OBJECTIVES: For reimbursement, economic evaluations such as healthcare costs are often as important as clinical outcomes. However, estimating the standard error (SE) of the difference in mean costs between two treatments using the zero-inflated Gamma distribution is currently limited to bootstrapping and the method of Mills (2013). We propose an exact, closed-form solution for estimating SE based on a parametric model, and compare the results to those of bootstrapping and Mills' method. **METHODS:** Data were generated using the gamma distribution with varied shape parameters (0.5, 1, 2, 5, 7.5, 10, 20) and a fixed scale parameter(1000) The proportion of zeroes for costs in each data set were also varied (0.5, 0.7, 0.9). Each simulation data set contained 10,000 observations. We calculated the SE's of the mean difference in cost between two treatments, using three methods. **RESULTS:** For small shape parameters (0.5, 1, 2), the SEs of all methods showed similar results across different proportions of zeroes (0.5, 0.7, 0.9), with the SEs from the Mills method being slightly larger than that of the bootstrap method (10000 repetitions) or our exact method. However, when the shape parameter was greater than 2, the SEs estimated by the Mills method were significantly larger than that of the other methods. Our proposed approach produced similar SEs compared to those of the bootstrap method across all different proportions of zeroes. **CONCLUSIONS:** Our proposed method has a nice property: it is not only an exact method based on a parametric model, but also produces the similar SEs to the bootstrap method. The Bootstrap method is a consistent estimator, though it is not exact and computational costs are a consideration. One should use the Mills method with care, because the estimated SE may be significantly larger across a range of parameters.

PRM109

A COMPARISON OF PATIENT LEVEL DATA (IPD) REGRESSION TECHNIQUES IN ADJUSTING FOR CROSS TRIAL DIFFERENCES IN AN INDIRECT TREATMENT COMPARISON (ITC): A WORKED EXAMPLE OF SIMULATED TREATMENT COMPARISON (STC) VS BAYESIAN IPD REGRESSION

Shukla P, Sharma A, Siddiqui MK
PAREXEL International, Chandigarh, India

OBJECTIVES: In the absence of Head-to-head clinical trials population adjusted ITC such as matching adjusted indirect comparison (MAIC) and STC can account for between-trial imbalances in the distribution of effect modifiers. STC is a regression-based approach which predicts the outcome in the target population. As ITC conducted in a Bayesian setting has been increasing in popularity, our objective is to bring together the flexibility of Bayesian modeling with STC model by NICE and compare the output. **METHODS:** We simulated the hypothetical data from three trials (treatment B vs. A) and a comparator trial (treatment C vs. A), using R package Wakefield. In line with NICE worked example, we generated two variables, age, and gender. We developed a logistic regression model in WinBUGS for AB trial population and used this to predict the outcome in AC trial population. To take into account the clustering of individuals within the component trials, we introduced a study-level baseline risk term in the outcome model. To compare the results, the log OR of B vs. A in the AC population were then estimated from the two methods. **RESULTS:** Treatment effect estimates obtained from the Bayesian model after adjusting for the different baseline risk were logOR B vs. A: -2.49 (-3.06, -1.95). These results were aligned with the pooled estimates obtained from STC (as per the NICE methods) logOR B vs. A: -2.45(-2.99, -1.90). **CONCLUSIONS:** Bayesian IPD regression with study level baseline terms added in the outcome model can be used to combine IPD from multiple trials when population adjusted ITCs are required. Our findings showed that STC results from both frequentist and Bayesian frameworks are aligned with the latter offering more flexibility. STC can also be an option to explore when MAIC is not feasible due to small effective sample size.

PRM110

HUMANISTIC AND ECONOMIC BURDEN ASSOCIATED WITH ANXIETY AND DEPRESSION AMONG ADULTS WITH DIABETES AND HYPERTENSION

Wallace K, Misra R, Sambamoorthi U
West Virginia University, Morgantown, WV, USA

OBJECTIVES: This study estimated the humanistic and economic burden associated with co-morbid depression and anxiety among adults with diabetes and hypertension. **METHODS:** A retrospective cross-sectional study was conducted among adults (>= 18 years) with diabetes and hypertension. Respondents were classified into four groups: 1) depression and anxiety (N=309), depression only (N=561), anxiety only (N=366), and no depression and no anxiety (N=3,324) using data from the Medical Expenditure Panel Survey for the years 2013 and 2015. The humanistic outcomes were: health-related quality of life (HRQoL) measures obtained from the SF-12-V2. Economic outcomes were: total annual healthcare expenditures, third party spending, out-of-pocket (OOP) spending and proportion of income spent out-of-pocket. Counter-factual recycled prediction method was used to assess the incremental burden associated with comorbid depression and anxiety. **RESULTS:** Nationally, 1.57 million (7.7%) adults with diabetes and hypertension reported having comorbid depression and anxiety and incurred $45.3 billion in total healthcare expenditures. The per-capita total expenditures for those with comorbid depression and anxiety was nearly 250% higher ($28,832 vs $11,543) compared to those without depression and without anxiety. They also experienced high OOP burden (AOR=1.57; 95% CI=1.08, 2.27). The relationship between comorbid depression and anxiety and HRQoL measures were insignificant after adjustment. **CONCLUSIONS:** Management of diabetes and hypertension needs to account for comorbid depression and anxiety.

