A30

PRM96

OPTIMIZING THRESHOLDS FOR A CLINICAL RECOGNITION ALGORITHM

Hays WD1, Debes R2, Sutantia B3, Rosen J4
1Cerner Research, North Kansas City, MO, USA, 2Cerner Corporation, North Kansas City, MO, USA, 3Cerner Research, North Kansas City, MO, USA, 4Cerner Research and the University of California, San Diego, Culver City, CA, USA

OBJECTIVES: Cerner has developed the St John’s Sepsis Alert, an evidence-based real-time algorithm that alerts clinicians to the presence of the Systemic Immune Response Syndrome or sepsis. Using simulation, we estimated the Alert’s performance and determined the optimal cut-offs for ~17 included parameters (e.g., blood pressure). METHODS: We estimated the operating characteristics of the alert by applying its logic to 3 years of real-world data on adults from Cerner Health Facts, a time-stamped database extracted from electronic medical records. We evaluated the base case and performed bootstrap iteration uncertainty analysis. Each run used a different set of thresholds, each drawn randomly from the range of reasonable values using a Latin hypercube sampling design under the assumption of an independently distributed joint beta distribution. Each run provided a point on the Receiver Operating Characteristic curve. We constructed an extended dominance curve from the resulting point cloud and determined the optimal values as those associated with that curve. RESULTS: Data from ~69,000 hospitalizations with a 5% incidence of sepsis were available.Using baseline values for the Alert, we estimated a Sensitivity of 56%, Specificity of 90%, Positive Predictive Value of 22%, and Negative Predictive Value of 98%. The uncertainty analysis found that 10 sets of cut-offs dominated all the others. The C-statistics for these ranged from 67% to 75%, Sensitivity from 39% to 73%, Specificity from 96% to 76%, PPV 33% to 13%, NPV 95% to 98%, and the posterior probability of sepsis increased from 6% to 7-fold. The most accurate thresholds were not necessarily optimal for implementation; certain sets with only slightly lower C-statistics greatly decreased false positive rates. CONCLUSIONS: Simulation can usefully inform the clinical implementation of new algorithms. Using a range of acceptable thresholds can identify statistical optimality and characterize trade-offs deviating from it.

PRM97

EVALUATING THE COST-EFFECTIVENESS OF MULTICOMPONENT REHABILITATION GUIDELINES

Mewes JC, Steuten LMG, IJzerman MJ, van Harten WH

University of Twente, Enschede, The Netherlands

OBJECTIVES: The Dutch guideline for cancer rehabilitation recommends patients to engage in multicomponent interventions, i.e., several single interventions combined in intervention programs to form a health care strategy. In order to evaluate the cost-effectiveness of these multicomponent interventions, we performed a health economic evaluation of this guideline, data on the cost-effectiveness of these multicomponent interventions is required. However, to date, the interventions (cost-)effects were almost exclusively assessed for the single interventions rather than for the multicomponent intervention, which challenges the health economic analysis of the multicomponent interventions. The objective of this study was to identify or develop a method that allows to deduct the cost-effectiveness of multicomponent interventions from published data of the single interventions. METHODS: We searched the literature for articles offering a method or ideas for the development of a method for assessing the cost-effectiveness of multicomponent interventions. We selected a few, extrapolated cost-effectiveness data of the single interventions. RESULTS: Cost-effectiveness gap analysis was identified in the literature as being a suitable method, with further refinement. We suggest filling first calculate the costs of all interventions. Given the effectiveness of one intervention it is then possible to estimate how much additional effectiveness a second (or any subsequent) intervention would have to provide so that the multicomponent intervention remains cost-effective, given a range of ceiling ratios. Recommendations for methods for estimating the additional effect of subsequent interventions were deducted from the literature identified. CONCLUSIONS: We suggest estimating the cost-effectiveness of the combined interventions as recommended in clinical guidelines by performing a refined cost-effectiveness gap analysis method.

