0.40, and 0.72 (0.50-0.90), respectively, but were stable over time (re-test). For data available in 2009, incidence of PML (cases per 100,000 patients) over two years was estimated at 0.22 (95% CI 0.0078-3.13) for patients initiated on efalizumab and 1.21 (0.0052-14.81) for patients remaining on treatment. ORs versus placebo were 1.7 (1.4–2.2) for AAs and 1.5 (0.6–5.1) for SF-36. Combination of normalized weights and adjust treatment set size with the result of efalizumab treatment was 0.60 (stdev 0.28) on a scale of -1 to +1. The same approach was applied to compare efalizumab with three active comparators. CONCLUSIONS: The method supports comparative safety assessment and captures uncertainty in judgments and in data, allowing the inclusion of rare, disparate and newly emergent safety data in quantitive assessments.

PRM17 LONGITUDINAL AND CROSS SECTIONAL ASSESSMENTS OF HEALTH UTILITY IN ADULTS WITH HIV/AIDS: A SYSTEMATIC REVIEW AND META-ANALYSIS

Yucel A., Ferries E.A., Sharma M., Johnson M., Aparasu R.

Hanoi Medical University, Hanoi, Vietnam

OBJECTIVES: Utility estimates are important health outcomes for economic evaluation of care and treatment interventions for patients with HIV/AIDS. We conducted a systematic review and meta-analysis to examine the performance of preference-based instruments, estimate health utility of patients with HIV/AIDS by disease stages, and investigate changes in their health utility over the course of antiretroviral treatment. METHODS: We searched PubMed/MedLine, Cochrane Database of Systematic Review, NHS Economic Evaluation Database and Web of Science for English-language peer-reviewed papers published during 2000-2013. We selected 49 studies that used 3 direct and 6 indirect preference based instruments to make a total of 218 utility measurements. Random effect models with robust estimation of standard errors and multivariate fractional polynomial regression were used to obtain the pooled estimates of utility and model their trends. RESULTS: Utility depending on direct-preference measures tended to be lower than other types of measures. Utility elicited by two of the indirect preference measures - SF-6D (0.171) and EQ-5D (0.114), and that of Time-Trade off (TTO) (0.151) was significantly different than utility elicited by the Standard Gamble (SG). Comparison of the pooled health utilities of asymptomatic HIV patients, symptomatic and AIDS patients reported a decrease of 0.025 (p=0.40) and 0.176 (p=0.001) in utility scores, adjusting for method of assessment. In longitudinal studies, the pooled health utility of HIV/AIDS patients significantly decreased in the first 3 months and then increased afterwards. Magnitude of change varied depending on the method of assessment and length of antiretroviral treatment. CONCLUSIONS: The study provides an accumulation of evidence on measurement properties of health utility estimates that can help inform the selection of instruments for future studies. The pooled estimates of health utilities and their trends are useful in economic evaluation and policy modeling of HIV/AIDS treatment strategies.

PRM18 IMPROVING PARAMETER ESTIMATION FOR A DECISION-ANALYTIC MARKOV MODEL TO EVALUATE THE USE OF NOVEL BIOMARKER LED STRATEGIES FOR PREVENTION OF CARDIOVASCULAR DISEASE

Lamocki P1, Cairns K2, Keel F1, Conrads-Frank A1, Schnell-Inderst P1, Salomaa V1, Blankenberg P1, Zeller T1, Saarela O1, Kiviluoto K1, Kihre F1, Siibert U1

1Queen’s University, Kingston, ON, Canada, 2Cerner Corporation, Culver City, CA, USA

OBJECTIVES: This study sought to compare and improve the methods for estimating input parameters to populate a decision-analytic Markov model for use in the health economic evaluation of novel biomarkers for primary prevention of Cardiovascular Disease. METHODS: An initial five state Markov model is built using the FINRISK97 cohort and follow-up for coronary and stroke events. Two different approaches to estimating the transition rates between health states are used. The first approach involves a separate analysis for each of the transitions between different health states each using a separate Cox Proportional Hazards model (censoring any movement into a separate analysis for each of the transitions between different health states). In the second approach, all the transitions using a separate Cox Proportional Hazards model (censoring any movement into a separate analysis for each of the transitions between different health states), and investigate changes in their health utility over the course of antiretroviral treatment. T wo different approaches to estimating the transition rates varied depending on the method of assessment. INTERNATIONAL: Magnitude of change varied depending on the method of assessment and length of antiretroviral treatment. CONCLUSIONS: The study provides an accumulation of evidence on measurement properties of health utility estimates that can help inform the selection of instruments for future studies. The pooled estimates of health utilities and their trends are useful in economic evaluation and policy modeling of HIV/AIDS treatment strategies.

