

2010 Workshops – Wednesday, October 27, 2010

Full Day Workshop 9 am to 4 pm

Analysis of Longitudinal Studies of HRQOL in SAS, SPSS and R

Main author: Diane Fairclough, DrPH Colorado School of Public Health, Denver, CO USA

This morning/afternoon workshop will provide a tutorial of the analysis of longitudinal studies with missing data. During the morning we will start with models that assume the dropout is ignorable given the analysis incorporates all available data. We consider two designs: where time is an ordered category (e.g. pre, during, post) or time is a continuous construct. For both, we will first discuss the assumptions underlying each model and then will go through the steps of fitting the model in SAS, SPSS and R. During the afternoon, we will continue to work with the above models and also consider a few analysis techniques that can be used as sensitivity analysis when there is concern about the non-ignorable assumptions. The workshop will include time for practice with datasets that will be provided. Participant are encouraged (but not required) to bring laptop computers with SAS, SPSS or R installed.

Level: Advanced

Half Day Workshops

Morning Workshops (#1-#8) 9 am to 12 pm

Workshop 1

An Introduction to Pediatric Patient Report Outcomes and their Inclusion as Endpoints in Pediatric Clinical Trials

Instructors: Rob Arbuckle, MSc Mapi Values Ltd, Bollington, Cheshire United Kingdom; Linda Abetz-Webb, MA Mapi Values Ltd, Bollington, Cheshire United Kingdom; Robyn Carson, MPH Forest Research Institute, Jersey City, NJ USA; Barbara Edelman Lewis, PhD Ironwood Pharmaceuticals, Cambridge, MA USA and Olivier Chassany, PhD Hopitaux de Paris, Paris, France

Patient reported outcomes (PROs) are often required as primary endpoints in pediatric clinical trials, or are of value as key secondary endpoints. However, pediatric PRO selection/development and pediatric clinical trial design are both complex and challenging. When, in relation to the adult program, should you plan for pediatric trials? Are all PRO development methods applicable to pediatrics PROs, including observer ratings? How do you design developmentally appropriate PROs when the age range spans 2-17 years? How do we best elicit content from children to develop pediatric PROs? At what age must we rely on observer ratings and from what age can children reliably complete a PRO themselves? What recall periods and response formats are appropriate? When should you engage with regulators and what standards do they expect? Is pooling of data across age ranges appropriate? We will explore the above issues with input from industry, regulatory and agency perspectives and with real-life examples from several disease areas including constipation/IBS, restless legs syndrome, and allergic rhinitis. Recordings of challenging pediatric interviews will be shared and creative methods for engaging with and eliciting content from children discussed. There will be the opportunity to code child interview transcripts.

Level: Basic

Workshop 2

Conducting Cognitive Interviews in Special Populations

Instructors: Carla Demuro, MS RTI Health Solutions, Research Triangle Park, NC USA; Dana DiBenedetti, PhD RTI Health Solutions, Research Triangle Park, NC USA; Sandy Lewis, BSRN RTI Health Solutions, Research Triangle Park, NC USA; and Sheri Fehnel, PhD RTI Health Solutions, Research Triangle Park, NC USA

The cognitive interview was initially developed in the early 1980's to identify sources of potential response error within surveys or questionnaires. With the release of the final FDA guidance 'Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims' (US Department of Health and Human Services, 2009), cognitive interviews have gained importance and relevance in both concept elicitation and cognitive debriefing. Standard approaches for the conduct of a cognitive interview such as think aloud, verbal probing, and concurrent and retrospective probing abound and all provide valuable insight. However, such standard approaches may lose utility in a sick or otherwise challenged population. For example, what approach is best for use in pediatric subjects to sustain attention? How does the interview environment impact subjects with dementia, schizophrenia or other cognitive impairment? Are there special considerations in interviewing patients with communication challenges (e.g., those with severe stuttering, visual impairment)? How do we ask questions about sensitive topics such as sexual functioning, risk behaviors or sexually transmitted infections without introducing bias? This workshop seeks to provide tips and techniques to guide the researcher in methods to foster a successful cognitive interview under suboptimal conditions that will meet the rigor described in the FDA guidance.