PRM98

THE USE AND IMPACT OF VALUE OF INFORMATION ANALYSIS IN DECISION-MAKING

Chuang LH1, Treur M2, Heeg B3, Van Hout B4, Filippon A5
1Pharmaceutische Universiteit, Rotterdam, The Netherlands, 2Pharmaceutische Universiteit, York, UK, 3Bayor Pharma AG, Berlin, Germany, 4University of Amsterdam, Netherlands, 5University of Amsterdam, Amsterdam, The Netherlands

OBJECTIVES: Along with uncertainty around reimbursement decisions, one should determine the worth of additional research to reduce the probability of making wrong decisions. Value of information analysis (VOI) provides an explicit framework to inform future research. The objectives of this research were to assess published evidence of VOI and to evaluate its impact on decision making and research agendas. METHODS: A literature review was conducted in MEDLINE and EMBASE to collect studies with VOI applications published until November 2012. Data extracted included study indication, year, country, sponsorship, type of VOI analysis, research impact and quality of the study. HTA guidelines of developed countries were checked and unclassified VOI studies were compared and discussed the relevance of VOI. RESULTS: One hundred studies with VOI applications were identified. Amongst these, cancer was the most popular indication (25%) and the majority had a UK perspective (49%). The number of publications gradually rose after 2005 but remained steady since 2009. Rarely, studies were sponsored by industry (8%). Expected value of perfect (parameter) information (EVPI/PPI) was reported in 92% (59%) of the articles, respectively. Only 10% of studies reported expected value of sample information (EVS) and 4% reported expected value of perfect implementation (EVPI). Finding a lack of actual application on future research had been reported. Only the UK and Netherlands recommend the use of VOI in the HTA guideline. However VOI is not a formal requirement in the manufacturer/sponsor submission of evidence to NICE or Czv. All experts reported lack of influence of VOI in research decision. CONCLUSIONS: The application of VOI to inform HTA and research is limited. Even in the UK and Netherlands, where VOI analysis is recommended in the HTA guidelines.

PRM99

PREDICTION OF MORTALITY IN THE PRESENCE OF TIME-DEPENDED COVARIATES: AN APPLICATION FOR HEALTH ECONOMIC PROJECTIONS

Exevides A, Colly C

ECO Lane Phase & Outcomes Research, San Francisco, CA, USA

OBJECTIVES: We want to develop a parametric model to predict mortality for future patients with a disease that have specific demographic and clinical characteristics while considering patient trajectories over time for a set of biomarkers, which are critical predictors of disease progression. This is an important tool for many health economic projections. METHODS: In time-to-event studies, longitudinal measures are collected for important disease progression biomarkers. Using only the last available value of these measures in survival models discards important information from the longitudinal evolution. We used data from a 3-year observational study to estimate the covariance coefficients in a Cox-proportional hazards model in the presence of time-dependent biomarkers. In addition, we used a real-world accelerated failure time model to estimate the scale/shape of a parametric survival distribution using SAS®LIFEREG, which, unlike PHREG, does not allow for timedependent measures. In dependence, we, first, applied this Cox/LIFEREG model, we combined the coefficients from PHREG and the scale/shape from LIFEREG to compute the probability of survival. RESULTS: By applying the Weibull model without considering patient trajectories over time, we predicted a 3-year, survival, rate of 55.1%. However, our hybrid combination approach of the Cox/Weibull model, predicted a more accurate 3-year survival rate of 46.7%, which fell within the confidence bounds of the original observational study. CONCLUSIONS: By ignoring the additional variability of patient trajectories over time, when modeling survival, can lead to biased estimates. We have implemented a hybrid approach by which we incorporated the impact of time-dependent biomarkers of the disease along with the scale/shape of a parametric survival distribution to more accurately project survival time in health economic modeling.

PRM100

MIXTURE SURVIVAL MODELING FOR HETEROGENEOUS PATIENT POPULATION

Qian Y1, Dizon D2
1Amgen Inc., Thousand Oaks, CA, USA, 2University of Miami, Miami, FL, USA

