PO4  TRANSLATABILITY ASSESSMENT OF THE YOUTH QUALITY-OF-LIFE INSTRUMENT–WEIGHT MODULE (YQOL-W)

Conway K1, Patent M1, Conti A2, Acquadro C3, Fuller D4

1MAPI Research Trust, Lyon, France, 2University of Washington, Seattle, WA, USA, 3MAPI Institute, Philadelphia, PA, USA

OBJECTIVES: 1) To assess if a translatability assessment (TA) conducted prior to a psychometric evaluation in PRO instrument development could predict items subsequently eliminated, and 2) to provide evidence to the extent to which a TA adds value to the translation and cultural adaptation of PRO instruments.

The Youth Quality-of-Life Instrument–Weight module (YQOL-W) was chosen as a candidate questionnaire for this exercise. METHODS: A team of two linguists, blinded to the results of psychometric analyses and decisions on item reduction, conducted a TA on the 32-item pre-final version of the YQOL-W. Results were categorized into several types of issues. Items for possible deletion were identified and compared to the results of the item-reduction phase, using both qualitative and psychometric methods performed by the developer. RESULTS: During the item-reduction phase, 11 items were dropped, seven for qualitative reasons (items 10, 30, 31, 27-32) and four for weak factor loading (items 14, 16, 18, 19). Out of the 11 items dropped, TA identified nine problematic items (82%) and, among them, advised dropping five items (45.4%), either for redundancy with others (items 14, 19, 29, and 32) or inconsistency of construct with respect to concept (item 27). For item 28, TA recommended changing the original for semantic reasons. For items 10, 30, and 31, TA recommended using alternative wording for translation purposes but did not suggest changing the original. Finally, for items 16 and 18, TA recommended no changes.

CONCLUSIONS: Our study showed that a translatability assessment can anticipate results of the item-reduction phase. Although more empirical studies are needed, we have demonstrated that the involvement of linguists in the early development phase could be a plus in detecting irrelevant or inappropriate items, and that a translatability assessment can be a useful step in PRO instrument development.

POD IUM SESSION II: COMPARATIVE EFFECTIVENESS RESEARCH STUDIES

CE1  DOES COMPARATIVE EFFECTIVENESS RESEARCH INCREASE ECONOMIC EFFICIENCY?

Saxton CM1,2, Weinberger SE3,4,5

1Children’s Hospital Boston, Boston, MA, USA

OBJECTIVES: We explore whether comparative effectiveness research (CER) can in principle increase economic efficiency. METHODS: The definition and goals of CER are reviewed. The principle of CER is compared, when the two diverge are discussed. RESULTS: CER is frequently defined to include measurement of clinical benefits and harms. Economic efficiency on the other hand depends on individuals’ preferences for the health technology under consideration relative to other health or non-health commodities (e.g., following the Grossman’s utility model). The distinction is important because choosing the most clinically effective interventions as the end goal of CER may increase but fail to maximize economic efficiency, or in other words, may result in greater marginal impact on patient or family utility. We present examples which may help explain why preferred care is not reimbursable and prescribed treatment may not adhere to economic theory. The situation is analogous to the one we would have adopted for ALLHAT, including adaptive randomization with the potential for early stopping. CONCLUSIONS: One design was executed using the ALLHAT data, preserving accrual order and re-sampling patients when adaptive randomization called for additional patients in a given arm. RESULTS: ALLHAT enrolled 42,418 patients in four years, following patients for an additional four years and showed that none of the comparator treatments differed significantly from diuretics. The recommendation for all arms was calcium channel blockers, because it initially outperformed other arms during the accrual stage of the trial. CONCLUSIONS: CER trials are important to clinical research. However, conventional methods - even with early stopping rules as in ALLHAT – can be so time-consuming, that the clinical relevance of the study may have diminished by the time results are published. CER or adaptable methods, tailored to CER, can be more efficient, and produce more timely evidence.