Level: Basic

Workshop 3

Interpreting Utility (Preference-Based) Measures of Health-Related Quality of Life

Instructors: David Feeny, PhD Kaiser Permanente Northwest, Portland, OR USA; Allyson Jones, PhD University of Alberta, Edmonton, AB Canada and Maria Santana, PhD University of Alberta Hospital, Edmonton, Alberta Canada

The Workshop will be at an advanced level, focusing on the interpretation of utility scores from direct and multi-attribute (indirect) approaches to preference-based measurement. The direct approaches will include the visual analogue scale (Feeling Thermometer), time tradeoff, and standard gamble. Major multi-attribute utility measures will include the EQ-5D, Health Utilities Index (HUI), and Short-Form 6D. The Workshop will include hands-on experience in the direct elicitation of preference scores for a health state, the completion of questionnaires from several instruments (EQ-5D, HUI2 and HUI3, SF-6D), and the analysis and interpretation of the results. The interpretation of scores will be considered in the context of comparing groups at a point in time as well as comparing within-person change over time. Applications will be drawn from diverse clinical and population health applications including osteoarthritis of the knee, multiple sclerosis, total hip arthroplasty, lung transplantation, acute lymphoblastic leukemia, hip fracture, data from population health surveys, and the use of preference-based measures in routine clinical practice settings. Evidence on clinically important differences will be discussed. Attendees should, at a minimum, have a basic knowledge of the conceptual foundations and practical approaches of the utility approach to assessing health-related quality of life

Level: Advanced

Workshop 4

Everything you wanted to know about modern psychometric methods but were too afraid to ask: Part 1

Instructors: Jeremy Hobart, MD PhD Peninsula Medical School, Plymouth, Devon UK and Stefan Cano, PhD Peninsula Medical School, Plymouth, Devon UK

Confused by classical test theory? Finding item response theory inaccessible? Wrestling with Rasch analysis? Struggling with the statistics? Muddled by the math? Are you like we were: wannabe psychometricians trying to get into modern psychometric methods, but who have attended other workshops and left feeling none-the-wiser? Then this is the modern psychometric methods workshop for you. In reality the basic principles are very simple. So, we have put together a straightforward, non-technical, non-mathematical approach to modern psychometric methods that will help you begin to access the fantastically powerful and illuminating information they have to offer. In this workshop we will describe: the role of quality of life and other types of health rating scales as outcome measures; the basic principles underlying traditional psychometric methods, their limitations and drivers for modern methods; the principles underpinning modern psychometric methods; the similarities and difference between traditional and modern psychometric methods; the similarities and important differences between Rasch analysis and Item Response Theory. This workshop will also provide a hands-on instruction in using the Rasch Measurement Model (RUMM2020) software package in order to analyze data for internal construct validity; reliability; category probability curves; differential item functioning; and scaling characteristics. The three-hour workshop will be divided into three equal sub-sessions: lecture including descriptions of background, principles and comparison of methods; demonstration of Rasch analysis in practice; hands-on class exercise including data analyses using the RUMM 2030 program, and question & answer session (attendees should bring their laptops to the session to fully benefit from the hands on data analysis exercise).

