ent to materially affect ICERs. More conservatively, assuming no treatment effect beyond 10-years still resulted in 0.35 life years gained, a significant gain compared to the 10-year horizon. Varying assumptions in different ways altered the magnitudes of the gains in LYs (and potentially cost-effectiveness), but not the essential conclusions. **CONCLUSIONS:** This research confirms Gray's suggestion of the importance of extending analysis time horizons when differential mortality is observed at the end of a study. Under any reasonable assumption applied to the extrapolation, any survival difference at end of study must persist to some degree beyond that time and therefore add to the treatment benefit observed up to the point of extrapolation. Ignoring the post-study period biases clinical and cost-effectiveness results.

PRM5

COMPARING DIAGNOSIS-BASED AND PRESCRIPTION-BASED COMORBIDITY MEASURES FOR PREDICTING HEALTH SERVICE UTILIZATION AND COSTS

Gangan N, Banahan B III
University of Mississippi, University, MS, USA

OBJECTIVES: Comorbidity scores are frequently used for controlling confounding in observational studies using administrative claims. Several comorbidity measures have been developed and evaluated for predicting a variety of outcomes. However, scores are often used with outcomes other than those included in validation studies. Few publications have compared the performance of different scores in predicting different types of outcomes. The objective of the present study was to compare the Charlson Comorbidity Index (CCI), a frequently used diagnosis-based comorbidity score, and Rx Risk, a prescription-based comorbidity score, as predictors of three outpatient outcomes. **METHODS:** A retrospective analysis was conducted using Mississippi Medicaid medical and pharmacy claims data for the period January 2010 to December 2011. Inclusion criteria were continuous enrolment during the observation period, not dual-eligible, and having both medical and pharmacy claims in both calendar years. CCI scores and Rx Risk Scores were calculated using 2010 claims. Scores were evaluated as predictors of outpatient visits, total outpatient costs and total pharmacy costs in 2011. Costs were log transformed. A base general linear model with age, gender, race and ethnicity was developed. Predictive ability of each comorbidity score was measured as the change in R2 when the score was added to the base model. **RESULTS:** R2s for the base model were visits - 0.07, outpatient costs - 0.01, and pharmacy costs - 0.03. CCI and Rx Risk improved prediction for visits and pharmacy costs (CCI R2s; 0.10, 0.05; Rx Risk R2s; 0.13, 0.07). **CONCLUSIONS:** Although CCI is often used for outpatient outcomes, Rx Risk provides a better measure of comorbidity when the dependent variables are outpatient utilization or costs. The CCI was developed for predicting mortality during hospitalization. These results indicate that comorbidity scores developed for predicting outpatient outcomes would be better for controlling for comorbidity in outpatient based studies.

PRM6

MATCHING-ADJUSTED INDIREC TREATMENT COMPARISON AND SURVIVAL EXTRAPOLATION IN RADIOIODINE-REFRACTORY DIFFERENTIATED THYROID CANCER (RAI-REFRACTORY DTC)

Tremblay G[1], Holbrook T[2], Milligan G[2], Pelletier C[1]
[1]*Eisai, Woodcliff Lake, NJ, USA,* [2]*Adelphi Real World, Manchester, UK*