CE2  RECONCILING VARIATIONS IN PROPENSITY SCORE MATCHING (PSM) ALGORITHMS IN OBSERVATIONAL COMPARATIVE EFFECTIVENESS RESEARCH (OCR)

Baker BJ1, Heien HC2

1Mayo Clinic, Rochester, MN, USA

OBJECTIVES: The OCR literature that evaluates comparative effectiveness of alternative medical interventions using observational data and matching, and other non-randomization methods has seen explosive growth in the use of PSM in recent years. However, different PSM algorithms (e.g., one-to-one, one-to-many, radius matching etc.) yield different estimates of the treatment effect. METHODS: Data from a physiological example come from health care claims and clinical information from electronic health records in a large U.S. hospital on patients that underwent hysterectomy with robotic (N=315) and abdominal (N=265) approaches. The ATE of interest is the predicted difference in 6-week all-cause costs starting from index surgery date. Baseline characteristics including age, race, insurance type, BMI, comorbidities and other clinical characteristics adjusted in each of the 8 PSM models. 95% confidence interval (CI) for the range of ATE is obtained through 500 bootstrap replications. Internalized model as the basis for testing adaptation methods, and alternative methods and predicted costs. RESULTS: Following eight PSM algorithms were implemented: one-to-one, one-to-many (with replacement), 3 radius matching with caliper=0.01, 0.001 and 0.0001, and linear regression. The usual radius matching approach was found costlier with its costs varying between $16,713 and $17,522, while abdominal costs varied between $13,326 and $14,615. The minimum predicted cost difference was $2,812 (95% CI: $1,186, $4,437) while the maximum difference was $3,854 (95% CI: $2,183, $5,525). The paired bootstrap sample size was highest in radius matching with caliper=0.001. CONCLUSIONS: The range for ATEs instead of a single ATE from a specific PSM algorithm may help explain why preferred care is not reimbursable and prescribed treatment not adhered to. In addition, economists define costs as the forgone greater marginal impact on patient or family utility. We present examples which may help explain why preferred care is not reimbursable and prescribed treatment may not adhere to economic theory. The situation is analogous to the one we would have adopted for ALLHAT, including adaptive randomization with the potential for early stopping. One design was executed using the ALLHAT data, preserving accrual order and re-sampling patients when adaptive randomization called for additional patients in a given arm. RESULTS: ALLHAT enrolled 42,418 patients in four years, following patients for an additional four years and showed that none of the comparator treatments differed significantly from diuretics. The recommendation for all arms was calcium channel blockers, because it initially outperformed other arms during the accrual stage of the trial. CONCLUSIONS: CER trials are important to clinical research. However, conventional methods - even with early stopping rules as in ALLHAT – can be so time-consuming, that the clinical relevance of the study may have diminished by the time results are published. CER or adaptable methods, tailored to CER, can be more efficient, and produce more timely evidence.

CE3  ADVANTAGES OF BAYESIAN ADAPTIVE TRIALS FOR COMPARATIVE EFFECTIVENESS RESEARCH (CER): ‘RE-ADAPT’ ALLHAT

Woolf P1,2,3, Isahk J1, Mullins CD1,2,3, Jocson J1,2, Luce BR1,2, Davis B1

1Berry Consultants, Austin, TX, USA, 2United BioSource Corporation, Bethesda, MD, USA, 3University of Maryland School of Pharmacy, Baltimore, MD, USA, 4The University of Texas School of Public Health, Houston, TX, USA

OBJECTIVES: Traditional randomized trials are the gold-standard for clinical research but may be long, expensive and, especially for CER, inefficient in providing evidence to prepare next-generation treatments. METHODS: Pragmatic Trials (RE-ADAPT) investigators sought to test efficiency gains in CER trials by redesigning, then re-executing ALLHAT using adaptive trial methodologies. RESULTS: ALLHAT was a traditionally designed, large comparative trial looking at calcium channel blockers compared to diuretics across classes of anti-hypertensive drugs compared with diuretics. We prospectively developed seven Bayesian adaptive designs incorporating combinations of early stopping, adaptive randomization, and early stopping. Only information available when ALLHAT was designed, we examined operating characteristics of the designs to evaluate clinical and economic efficiencies (e.g. shorter trials, greater allocation to superior treatments). One design was chosen as the common conclusion as the original ALLHAT trial but would have stopped nearly one year earlier. Also, the simulated trial randomized 28% more patients to calcium channel blockers, because it initially outperformed other arms during the accrual stage of the trial. CONCLUSIONS: CER trials are important to clinical research. However, conventional methods – even with early stopping rules as in ALLHAT – can be so time-consuming, that the clinical relevance of the study may have diminished by the time results are published. CER or adaptable methods, tailored to CER, can be more efficient, and produce more timely evidence.