Workshop level: Basic

Workshop 5

Interpretation of PRO Data within Clinical Trials

Instructors: Josephine Norquist, MS Merck Sharp & Dohme, Corp., North Wales, PA USA and Kathleen Wyrwich, PhD United BioSource Corporation, Bethesda, MD USA

Learning Objectives: (1) Understand the FDA and EMA regulatory recommendations for interpreting PRO data to support product labeling and promotional claims (2) Describe current methodology to derive a responder definition (3) Understand use of cumulative response distribution (4) Pitfalls when applying within-person changes to between-group population differences and (5) Address the need for interpretation of between-group treatment differences. The final FDA Guidance on PRO measures no longer proposes the minimal important difference (MID) as an approach for the interpretation of PRO data in contrast to the EMA Health-Related Quality of Life Reflection Paper. Instead, an a priori responder definitions or, alternatively, cumulative distribution functions are recommended by the FDA. This methodology, however, does not assist in the identification of what should be considered a 'clinically meaningful' treatment difference. A challenge in this emerging area is that 'clinical meaningfulness' is an undefined subjective concept that may vary depending on the perspective of who is judging the clinical meaningfulness. In addition, linking treatment 'differences' on a PRO measure to treatment differences on a clinical parameter in clinical trials has methodological issues.

Level: Advanced

Workshop 6

Applications of Item Response Theory Modeling for Enhancing Health-related Quality of Life Measurement

Instructors: Bryce Reeve, PhD National Cancer Institute, Bethesda, MD USA and
Chih-Hung Chang, Ph.D. Northwestern University, Chicago, IL USA

There is a great need in health outcomes research to develop instruments or assessment tools that can accurately measure a person's health status with minimal response burden. Further, these measurement tools must be free of bias due to possible cultural/racial differences which may affect responses to questionnaires administered in multi-national settings. This need for psychometrically sound and clinically meaningful measures calls for better analytical tools beyond the methods available from traditional measurement theory. Applications of item response theory (IRT) modeling have increased considerably because of its utility for instrument development and evaluation, assessment of measurement equivalence, instrument linking, item banking, and computerized adaptive testing (CAT). IRT models the relationship, in probabilistic terms, between a person's response to a survey question and their standing on a health construct (e.g., fatigue or depression) being measured. This information allows instrument developers to create reliable and efficient quality of life measures tailored for an individual or group or specific application.

Level: Basic

Workshop 7

Advanced Psychometric Methods, Part 1: Use of Exploratory and Confirmatory Factor Analyses in PRO Instrument Development and Evaluation

Instructors: Donald Stull, PhD United BioSource Corporation, Bethesda, MD USA; Margaret Vernon, PhD United BioSource Corporation, Bethesda, MD US; Elizabeth Merikle, PhD United BioSource Corporation, Bethesda, MD USA

The development and psychometric evaluation of PRO instruments requires the application of a number of different techniques, including exploratory and confirmatory factor analysis (FA), and structural equation modeling (SEM). We will provide a brief overview of psychometric analyses and will then focus on the application of (1) exploratory and confirmatory factor analysis for understanding of new measures and (2) use of SEM for testing construct validity. Exploratory and confirmatory FA can be used to examine the relationships among items with a PRO measure or among different domains or multiple PRO measures. These techniques are useful for understanding the internal structure of PRO instruments and for understanding construct validity. This half-day workshop will describe the main methods of FA and illustrate these methods with examples from the instrument development literature. SEM is a powerful analytic technique that combines FA and path analysis in a simultaneous, confirmatory approach. Using SEM, the researcher can specify and evaluate hypothesized relationships between observed and latent (unobserved) constructs as well as relationships among the latent variables. SEM can also estimate the reliability and validity of measurement models while explicitly modeling measurement error. A researcher specifies a measurement model and a structural model which specifies relationships among the latent variables to examine construct and criterion-related validity. If the observed covariances are consistent with the model-implied covariances, the researcher has evidence supporting the construct validity of the PRO measure. This workshop will demonstrate the main methods, testing assumptions and criteria, and provide examples to illustrate the methods of SEM.

