THE NATURE AND SCALE OF INADEQUATE REPORTING OF DICHOTOMOUS OUTCOMES FROM SIX SYSTEMATIC REVIEWS

OBJECTIVES: Inadequate reporting of outcomes is a major problem while performing meta-analysis. The objective of this study is to estimate the nature and scale of inadequate reporting of dichotomous outcomes. METHODS: Reporting quality of dichotomous outcomes (efficacy and safety) was assessed across six systematic reviews conducted in four disease areas namely oncology (non-small cell lung cancer, colorectal cancer and ovarian cancer), overactive bladder, multiple sclerosis and rheumatoid arthritis. Reporting quality was considered inadequate when either number of patients analysed was not reported or percentage was reported improperly. Reporting quality was considered as improper when the number of patients with outcome (n) could not be calculated from the given percentage. This may be the case when more than one value corresponds to same reported percentage (example 5% of 498 could have values between 23 and 27). Analyses were conducted using STATA 9.2.

RESULTS: In total, 6408 reported outcomes were included in the analysis. Inadequate reporting of dichotomous outcomes was frequent and observed for 691 (10.78%) outcomes. The estimate of inadequate reporting varied across selected reviews and ranged from 7.36% to 17.41% in overactive bladder and multiple sclerosis, respectively. Estimate of inadequate reporting was similar for the safety and efficacy outcomes (10.53% and 11.19%, respectively). Improper reporting of percentage contributed for approximately two-thirds (65.56%) of the inadequately reported outcomes whereas for the remaining outcome N was missing. CONCLUSIONS: Inadequate reporting of dichotomous outcome was frequently observed among the selected reviews. It was observed that reporting of percentages in published reports are often imperfect (due to rounding of values). This leads to loss of data available for meta-analyses as number of patients with outcome could not be calculated accurately. Our results are indicative of outcome reporting bias which needs to be investigated further.


OBJECTIVES: ATLAS.ti software was designed for the qualitative analysis of textual, graphical, audio, and video data. Using ATLAS.ti, qualitative data can be systematically analyzed and the relevant data extracted. It is used extensively in the analysis of data from in-depth interviews and focus groups. We explored its use in extracting, analyzing, and synthesizing literature from a large, focused review of pain relief.

METHODS: A comprehensive search of studies published between 1998–2008 was conducted in Embase.com comparing outcomes and adverse events following different pain relief treatments across a range of therapeutic areas. Relevant studies were imported into ATLAS.ti and were coded and analyzed. On completion of each study review and data coding, ‘outcomes’ for codes were extracted. ‘Outcomes’ contained all data related to a code and could consist of either singular codes such as ‘therapeutic area’ or combinations of codes such as ‘outcome’ and ‘superiority’ (of one drug against another). Data were categorized according to emergent themes regarding outcomes and adverse events (AEs) associated with the pain relief treatments included in the review.

RESULTS: Following a 2-stage screening process, 232 relevant, published studies were imported into ATLAS.ti for full review and coding. Information on sample size, therapeutic area, outcomes, adverse events, and other trial data was coded. ATLAS.ti enabled identification of the most common therapeutic areas where the target product had been studied: rheumatology, acute pain, chronic pain and urology. Further, specific comparisons between the target product and other pain relievers (COX-2, NSAIDs, and opioids) on efficacy and AEs were possible by coding key results from the studies. Versatile cross-comparisons were possible because of the detailed coding. CONCLUSIONS: ATLAS.ti can be a useful tool for extracting and synthesizing a large volume of literature. It enabled development of new key product messages by allowing fore emergent themes from a large, published literature.

REDUCING BIAS IN A RETROSPECTIVE CASE-CONTROL STUDY: AN APPLICATION OF PROPENSITY SCORE MATCHING.

OBJECTIVES: The goal of this study was to perform propensity score matching analysis to evaluate the treatment effects of varenicline. We present the derivation of propensity scores, selection of controls, and comparison of key patient characteristics before and after matching.

METHODS: We utilized the largest hospital service-level database in the U.S. We identified 2,588,722 adult patients undergoing inpatient echocardiography between January 2003 and October 2005 of which 2,900 had diagnoses for critical illness (heart failure, acute myocardial infarction, arrhythmia, respiratory failure, pulmonary embolism, emphysema, and pulmonary hypertension) and who also received a contrast agent (Perflutren Protein-Type A Microspheres Injectable Suspension, USP). Patients receiving other contrast agents were excluded from the study. To control for differences between patients receiving contrast echocardiography (cases) to those who received non-contrast echocardiography (controls), we used propensity score matching. A stepwise logistic regression was used to model treatment choice (contrast vs. non-contrast). Variables used in the construction of the propensity score included comorbidities, demographic factors, hospital-specific factors, level of care, and mechanical ventilation status. Cases were matched to 4 control patients by nearest neighbor matching algorithm based on differences in propensity scores among cases and controls. RESULTS: The nearest neighbor matching algorithm successfully identified 4 matches for each of the 2900 contrast patients. Prior to matching, 23 of the 2900 statistically significant differences between cases and controls (P < 0.01). These characteristics included mechanical ventilation status, ICU status, and the Deyo-Charlson Comorbidity Score. After matching, one variable remained statistically significant (higher concomitant medication usage among cases; P = 0.006). CONCLUSIONS: The proposed score matching using the nearest neighbor matching algorithm was able to reduce the differences between cases and controls. This improved the precision of the estimates.