PRM20 COMPARING THE PREDICTIVE PERFORMANCE OF TWO VARIANTS OF THE ELIXHAUSER COMORBIDITY MEASURES FOR ALL-CAUSE IN-HOSPITAL MORTALITY IN A LARGE MULTI-PAYER ADMINISTRATIVE DATABASE

Perez J.R., Hao Y., Emons M.P., Mattison D.R., Krewski D1

1University of Ottawa, Ottawa, ON, Canada, 2Cerner Corporation, Culver City, CA, USA, 3Risk Sciences International, Ottawa, ON, Canada

OBJECTIVES: Tools used to address confounding, such as measures of patient comorbidity, have been important tools in patient outcomes research. The goal of this study was to compare the two most common comorbidity measures in administrative datasets: the Quan (Quan et al. 2005) and the AHRQ (AHRQ 2015) variants. METHODS: Of 8,128,713 patients with a mean age of 51.4 years and a male composition of 61% 41.6%, 76,215 (9.9%) were deceased at 1-year. Age, sex and all comorbidities were significantly correlated with mortality in the two Elixhauser variants irrespec-tively of lookback period. Excellent predictive performance (c-statistic) was found for the Quan (c.stat=0.874, 95%CI=0.867-0.873) and AHRQ (c-stat=0.876, 95%CI=0.870-0.879) variants and both exceeded the performance of LCA (c-stat=0.70-0.94). CONCLUSIONS: The two variants produced similar rankings for top performers, the MLM approach was more conservative in its c-stat ratio estimates which were shrunken towards the overall mean. The MLM method is recommended to better statistically adjust and avoid false positive identification of outliers.

PRM19 DIFFERENT MODELING TECHNIQUES CHANGE RANKINGS OF HOSPITAL PERFORMANCE? : MULTILEVEL MODELING VS. STANDARD LOGISTIC REGRESSION

Yacali A., Ferries EA, Sharma M, Johnson M, Aparasu R.

University of Houston, Houston, Texas, USA

OBJECTIVES: There are conflicting results about whether using multilevel modeling (MLM) produces different rankings in hospital performance compared to using traditional statistical techniques. We would like to compare hospital rankings obtained from standard logistic regression (LR) relative to MLM modeling. METHODS: The 2012 U.S. Nationwide Inpatient Sample (NIS) was used to identify patients with a primary diagnosis for ischemic stroke, using ICD-9 diagnosis criteria. Stepwise backward selection technique with logistic regres-sion was employed to examine the variation among hospitals in the south by adjusting for patient-level and hospital-level risk factors. The analysis included hospitals with more than 30 stroke cases in the southern region of the U.S. The hospital mortality rates for stroke among hospitals in the south. INTERNATIONAL: The hospital IDs have been masked in order to comply with the NIS data user agreement. RESULTS: There were 19,071 stroke hospitalizations in 320 hospitals. The same rankings were observed for top performers in hospitals’ observed/expected ratios (O/E) by LR and MLM. However, high O/E outlier status differed between the two statistical methods, yielding different rankings for worst performing hospital. Hospital K, L were ranked as the worst 3 performers in LR whereas K, L were ranked the same in MMO/E of Hospital K were 4.60 (CI=1.92–8.94) and 0.62–8.64) and 3.27 (CI=0.70–9.54) in LR and respectively. O/E of Hospital L were 2.44 (CI=1.90–8.13) and 3.33 (CI=0.70–9.54) in LR and respectively. LR identified Hospital M’s O/E ratio as 3.5 (1.477–2.6). MLM identified Hospital N’s O/E as 3.33 (0.70–9.54). CONCLUSIONS: Although both logistic regression and multilevel modeling produced similar rankings for top performers, the MLM approach was more conservative in its O/E ratio estimates which were shrunken towards the overall mean. The MLM method is recommended to better statistically adjust and avoid false positive identification of outliers.

PRM21 STATISTICAL IDENTIFICATION OF PATIENT SELECTION BIAS IN RETROSPECTIVE CHART REVIEWS

Signorovitch J1, Reichmann WM2, Li N1, Liu Z1, Hao Y1, Perez J2

1Analytiq Corp., Inc., Boston, MA, USA, 2Novartis Pharmaceuticals Corporation, East Hanover, NJ, USA

OBJECTIVES: Retrospective chart reviews are often used to generate real-world evidence of patient outcomes. Physicians participating in these studies may be asked to randomly select patient charts for inclusion among all charts meeting eligibility criteria. Random selection is important for obtaining representative samples and unbiased estimates of patient outcomes. This study examined a statistical method for detecting non-random selection of patient charts with an application to real data. METHODS: Example data were drawn from a recent retrospective oncology chart review study that estimated overall survival (OS) and progression-free survival (PFS) following the start of treatment. In the study, participating physicians were asked to provide a random sample of their eligible patients. Late class analysis (LCA) was used to test whether the distribution of FFS and OS was consistent with random selection for all charts [a one class model] versus a mixture in which some physician selected random charts, and all others selected convenient charts (models with > 1 class). The best fitting model was identified using information criteria. Physician characteristics, FFS and OS were compared across latent classes (LCA). RESULTS: In overall 40% of charts, media distribution for FFS and OS were significantly lower than published benchmarks drawn from randomized trials. LCA identified evidence for two classes of physicians. One class was consistent with biased selection for recently seen patients who were still alive at the time of chart review completion; the other was consistent with unbiased selection. CONCLUSIONS: In this example using data from a retrospective oncology chart review, LCA was helpful in identifying physicians who may have selected convenient but non-random patient charts. LCA warrants