Level: Advanced

Workshop 8

How many subjects do I need for my study? Sample size determination for studies with QoL outcomes

Instructor: Stephen Walters, PhD University of Sheffield, Sheffield, United Kingdom

Quality of Life measures are now frequently used in clinical trials and health services research, as primary endpoints. Thus investigators are now wanting to plan studies using QoL measures, which includes questions on sample size. Sample size calculations are now mandatory for many research protocols and are required to justify the size of studies in papers before they will be accepted by journals. Sample size is critically dependent on the purpose of the study, the QoL outcome measure and how it is summarised, the proposed effect size and the method of calculating the test statistic. Whatever type of study design is used the problem of sample size must be faced. This workshop will describe how sample sizes may be estimated for a variety of different study designs using QoL outcomes. The study designs will include: two group clinical trials, cross-over trials, cross-sectional surveys and reliability studies. The consequences of comparing more than two groups or investigating several QoL outcomes simultaneously are discussed. This course will provide practical guidance on the estimation of sample size for a variety of study designs. It will feature many real and original examples, taken from the instructor's experience of analysing PROs. Content: Introduction to significance tests, P-values and power; Samples sizes for comparison of two independent groups; Sample sizes for paired designs; Sample sizes for equivalence/non-inferiority studies; Unequal groups; Multiple comparisons; Sample sizes for surveys; Sample sizes for reliability and method comparison studies; Use of computer simulation. It is expected that course participants will be familiar with statistical concepts such as: hypothesis testing; confidence intervals; simple statistical tests (e.g. t-test).

Level: Basic

Afternoon Workshops (9-16)

1 pm to 4 pm

Workshop 9

Everything you wanted to know about modern psychometric methods but were too afraid to ask: Part 2

Instructors: Jeremy Hobart, MD PhD Peninsula Medical School, Plymouth, Devon UK and Stefan Cano, PhD Peninsula Medical School, Plymouth, Devon UK

Part 2 builds on Part 1 and is a direct response to last year's feedback for more time in general, more detail on the similarities and differences between Rasch measurement and IRT, more examples of practical applications and their interpretation, and the call for whole day devoted to this critically important topic. This allows the workshop to develop further the ideas behind modern psychometric methods a little deeper but still in a non-technical and basic manner. We will provide multiple examples of the applications of Rasch measurement to scale evaluation, scale modification, new scale development, head to head scale comparisons, item banking, and the all important issue of determining clinically significant change.

Workshop level: Basic

Workshop 10

The use of R and WinBUGS in fitting Item Response Theory models

Instructor: Yuelin Li, PhD. Memorial Sloan-Kettering Cancer Center, New York, NY USA

This workshop is developed for behavioral scientists who want to learn how to fit IRT models with open-source statistical computer languages R and WinBUGS. This workshop is a How-To guide to IRT basic theories and a step-by-step tutorial on how to use R and WinBUGS to analyze IRT data for measurement development. R and WinBUGS are well suited for learning IRT because (among other advantages) their syntax closely resembles the IRT theories. They promote a clear understanding of the fundamental theories. Additionally, support can be sought from active online user forums. We will capitalize on these advantages to learn how to make R and WinBUGS work for your research and teaching. Participants are encouraged to bring a laptop with R and WinBUGS already installed. Participants are encouraged to go to my website (<http://idecide.mskcc.org>), download the sample R and WinBUGS codes, and try them before the workshop. Additional software requirements are also listed on the website. The workshop is organized as follows. We will first cover a few basic techniques in using R for general statistical analysis and data visualization. For IRT, we will use real examples from my own research to cover the basic concepts as well as the more advanced topics. The basic skills include how to select/modify items and response categories by visually inspecting the Item Characteristic Curves and the Item Information Curves. The more advanced topics will include Generalized Partial Credit Model by WinBUGS, Differential Item Function (DIF) and Computerized Adaptive Testing (CAT). CAT addresses the interactive administration of an instrument, e.g., over the web.