OBJECTIVES: It is not uncommon to have heterogeneous patient populations in clinical trials while lacking established biomarkers/factors to identify heterogeneity. In estimating average survival time, a critical component in decision analysis of limited trial data is often necessary and standard parametric survival models are usually regularized without considering population heterogeneity. Our project was to introduce mixture modeling that incorporated population heterogeneity. METHODS: Heterogeneous survival data were simulated. Mixture modeling was applied to address population heterogeneity via Bayesian inference, where clinical inputs (e.g., survival of one sub-population is longer than the other) can also be incorporated. The model fit was evaluated and the estimated survival was assessed. RESULTS: Two-hundred patient level survival data were simulated from a mixture of two exponentials with average survival of 3 (30% of patients) and 0.6 year respectively and overall average survival of 1.32 years. The simulation was run 100 times. For each dataset, Bayesian inference was based on 10000 iterations after 1000 burn-in. In one exercise, the standard approach, including exponential, Weibull and lognormal models, and a mixture of two exponential models were applied. The mixture model provided a good fit per DIC and the estimated proportion was 0.372 (SD: 0.13). The estimates for average survival were compared across models (true mean: 1.32; exponential: 1.238 [SD: 0.142]; Weibull: 1.260 [SD: 0.151]; lognormal: 1.760 [SD: 0.311]; exponential mixture: 1.311 [SD: 0.129]). The extension to mixture models with different components (number or structure) could be implemented. CONCLUSIONS: The mixture modeling could potentially improve the estimates for average survival time in heterogeneous populations. Though this doesn’t apply in every situation, heterogeneity should be considered based on clinical inputs, and tools that address heterogeneity from both clinical and statistical perspectives should continue to be developed to support survival modeling.

PRM101

A METHOD TO ESTIMATE DISEASE-STAGE-SPECIFIC SURVIVAL USING DATA OBTAINED FROM A PREVALENT COHORT

Tabaettit1, Glob MD2, Barcelona, Spain

OBJECTIVES: Technical, ethical, and practical reasons may hamper efforts to conduct clinical trials while lacking established biomarkers/factors to identify heterogeneity. In estimating average survival time, a critical component in decision analysis of limited trial data is often necessary and standard parametric survival models are usually regularized without considering population heterogeneity. Our project was to introduce mixture modeling that incorporated population heterogeneity. METHODS: Heterogeneous survival data were simulated. Mixture modeling was applied to address population heterogeneity via Bayesian inference, where clinical inputs (e.g., survival of one sub-population is longer than the other) can also be incorporated. The model fit was evaluated and the estimated survival was assessed. RESULTS: Two-hundred patient level survival data were simulated from a mixture of two exponentials with average survival of 3 (30% of patients) and 0.6 year respectively and overall average survival of 1.32 years. The simulation was run 100 times. For each dataset, Bayesian inference was based on 10000 iterations after 1000 burn-in. In one exercise, the standard approach, including exponential, Weibull and lognormal models, and a mixture of two exponential models were applied. The mixture model provided a good fit per DIC and the estimated proportion was 0.372 (SD: 0.13). The estimates for average survival were compared across models (true mean: 1.32; exponential: 1.238 [SD: 0.142]; Weibull: 1.260 [SD: 0.151]; lognormal: 1.760 [SD: 0.311]; exponential mixture: 1.311 [SD: 0.129]). The extension to mixture models with different components (number or structure) could be implemented. CONCLUSIONS: The mixture modeling could potentially improve the estimates for average survival time in heterogeneous populations. Though this doesn’t apply in every situation, heterogeneity should be considered based on clinical inputs, and tools that address heterogeneity from both clinical and statistical perspectives should continue to be developed to support survival modeling.
their disease stage. The methodological framework here presented uses prevalent cohort study data to reconstruct stage-specific survival rates, as they would be if obtained from an incident cohort. METHODS: This approach assumes incident individuals and, under certain caveats, can be generalized to non-proportional hazards. Through a piece-wise back calculation approach recent survival data are fed and integrated into the model to calculate survival at earlier time periods. Thus, this allows us to reconstruct the evolution of mortality probability as a function of the time elapsed from the onset of illness. This method is then generalized to allow comparison between ‘placebo’ and ‘treatment’ prevalent-cohorts. RESULTS: The method is very general and can be applied to any databases. As an example, the only published data on an untreated tuberculosis prevalent cohort is used to infer the actual mortality rates of untreated tuberculosis through the course of the illness. The results show that mortality was significantly higher in the first two years of illness, declining dramatically immediately after and increasing again to intermediate rates after one year, suggesting population heterogeneity with respect to the risk of dying from tuberculosis. Further analysis of these data on a sub-sample of patients with ‘treatment’ cohorts is available, this method can be used to infer the potential stage-specific effectiveness of a treatments. CONCLUSIONS: Under specific modeling assumptions, incident-cohort survival and treatment effect can be inferred from prevalent-cohort studies.

