Abstracts

USE OF CLASSIFICATION TREES AS AN AID IN UNDERSTANDING MISSING DATA: AN EXAMPLE FROM AN INTERNET-BASED SURVEY OF PATIENT CHARACTERISTICS AND RESOURCE UTILIZATION

Lily A. J. Khan, S. A. Hayes C* 
Lilly USA, LLC, Indianapolis, IN, USA; Eli Lilly and Company, Indianapolis, IN, USA

OBJECTIVES: Cross-sectional research designs are often employed to answer questions in a timely, cost-effective manner. A common situation is nonresponse, where missing or incomplete data may introduce bias when only the complete data are analyzed. In conducting a survey of type-2 Diabetes Mellitus (T2DM) patients with and without symptoms of ADHD, HbA1c values, a primary measure of diabetes control, were obtained from 314 of 567 (55%) cases. Classification trees, a data-mining tool used for predicting categorical responses, were used to examine the covariates associated with those patterns of missing data. METHODS: The study design utilized an internet-based panel of 567 T2DM cases that consented to participate in a survey assessing the level of ADHD symptoms, when present, the impact the ability to manage HbA1c and other diabetes symptoms. In addition to de-identified demographic data, respondents were asked to provide information on diabetes self-care, HbA1c values, ADHD symptoms, and health care utilization resources. Classification trees were used to develop a predictive model for the occurrence of missing HbA1c values. RESULTS: A classification tree analysis revealed that the diabetes self-care score (SC-R total score: n = 567, mean = 62.3, SD = 15.9) was an important predictor of missing HbA1c. Those respondents who had diabetes self-care scores below 51 were less likely to have provided an HbA1c. A high SC-R total score indicates a high level of self-care, and the SC-R total score has been found to correlate negatively with HbA1c, suggesting that better self-care is associated with better control. CONCLUSIONS: Classification trees can be useful in understanding the structure of missing data. This understanding is a first step in an analysis of a survey and can influence decisions on the need for imputation of missing data and the imputation method.

THE POPULATION IMPACT OF CHEMOTHERAPY IN LATE-STAGE PROSTATE CANCER: A SIMULATION STUDY USING TAX327 AND SEER-MEDICARE DATA

Zagar AJ1, Wilks M, Persson U, Nelson C2, He J* 
14th. - The Swedish Institute for Health Economics, Lund, Sweden; 2jonson & jonson Pharmaceutical Services LLC, Rumson, NJ, USA

OBJECTIVES: Suitable long-term data are rarely available when reimbursement decisions are made for T2DM drugs. If properly constructed and applied, economic simulation models can be useful tools for decision-making. To provide confidence that model results reflect reality, the ability of the models to replicate real-world outcomes should be assessed with formal validity tests. The objective of this study was to assess the validity of a new economic simulation model of T2DM. METHODS: The economic simulation model of T2DM is a stochastic, micro-simulation model that considers the Markov health states representing the development and concomitant consequences of key micro- and macro-vascular diabetes complications. Model predictions for 82 outcomes covering 16 different patient populations from six published outcomes studies were generated by replicating the health and demographic profiles of each patient population, applying the study inclusion/exclusion criteria and treatments, and then simulating outcomes for an equivalent length of time. Where published data were incomplete for a study, the most suitable data from other studies were used. Predicted outcomes were plotted against observed outcomes to assess model fit, as all points would lie on the unity line when the predicted and observed values match exactly. The predictive fit was assessed both overall and separately by whether the source data were or were not used in model construction. RESULTS: The best-fitting line through the cluster plot of predicted versus actual outcomes emanated from the origin with slope 1.07. The resulting R2 was 0.95. The model predicted outcomes for internal and external studies equally well, with R2 values of 0.95 and 0.97, respectively. CONCLUSIONS: Results suggest the model corresponds well with actual clinical data, even considering studies not used in model construction. Though model-based results must always be interpreted cautiously, the model can be considered a useful decision-making tool.

A TRANSMISSION-DYNAMIC MODEL TO PREDICT THE INDIRECT BENEFITS OF INFANT VACCINATION WITH 7-VALENT PNEUMOCOCCAL CONJUGATE VACCINE (PCV7) TO OLDER POPULATIONS

Sneden S1, Strutton DM1, Ciuryla V2, Schwartz BJ3, Bottsma M* 
1PharmNet North America LLC, Bethesda, MD, USA, 2Wyeth Research, Lelystad, NL, USA, 3Washington State University, Pullman, WA, USA

OBJECTIVES: Since its introduction in 2000, PCV7 vaccination in infants and young children has dramatically reduced the overall incidence of invasive pneumococcal disease (IPD) in children and adults. We developed an age-structured transmission dynamic model to quantify these direct and indirect benefits of infant PCV7 vaccination based on US surveillance data from the years 2000–2005. In the present analysis, the model was updated to include data from the years 2006 and 2007. METHODS: The mathematical model simulates the acquisition of asymptomatic carriage of pneumococci and the development of fatal and non-fatal IPD among vaccinated and unvaccinated individuals aged <2, 2–4, 5–6, and >65 years old. The model was parameterized by approximating IPD surveillance data, supplemented with published literature. The original model was able to adequately replicate the observed incidence of IPD cases and deaths from the years 2000–2005; By modifying key transmission parameters, the model was recalibrated to fit the most current surveillance data available (2000–2007). RESULTS: From 2001–2005, a decline in IPD was observed in nearly every age group. Predictions from the original model indicated the IPD decline would continue in years 2006 and 2007. However, actual surveillance data show IPD incidence remained stable from 2005–2007, potentially a result of very high – but not complete – vaccine coverage and effectiveness or the continued or increased presence of IPD caused by pneumococcal serotypes not included in PCV7. The recalculation of indirect benefit of infant vaccination to the older, unvaccinated population demonstrating ~35% fewer IPD cases in the >65 year age group five years after PCV7 introduction. CONCLUSIONS: The improved epidemiologic model simulates the observed US surveillance data to date. By modifying key parameters, this model can be used to simulate the impact of various vaccination strategies in the US and other developed countries.