Level: Advanced

Workshop 11

Methods for Estimating Longevity and HR-QOL Using Data from Medical Records

Instructor: Michael Molla, Ph.D. National Center for Health Statistics / CDC, Hyattsville, Maryland USA

This workshop is suitable for those who are not familiar with methods of estimating longevity and HR-QOL using the life table and the illness-death models. The Workshop will have three parts. Part one, will introduce a simple single decrement life table model. Part II, will present the more advanced -illness-death- model. Part III will cover the integration of fatal and non-fatal health outcomes using each of the two models. Participants will be equipped with techniques that will enable them to estimate and evaluate the expected longevity and HR-QOL of a small or large group of patients based on data from sources such as medical records. Participants will be provided with hard copies of PowerPoint slides, copies of two NCHS publications that illustrate and explain the two models in detail and internet addresses where they can download the two programs free of charge. In the planned workshop, we will show how both longevity and HR-QOL can be estimated using data from sources such as medical records. We will discuss two models, two types of data and two PC-based computer programs. Longevity and HR-QOL will be estimated using data with single record per patient and data with multiple records per patient. The application will be illustrated using two PC-based computer programs, the first in Microsoft Office Excel and the second in PC-SAS.

Level: Basic

Workshop 12

Documenting PRO Evidence for EMA and FDA Submissions to Support Promotional and Labeling Claims

Instructors: Josephine Norquist, MS Merck Sharp & Dohme, Corp., North Wales, PA USA; Annabel Nixon, PhD Oxford Outcomes, Oxford, Oxon UK.; and Kati Copley-Merriman, M.S., M.B.A RTI Health Solutions, Research Triangle Park, NC USA.

This workshop will provide a detailed overview of the required evidence related to regulatory submissions in order to achieve a patient-reported outcome (PRO) label claim, and insights into procedures for documenting this evidence in the format required by regulatory authorities. The discussions will be framed by recent regulatory guidances (FDA 2009, EMA 2005) and will have a practical focus in order to guide what is, and what is not, required in order to successfully achieve a desired claim. Differences between the two regulatory guidances will be highlighted during the workshop. The first part of the workshop will examine the required evidence for the PRO instrument itself including time focusing on the key issues of content validity, psychometric evidence for validity and reliability, and interpretation of scores. The second part of the workshop will move on to look at strategies for documenting PRO evidence in the wider context of the clinical trial evidence including endpoint models and PRO components of the statistical analysis plan. The workshop will finish by examining how to address PRO instrument modification both in terms of linguistic validation and ePRO adaptation. Because a standard and well-organized PRO Evidence Dossier to support labeling and promotional claims for FDA and EMA regulated medical products is critical, increases efficiency of providing relevant PRO information to the regulatory agencies and increases efficiency of the regulatory review process of the PRO information, participants will be introduced to the 'When', 'What', and 'How' of a PRO Evidence Dossier. The workshop will finish with discussions on conducting a gap analysis and on when to engage regulatory agencies to discuss the PRO endpoint(s). Throughout,

strategies for maximizing opportunities for achieving a PRO label claim will be highlighted and discussed within the group. The workshop will include an example claim and corresponding PRO instrument to work through as a class.

Level: Basic

Workshop 13

Advanced Psychometric Methods, Part 2: Executing and Interpreting Exploratory and Confirmatory

Instructors: Donald Stull, PhD United BioSource Corporation, Bethesda, MD USA; Margaret Vernon, PhD United BioSource Corporation, Bethesda, MD US; Elizabeth Merikle, PhD United BioSource Corporation, Bethesda, MD USA

This half-day workshop will build on Advanced Psychometric Methods, Part 1, by presenting results from examples of exploratory and confirmatory factor analyses; executing live, interactive analyses; and interpreting results of output, particularly for confirmatory factor analyses and structural equation models (including latent growth mixture models). We will work through examples of analyses by presenting hypothesized models, discussing key analytic criteria (e.g., sample size, factor loading size, extraction, rotation, key parameter estimates, cross-loadings and correlated errors, model specification and identification, fit indices, indications of model misfit), and how to interpret output. Annotated examples will be presented from output from selected software (e.g., SAS, Stata, EQS, and Mplus), but the issues are relevant regardless of the users software.