PRM102 A SOLUTION FOR UNDER-DIAGNOSED POST-MENOPAUSE: PROBABILISTIC LINKAGE OF CLAIMS AND REGISTRY DATA

Baser Q1, Wu J, Kariyubo F, Xie L1

1StatInMed Research/The University of Michigan, Ann Arbor, MI, USA, 2StatInMed Research, Ann Arbor, MI, USA

OBJECTIVES: Post-menopausal women are under-diagnosed in claims data even though they exhibit symptoms. Probabilistic linkage was conducted between claims data from the University of Michigan Women’s Registry data to control for post-menopausal symptoms as risk factors in the analysis. METHODS: Women with a menopause diagnosis who used estrogen only hormone therapy (HT) were selected from the registry data. A separate group of patients aged 45 or older who were prescribed estrogen only HT was selected for a larger U.S. claims database. Logistic regression was conducted to calculate the propensity score for each patient, controlling for women’s health-related comorbidities in each population. Patients with the closest propensity score from each group were matched, and the menopause symptoms of Registry patients were added to the claims records of the corresponding matched patients in the claims data. Linkage process was repeated 250 times, and the mean and 95% confidence interval (CI) of health care costs in claims data. Probabilistic linkage can bridge the gap between claims and registry data information.

RESEARCH ON METHODS – Patient-Reported Outcomes Studies

PRM103 A STUDY TO EXAMINE THE EQUIVALENCE OF THE PAPER AND ELECTRONIC VERSIONS OF THE PSORIASIS SYMPTOM INVENTORY (PSI)

Bushell DM1, Martin ML1, Scanlon M1, Chau D1, Viswanathan HM1

1Health Research Associates, Inc., Seattle, WA, USA, 2Amgen Inc., Thousand Oaks, CA, USA

OBJECTIVES: To evaluate the equivalence of electronic and paper versions of the Psoriasis Symptom Inventory (PSI) and examine measurement properties of the electronic version. METHODS: In a prospective, randomized, cross-over, non-interventional study in adult subjects (age ≥ 18 years) with plaque psoriasis completing over (3.6) 30 days, subjects were randomized to two groups, each completing either the paper or electronic PSI daily for 7 consecutive days followed by the alternate version. The disease specific Dermatology Life Quality Index (DLQI)) and the generic SF-36v2 were also administered for evaluation of construct validity. Equivalence was assessed by the intra-class correlation coefficient (ICC) between both modes. Differences in scores between the administration modes were also tested using paired Student’s t-test. Measurement included internal consistency reliability, test-retest reliability, and construct validity. Convergent and discriminant validity were evaluated using pre-identified logical relationships between similar or distally related instruments. RESULTS: Eighty subjects (80/80) modestly psoriasis (26% [21/80] mild psoriasis receiving systemic treatment) were enrolled from 8 sites in the United States. Seventeen (21%) subjects experienced a change in sPGA and were excluded from analyses. The paper and electronic modes of the PDI model yields a latent utility index. Scoring system was reviewed and indices risk of order of completion for the paper and electronic versions. All mean score differences were non-significant (p > 0.05) except for “flaking”, however, based on its reliability, validity, and clinical relevance, the item was retained. Minimum values for reliability (>0.70) and validity (convergent, r ≥ 0.30) were exceeded for the electronic PSI. CONCLUSIONS: Equivalence between paper and electronic versions of the PSI and the two modes of comparable treatment and health-related outcomes indicated successful migration from paper to electronic format of the PSI.

PRM104 CLINICAL MEASUREMENT CONCEPTS IN ACUTE SKIN AND SKIN STRUCTURE INFECTIONS AND OTHER SKIN ABNORMALITIES: A COMPREHENSIVE LITERATURE REVIEW

Camin TA1, Debussk K1, Howard K1, Siciak JA1, Llorens L1, Crawley J, Haling K1

1The FHNI Biometrics Consortium CABP ASSBSSI Project BC, Powers JH1

2Outcomes Oxford Ltd, an ICON plc Company, San Francisco, CA, USA, 3Foundation for the National Institutes of Health, Bethesda, MD, USA; 4Cytiva, Inc; Oakland, CA, USA, 5AstraZeneca Pharmaceuticals LP, Wilmington, DE, USA, 6AstraZeneca, Malmö, Sweden, 7The Foundation for the National Institutes of Health, Bethesda, MD, USA, 8National Institute of Allergy and Infectious Diseases (NIAID) National Institutes of Health (NHI), Bethesda, MD, USA

