session - content, criterion and construct validity.

Learning Objectives:

1. Name and describe the various types of validity evidence.
2. Assess or analyze types of validity such as face content, criterion and construct.
3. Determine whether the validity evidence for a particular measure is adequate.

Readings:

1. Carmines and Zellar, pp 17-27
2. DeVellis, pp 49-58
3. Edwards, R., Telfair, J., Cecli, H., Lenoci, J. (2000) Reliability and Validity of a Self-Efficacy Instrument Specific to Sickle Cell Disease. *Behavior Research and Therapy*, 38, 951-963.
4. Fogg, L., Gross, D. (2000) Focus on Research Methods: Threats to Validity in Randomized Clinical Trials. *Research in Nursing Health*, 23, 79-87.

Session 5 (August 5th)

Focus:

Session 5 focuses on validity assessment when a gold standard exists. In this context, positive and negative predictive values, sensitivity and specificity and responsiveness can all be assessed. We will discuss how to examine and estimate each of these measurement properties as well as the range of issues that may affect them. In addition, we will discuss the usefulness of Receiver Operating Characteristic curves in determining cutpoints for cases versus noncases. Finally we will consider the impact of cultural issues such as language, ethnicity, age, socioeconomic status and education level on measurement and validity.

Learning Objectives:

1. Analyze validity in the context of a gold standard.
2. Determine positive and negative predictive values for a measure.
3. Calculate a measure's sensitivity and specificity.
4. Examine the effect of changing cutpoints on ROC curves on the distribution of cases and noncases.
5. Develop guidelines for cross-cultural adaptation of measures.

Readings:

1. Guyatt, G., Walter, S., Norman, G. (1987). Measuring Change Over Time: Assessing the Usefulness of Evaluative Instruments. *Journal of Chronic Diseases*, 40(2), 171-178.
2. Hoffman, JR., Mower, WR., Wofson, AB., Todd, KH., Zucker, MI. (2000). Validity of a Set of Clinical Criteria to Rule Out Injury to the Cervical spine in Patients with Blunt Trauma. *The New England Journal of Medicine*, 343, 94-99.
3. Lee, H., Kim, H., Kim, D., Suk, K., Park, J., Kim, N. (2000). Reliability of Magnetic Resonance Imaging in Detecting Posterior Ligament Complex Injury in Thoracolumbar Spinal Fractures. *Spine*, 25(16), 2079-2084.
4. McIsaac, WJ., Goel, V., To, T., Low, DE. (2000). The Validity of a Sore Throat in Family Practice. *Canadian Medical Association Journal*, 163(7), 811-815.
5. Lindberg, E., Elmsary, A., Janson, C., Gilslason, T. (2000). Reported Snoring: Does Validity Differ By Age? *Journal of Sleep Research*, 9(2), 197-200.
6. Jones, EG., Kay, M. (1992). Instrumentation in Cross-Cultural Research. *Nursing Research*, 41(3), 186-188.
7. Lawrence, VA., DeMoore, C., Glenn, ME. (1999). Systematic Differences in Validity of Self-Reported Mammography Behavior: A Problem for Intergroup Comparisons. *Preventive Medicine*, 29, 577-580.

Assignments:

1. Select a published instrument that is a possible candidate for including in a research project of your own and write a full summary (1-2 pages) of the evidence for the instrument's reliability and validity. Note whether the instrument has been used and evaluated in populations similar to the one used in your research and note reliability or validity evidence that is missing and suggest ways to gather that information. **Due August 12th, 2009 in class.**

Session 6 (August 12th)

Focus:

In Session 6 we will focus on factor analysis to examine how underlying constructs influence the responses on measured variables. For example, are the measures of cardiac function influenced by a single underlying construct or by multiple constructs? There are two basic types of factor analysis; exploratory and confirmatory. Exploratory attempts to discover the nature of the constructs influencing a set of responses. Confirmatory tests whether a specific set of constructs is influencing responses in a predicted way.

Learning Objectives:

1. Describe the nature and purpose of factor analysis
2. Conduct a factor analysis using statistical software
3. Interpret and factor analytic statistical output and describe the implications of the output for a particular measure.

Activities:

1. Computer exercise on conducting factor analysis.

Readings:

1. Carmines and Zeller, Appendix
2. DeVellis, pp 102-137
3. Tabachnick, G., Fidell, LS. (1996). Using *Multivariate Statistics*, (3rd ed.). New York, NY: Harper Collins.

Session 7 (August 19th)

Focus:

The central goal of this session is to introduce key concepts in survey/questionnaire design and development. This session will cover multiple aspects of item and response option wording, item sequencing, and questionnaire administration. In addition, key generic measures of health status will be introduced and discussed. Finally, we will review and discuss methods for presenting measures as part of a grant proposal.

Learning Objectives:

1. Describe key issues in questionnaire item development and sequencing.
2. Review and critique published questionnaires.

3. Identify key generic measures of health status that would be appropriate for your research.
4. Write a measurement section of a grant proposal.

Readings:

1. Guyatt, GH. (2000). Making Sense of Quality-of-Life Data. *Medical Care*. Vol. 38, Num 9, Suppl. II, pp II-175-II-179
2. Krosnick, JA. (1999). Survey Research. *Annu. Rev. Psychol.* 50:537-67
3. Patrick, DL., Yen-Pin, C. (2000). Measurement of Health Outcomes in Treatment Effectiveness Evaluations: Conceptual and Methodological Challenges. *Medical Care*. Vol. 38, Num. 9, Suppl. II, pp II-14-II-25
4. Stewart, AL., Nápoles-Springer, A. (2000). Health Related Quality-of-Life Assessments in Diverse Population Groups in the United States. *Medical Care*. Vol. 38, Num. 9, Suppl. II, pp II-102-II-124.

Assignments:

Due next session:

1. Select three measurements and write short summaries about them that would be appropriate for a grant application. Several examples of measurement sections from successful NIH grant proposals are included in this section on courseweb. **Due August 26th, 2009 in class.**

Session 8 (August 26th)

Focus:

Final exam

Learning Objectives:

1. Final exam will be administered **in Parkvale 305 A/B on Wednesday, August 26th at 1pm.**