Level: Advanced

Workshop 14

Systematic Reviews of Measurement Instruments

Instructors: Caroline Terwee, PhD VU University Medical Center, Amsterdam, the Netherlands; Lidwine Mokkink, PhD VU University Medical Center, Amsterdam, the Netherlands,; Henrica de Vet, PhD VU University Medical Center, Amsterdam, the Netherlands and Donald Patrick, PhD University of Washington, Seattle, Washington US

Systematic reviews of measurement instruments are reviews in which the measurement properties of all available instruments that measure a specific construct are systematically appraised and compared according to quality standards. Considering the large number of instruments available in many fields, such reviews are becoming increasingly important for choosing instruments for different applications. They can also aid to minimize the number of instruments used in a field by recommending a selection of best or the most promising instruments. Finally, systematic reviews are important for detecting gaps in knowledge on measurement properties and prioritizing new validation studies. The number of systematic reviews of measurement properties is increasing rapidly. However, the methodology and quality of such reviews varies widely, with regard to the search strategy, quality assessment of the included studies, and pooling and interpretation of the results. In this workshop we provide guidelines and practical tools for performing high quality systematic reviews of measurement instruments. Our methodology closely resembles the methodology of the Cochrane Collaboration for systematic reviews of randomized trials.

Level: Basic

Workshop 15

Innovative Methods for Electronically Capturing Patient-Reported Outcomes and Preference Data in Non-Interventional Studies

Instructors: Ingela Wiklund, PhD United BioSource Corporation, London, United Kingdom; Karin Coyne, PhD United BioSource Corporation, Bethesda, MD USA; and Shae Wilkins, Arrowhead Electronic Healthcare, Austin, TX USA

The increasing use of ePROs in a range of clinical and research settings has created opportunities to measure patient experience in innovative ways. This workshop focuses on best practices for leveraging electronic data collection methods to improve data quality and streamline capture of data in non-interventional studies. This workshop will draw from a range of examples from the literature, the field, and experience with the EXACT-PRO Initiative to inform a discussion of key issues associated with ePRO development and implementation. Specific topics to be addressed include: an overview of technology currently in use, new instrument development using electronic platforms innovatively and considerations for migrating existing validated instruments from paper to electronic platforms. The workshop will also address the use of technology to improve the administration of preference assessment methods including utilities, time trade-off and conjoint analysis. Other important topics include compliance enhancement strategies, and use of electronic methods with special populations such as older adults and children.

Level: Basic

Workshop 16

METHODS FOR DEVELOPING PREFERENCE-BASED MEASURES OF HEALTH FROM EXISTING MEASURES

Instructors: Tracey Young, PhD University of Sheffield, Sheffield, South Yorkshire United Kingdom; John Brazier, PhD University of Sheffield, Sheffield, South Yorkshire United Kingdom; and Donna Rowen, PhD University of Sheffield, Sheffield, South Yorkshire United Kingdom

Recent years has seen increasing reliance on a few generic preference-based measures of health (e.g. EQ-5D, HUI3, QWB or SF-6D) for calculating Quality Adjusted Life Years (QALYs) for economic evaluation. However, generic measures may not be used in key clinical studies. This may be due to a desire to reduce patient burden or a view that generic instruments are not valid for the condition or responsive to the effects of treatment. For these reasons there is interest in developing new preference-based measures of health. This workshop focuses on the estimation of the preference-based indices from existing generic and condition specific measures of health related quality of life. The focus is on the derivation of health states from existing measures rather than the other stages of valuing a sample of states and modeling the health state values. The workshop offers a practical introduction to the use of psychometric methods in the development and refinement of health state classifications. It will also examine the policy implications of using different descriptive systems to derive preference-based measures. It assumes a basic knowledge of HRQoL measurement and QALYs.

Level: Advanced