OBJECTIVES: Indirect treatment comparisons (ITCs) are an important part of any comparative effective demonstration in the absence of head-to-head clinical trials. Classic ITCs can lead to biased results due to differences in patient populations and trial designs. These differences can be corrected for by using matching-adjusted ITC (MAIC) technique. Furthermore, extrapolation of survival data beyond clinical trial results may be required for economic evaluations. The objective of this research was to compare lenvatinib and sorafenib in patients with RAI-Refractory DTC using MAIC and survival extrapolation techniques. **METHODS:** Mean overall survival (OS) and progression-free survival (PFS) outcomes were estimated by weighting lenvatinib's patient level data based on baseline characteristics from sorafenib phase 3 trial using logistic regression. Classic ITC was performed before and after adjustment. Extrapolation of OS and PFS was performed using proportional hazard, accelerated time failure, individual parametric models and piecewise models (Royston & Parmar). Results were presented as hazard ratios (HR) with confidence intervals (CI). **RESULTS:** Unadjusted ITCs for Lenvatinib vs. placebo were 0.746(0.497; 1.119) for OS and 0.213(0.158; 0.288) for PFS. The MAIC provided statistically significant estimates of 0.577 (0.347; 0.959) for OS and 0.170(0.118; 0.254) for PFS vs. placebo. Unadjusted ITCs vs. sorafenib were 0.933(0.529; 1.643) and 0.362(0.245; 0.536) respectively for OS and PFS; while MAIC results were 0.721(0.379; 1.373) and 0.325(0.201; 0.526) respectively for OS and PFS. Survival extrapolation provided estimates of 7.5-10 month of additional OS gain for Lenvatinib vs. placebo, with the MAIC extrapolation showing the largest gain and a good model fit. **CONCLUSIONS:** This analysis demonstrated that in absence of head-to-head trials, MITC offers important methodology to adjust for population and trial differences, especially in orphan diseases where limited data are available. MAIC can increase the reliability of comparative effectiveness data and support payers decision making.

PRM7

PROSPECTIVE BENEFIT-RISK MONITORING OF NEW DRUGS FOR RAPID ASSESSMENT OF NET FAVORABILITY IN ELECTRONIC HEALTHCARE DATA

Gagne JJ[1], Bykov K[2], Najafzadeh M[1], Choudhry NK[1], Martin DP[3], Kahler K[3], Rogers JR[2], Schneeweiss S[1]
[1]*Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA,* [2]*Brigham and Women's Hospital, Boston, MA, USA,* [3]*Novartis Pharmaceuticals Corporation, East Hanover, NJ, USA*

OBJECTIVES: Benefit-risk assessment (BRA) methods can combine measures of benefits and risks into a single value. We examined BRA metrics for prospective monitoring of new drugs in electronic healthcare data. **METHODS:** Using two databases, we emulated prospective monitoring of three drugs versus comparators (rofecoxib vs. non-selective non-steroidal anti-inflammatory drugs [ns-NSAIDs], prasugrel versus clopidogrel, and denosumab versus bisphosphonates) beginning at market entry of each drug of interest and using a sequential propensity score matched cohort design. We applied four BRA metrics: number needed to treat and number needed to harm (NNT|NNH); incremental net benefit (INB) with maximum acceptable risk [MAR], INB with relative-value adjusted life years [RVALYs], and INB with quality-adjusted life years [QALYs]. We determined whether and when the bootstrapped 99% confidence interval (CI) for each metric excluded zero, indicating net favorability of one drug over the other. **RESULTS:** For rofecoxib, all four metrics yielded a negative value, suggesting net favorability of ns-NSAIDs over rofecoxib, and the 99% CI for all but the NNT|NNH excluded the null during follow-up. For prasugrel, only the 99% CI for INB-QALY excluded the null, but trends in values over time were similar across the four metrics, suggesting overall net favorability of prasugrel versus clopidogrel. The 99% CI for INB-RVALY and INB-QALY excluded the null in the denosumab example, suggesting net favorability of denosumab over bisphosphonates. **CONCLUSIONS:** Prospective benefit-risk monitoring can be used to determine net favorability of a new drug in electronic healthcare data. In three examples, existing BRA metrics produced qualitatively similar results but differed with respect to alert generation. INB-QALY produced the most conclusive findings across the three examples.