HOW TO MAKE USE OF AVAILABLE SURVIVAL EVIDENCE IN AN INDIRECT COMPARISON

OBJECTIVES: The aim of this study was to evaluate the performance of different imputation techniques.

METHODS: A total of 2900 contrast patients were matched to 4 non-contrast patients. Survival time was defined as from time of echocardiography until death from any cause. Survival data was obtained in interviews and focus groups. We explored its use in extracting, analyzing and the relevant data extracted. It is used extensively in the analysis of data obtained in interviews and focus groups. We explored its use in extracting, analyzing, and synthesizing literature from a large, focused review of pain relief.

RESULTS: In this analysis, we had 2900 contrast patients and 4000 non-contrast patients. We identified 4 matches for each of the 2900 contrast patients. Prior to matching, 23 of the 2900 statistically significant differences between cases and controls (P = 0.001). These characteristics included mechanical ventilation status, ICU status, and the Deyo-Charlson Comorbidity Score. After matching, one variable remained statistically significant (higher concomitant medication usage among cases; P = 0.006). CONCLUSIONS: The proposed score matching using the nearest neighbor matching algorithm was able to reduce the differences between cases and controls. This improved the precision of the estimates.

OBJECTIVES: Therapies for oncology often effect time-to-event statistics like overall survival and progression free survival in literature. In this literature, these time-to-event statistics are summarized by median time-to-event, percentage of people having had the event at a specific point in time and hazard rates, among others. Although pooling based on individual patient data would be preferred, we still encounter situations where only aggregated data is available. For these situations, the hazard ratios can be pooled. However, many high quality papers would be ignored, if publications were only regarded once the hazard ratio is presented. Therefore we searched for methods to transform the different outcomes to the same scale and to pool based on as much information as possible. METHODS: A review was performed with respect to the pooling of different time-to-event outcomes. Mixed Treatment Comparisons were performed using the methods to assess their usability. RESULTS: For cost-effectiveness models, a distribution (exponential, weibull, among others) of the time-to-event statistic is often used to obtain the average of the time-to-event statistic for the comparison arm. This distribution can also be used to transform the median and the ‘percentage having had an event at a specific time point’ into hazard rates. Comparing the transformed medians with the hazard rates for publications in which both are presented implied a way to check the validity of assumptions. CONCLUSIONS: Although the way information about time-to-event statistics is presented may differ across publications, it is often possible to pool the different types of information. This implied the inclusion of papers which otherwise couldn’t be used and a reduction of the uncertainty in the cost-effectiveness outcomes.

CONCEPTUAL PAPERS & RESEARCH ON METHODS – Cost Methods

OBJECTIVE: To address the common issue of incomplete data in cost-effective analysis, we imputed missing cost components at the patient level using multiple imputation (MI) techniques. METHODS: A study cohort with concomitant medication, hospitalization and outpatient-visit costs was derived from the population of a randomized clinical trial comparing two treatments. On a total of 132 subjects without missing data, a pattern of missingness was created so that 25% of subjects had missing hospitalization costs and 50% (including the above 25%) had missing concomitant medication costs. The average total costs (sum of the three components) obtained using MI techniques (propensity score [PS], regression and Markov Chain Monte Carlo [MCMC]), complete case analysis (CCA) and available case analysis (ACA) were compared with actual costs. In imputation models, response variables were the log-transformed costs. The average total costs made up of three cost components are given by treatment A versus treatment B was actually more expensive than treatment A. CCA gave opposite results; while treatment B was actually more expensive than treatment A. CONCLUSIONS: Total costs made up of three cost components are given by treatment A versus treatment B. Actual total costs: $16,327 vs. $18,484. Estimated costs: CCA: $17,317 vs.$15,400; ACA: ($13,407 vs. $15,361; MI-Ps: $16,941 vs. $19,156; MI-Regression: $16,404 vs. $19,056; MI-MCMC: $16,584 vs. $18,947. CONCLUSIONS: Treatment B was actually more expensive than treatment A. CCA gave opposite results; while ACA underestimated total costs. Regression gave better results than PS, as a regression model was fitted for each missing cost component, with the previous variables as covariates. MCMC using Bayes’ theorem with a non-informative prior makes assumption of missing at random, MCMC could be a useful imputation technique compared with actual costs. In imputation models, response variables were the log-transformed costs. The average total costs made up of three cost components are given by treatment A versus treatment B was actually more expensive than treatment A. CCA gave opposite results; while treatment B was actually more expensive than treatment A. MI-MCMC using Bayes’ theorem with a non-informative prior makes assumption of missing at random, MCMC could be a useful imputation technique compared with actual costs.