OBJECTIVES: To identify and classify measurement concepts for other skin abnormalities that could be applied to ASSBSSI. The results were used to inform the development of a standardized measurement tool and/or a novel Clinician Reported Outcome (ClinRO) for ASSBSSI. METHODS: To identify relevant articles/searches for inclusion a search was conducted in OVID. MEDLINE (1946-present) and EMBASE (1988-2012) were searched using terms related to ClinRO, measurement tools and devices, diagnostic tools, and skin diseases/abnormalities. RESULTS: The search identified 428 abstracts. 381 were excluded from further scrutiny for eligibility, resulting in 30 that met the inclusion criteria. No ASSBSSI-specific ClinROs or measurement tools were identified in the literature. Several clinical concepts related to the measurement and evaluation of skin diseases/abnormalities were found e.g. Psoriasis Area and Severity Index, Vancouver Scar Scale, Pressure Ulcer Scale for Healing, Pressure Sore Status Tool and the Bates-Jensen Wound Assessment Tool. Disease-specific instruments were also identified. No severity index was found that measures the full spectrum of a skin infection. CONCLUSIONS: There is a paucity of ASSBSSI-specific clinical endpoints/ClinRO measures in the existing literature. This review highlights the need for further research to assess clinical measurement concepts in ASSBSSI. Robust scales, tools, and devices that measure or evaluate skin diseases and infections are identified. These could be considered as candidates for developing a standardized ASSBSSI measurement tool and/or a new ClinRO.

PRM105 VALUING EQ-5D-5L: CAN LATENT UTILITIES DERIVED FROM A DISCRETE CHOICE MODEL BE TRANSFORMED TO HEALTH UTILITIES DERIVED FROM TIME TRADE-OFF TASKS?

Xie F, Pullenayegum E

McMaster University, Hamilton, ON, Canada

OBJECTIVES: The EuroQol Group is evaluating the use of discrete choice experiments (DCEs) in valuing health states from the 5-level EQ-5D (EQ-5D-5L). Notably, the discrete choice (DC) model yields a latent utility index and unbounded, whereas health utilities must have interval properties and be anchored at 0 (representing death) and 1 (representing full health). Latent utilities must, therefore be transformed to health utilities. This study investigated the feasibility of performing such a transformation. METHODS: 545 respondents from Canada and 403 respondents from the UK each completed a series of DC and TTO tasks. Generalised linear mixed models were used to derive latent utilities. Linear regression models incorporating logitistic and polynomial terms, as well as non-parametric LOESS and spline models were assessed as candidate functions for transforming the latent utilities onto the health utilities. RESULTS: There was a high correlation between health utilities measured through TTO tasks and latent utilities derived from modelling of DC data (Spearman rho 0.79 in Canada and 0.86 in the UK). All transforming functions explained the between-state variation in health utilities, and, upon comparison, a small mean difference of 0.01 was found, suggesting that the transformation functions derived through linear regression had the desirable feature of being monotone, the LOESS transform in Canada and the spline transform in the UK lacked monotonicity. CONCLUSIONS: This pilot study suggests that transforming latent utilities to health utilities is feasible, and provides preliminary evidence that linear regression involving polynomial and logarithmic terms may be more desirable than non-parametric spline or LOESS functions.

PRM106 IS A PICTURE WORTH A THOUSAND WORDS? THE DEVELOPMENT OF A STANDARDIZED MEASUREMENT TOOL FOR SCAR SEVERITY

Heil A1, Jensen G1, Galipeau N2, Shields A2

1Amgen Inc., Thousand Oaks, CA, USA, 2Adelphi Values, Boston, MA, USA

OBJECTIVES: Photo-based questionnaires have been used to evaluate aesthetic treatment outcomes such as wrinkle reduction and eyelash growth. The objective of this study was to develop a photometric guide of scar severity intended for use by both clinicians and patients to measure scar treatment