PRM8

QT PROLONGATION IDENTIFICATION IN RETROSPECTIVE STUDIES

Ye Y, Caffrey AR
University of Rhode Island, Kingston, RI, USA

OBJECTIVES: To evaluate operational definitions for cardiac events related to QT prolongation, such as paroxysmal ventricular tachycardia, ventricular fibrillation and flutter, cardiac arrest, and sudden cardiac death, in retrospective studies using administrative databases. **METHODS:** Using PubMed, we searched for studies that retrospectively identified cardiac events related to QT prolongation in administrative or claims databases and were published between January 2000 and September 2014. Selection for full-text review was based on a preliminary review of titles and abstracts. **RESULTS:** Our initial search yielded 988 articles from which five were selected for inclusion after full-text review. Case report, clinical trial, congenital long QT syndrome, cardiac event not related to QT prolongation, and electrocardiography utilization are reasons for exclusion. Seven additional articles were identified from the references of these articles. The twelve included articles consist of four cohort studies (33%), three case-control studies (25%), three validation studies (25%), and two descriptive studies (17%). Nine studies (75%) utilized databases from the United States, five (42%) of which used Medicaid data, and three (25%) used European data. The most common operation definitions for cardiac events related to QT prolongation were primary discharge diagnosis of long QT-related cardiac events (75%) and sudden cardiac death (25%). The most common administrative codes utilized were ICD-9 (83%) and ICD-10 (17%). The most frequently utilized ICD-9 diagnosis code was 427.x (100%, cardiac dysrhythmias, ICD-10: I47-49), followed by 426.x (33%, conduction disorders, ICD-10: I44-45), and 798.x (33%, sudden death, cause unknown, ICD-10: R96). Six studies (50%) reviewed medical records to validate the diagnosis codes. Positive predictive values ranged from 77-94% when defining cardiac events related to QT prolongation using ICD-9 codes 426.x or 427.x. **CONCLUSIONS:** In administrative databases, ICD-9 codes 426.x and 427.x as the principle discharge diagnosis or underlying cause of death are commonly used to identify cardiac events related to QT prolongation.

PRM9

COMPARISON OF BENEFIT-RISK ASSESSMENT METHODS FOR PROSPECTIVE MONITORING OF NEWLY MARKETED DRUGS: A SIMULATION STUDY

Gagne JJ[1], Najafzadeh M[1], Choudhry NK[1], Bykov K[2], Kahler K[3], Martin DP[3], Rogers JR[2], Schneeweiss S[1]
[1]*Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA,* [2]*Brigham and Women's Hospital, Boston, MA, USA,* [3]*Novartis Pharmaceuticals Corporation, East Hanover, NJ, USA*

OBJECTIVES: We compared benefit-risk assessment (BRA) methods for determining whether and when sufficient evidence exists to indicate that one drug is favorable over another in prospective monitoring. **METHODS:** We simulated prospective monitoring of a new drug (A) versus an alternative (B) with respect to two beneficial and three harmful outcomes. We generated data for 1,000 iterations of six scenarios and applied four BRA metrics: number needed to treat and number needed to harm (NNT|NNH); incremental net benefit (INB) with maximum acceptable risk (INB-MAR); INB with relative-value adjusted life years (INB-RVALY); and INB with quality-adjusted life years (INB-QALY). We determined the proportion of iterations in which the 99% confidence interval (CI) for each metric included and excluded the null and we calculated mean time-to-alerting. **RESULTS:** With no true difference in any outcome between drugs A and B, the proportion of iterations including the null was lowest for INB-RVALY (64%) and highest for INB-QALY (76%). When drug A was more effective and the drugs were equally safe, INB-QALY indicated net favorability of drug A in 81% of iterations, INB-MAR and INB-RVALY indicated net favorability in 79% of iterations, and NNT|NNH indicated net favorability in 72% of iterations. When drug A was safer than drug B, NNT|NNH had the highest proportion of iterations indicating net favorability of drug A (65%). Mean time-to-alerting was similar among methods across the six scenarios. **CONCLUSIONS:** BRA metrics can be useful for identifying net favorability when applied to prospective monitoring of a new drug versus an alternative. INB-based approaches similarly outperform unweighted NNT|NNH approaches.

PRM10

USAGE OF PROPENSITY SCORE, INSTRUMENTAL VARIABLE, OR MACHINE LEARNING FOR REAL WORLD DATA ANALYSIS

Choi S
PAREXEL International, Billerica, MA, USA

OBJECTIVES: It is essential to reduce potential bias by adjusting for confounders when performing real world data analysis. It is informative to investigate usage of