PRM111

THE IMPACT OF CLUSTER SELECTION METHODS IN TWO-STAGE BOOTSTRAPPING TO ASSESS UNCERTAINTY IN HEALTH ECONOMIC OUTCOMES IN CLUSTER RANDOMIZED CONTROLLED TRIALS

Kip MM[1], Berghuis AS[1], Nijsten MW[2], IJzerman MJ[1], Koffijberg H[1]
[1]*University of Twente, Enschede, The Netherlands,* [2]*University of Groningen, University Medical Center Groningen, Groningen, The Netherlands*

OBJECTIVES: Bootstrapping is often used to assess uncertainty in outcomes of randomized controlled trials (RCTs) due to sampling variation and limited sample sizes. Although guidance is available on two-stage bootstrapping for cluster-RCTs, specific guidance is lacking on sampling clusters within bootstrap samples to address the uncertainty in variation across clusters. This study assesses the impact of using different selection approaches to sample clusters in two-stage bootstrapping in a case study on procalcitonin-based antibiotic treatment in IC patients with sepsis. **METHODS:** The case study was a cluster-RCT including 16 hospitals (4 academic, 12 non-academic) with on average 48 patients per hospital (range n: 1-185). Five cluster sampling approaches were investigated, based on random sampling of: 1) the intended number of patients, 2) 16 hospitals, 3) 16 hospitals maintaining the original ratio academic/non-academic hospitals, 4) as method 2 while maintaining the total number of patients, 5) as method 3 while maintaining the total number of patients. Additionally, a scenario analysis using half of the data was performed. Incremental cost differences and corresponding 95%CIs were determined based on 10,000 bootstrap samples. **RESULTS:** Different approaches of bootstrapping resulted in variation in the incremental costs per patient (data mean: €16, bootstrap range: €-24 - €183), with approach 5 deviating most from the observed mean incremental cost. 95%CIs also varied in size (smallest 95%CI: €-5,123 - €5,986 [method 5], largest 95%CI: €-5,699 - €6,566 [method 2]). Differences in outcomes were more pronounced when using half of the data. **CONCLUSIONS:** Using different approaches for sampling clusters in two-stage bootstrapping may influence the mean outcomes and 95%CIs. Determining the most appropriate sampling method based on outcomes and 95%CIs is dependent on the approach for selection used in the real-world trial. When the inclusion strategy is unknown, sensitivity analysis is recommended to assess uncertainty arising from this unknown cluster inclusion process.

PRM112

IS THERE A CONSENSUS REGARDING CLINICALLY RELEVANT NON-INFERIORITY MARGINS USED FOR KEY ONCOLOGY ENDPOINTS IN NON-INFERIORITY ONCOLOGY TRIALS?

Hashim M[1], He J[2], Hu P[3], Soikkeli F[1], Gebregergish S[1], Heeg B[1], Lam A[2]
[1]*Ingress Health, Rotterdam, The Netherlands,* [2]*Janssen Global Services, LLC, Raritan, NJ, USA,* [3]*Janssen R&D, Raritan, NJ, USA*

OBJECTIVES: Regulatory agencies such as the FDA and EMA have issued guidelines on how to evaluate non-inferiority. However, there is no clear guidance on how to select the non-inferiority margins (NIFm's) in oncology clinical trials. This study is to identify previously used NIFms in oncology clinical trials for key endpoints. **METHODS:** A systematic Medline literature review of NIFms in oncology clinical trials was conducted. Non-inferiority randomized clinical trials published in the English between Jan 2000 and Dec 2017 were included. Only studies reporting margins for the response rate (RR), progression-free survival (PFS), overall survival (OS), and safety endpoints were included. The following data items were extracted: indication, treatment setting, sample size, primary endpoint, treatment effect measure, and defined margin. Study selection and data extraction were performed by two independent investigators. **RESULTS:** Out of 635 screened search hits, 99 reports were included in this analysis. For both RR and safety, the rate difference was used as the measure of treatment effect. For RR, among 18 reports (median sample size: 293; range: 98 to 828), the median and mean of NIFms were 15% and 13%, respectively. For safety endpoint, we identified two studies with a NIFm of 10% and 15%. For both PFS and OS, the hazard ratio was used as the measure of treatment effect. For PFS, among 29 reports (median sample size: 402; range: 85 to 2,098), the median and mean of NIFms were 1.300 and 1.314, respectively. For OS, among 50 reports (median sample size: 586; range:22 to 1,725), the median and mean of NIFms were 1.250 and 1.272, respectively. On average, retrieved studies with a total sample size of 1000 or 500 had larger NIFms. **CONCLUSIONS:** There is no consensus on the appropriate NIFms for key endpoints in oncology trials. More guidance from regulatory agencies is needed.

PRM115

A WRINKLE IN TIME: WHEN SURVIVAL ESTIMATES CHANGE OVER THE OBSERVATION PERIOD

Smith BN
ICON Clinical Research, North Wales, PA, USA

OBJECTIVES: To illustrate the dynamic nature of time, it's potential impact on risk estimates, and approaches to account for time dependencies in time to event analyses. **METHODS:** Survival analysis techniques will be illustrated including time to event concepts, the relation between survival and hazard functions, censoring, and truncation. Further, exploration of time dependencies will be demonstrated and methodological approaches to account for and visualize variations in hazard estimates over time will be shown. **RESULTS:** Descriptive statistics and unadjusted product-limit estimators will be presented including frequencies, chi-square statistics, and Kaplan-Meier estimates to illustrate unadjusted associations. Graphical exploration of time-varying functions will be conducted to display time interactions and violation of proportionality. Cox proportional hazards modeling will be demonstrated and compared to the extended Cox model to account for time dependencies and covariate adjustment. **CONCLUSIONS:** Temporal deviations including loss to follow-up and time-dependencies can occur in even the most controlled research designs and accounting for time offers a unique and more granular view of the outcome in the context of treatment and other covariates. It is incumbent upon researchers to explore study outcomes as a function of time when possible in survival analytic approaches since time dependencies, when unaccounted for, can drastically alter the interpretation of results. Techniques exist to address time-varying hazards when identified and should be employed.