translated into 17 Asian languages (Chinese for China, Hong Kong, Singapore and Taiwan, Korean, Marathi, Malayalam, Punjabi, Telugu, Urdu, Tamil for India and Singapore, and Malay for Singapore). The objective of this study was to identify the main translation challenges, and which items needed the most references to conceptual definitions. METHODS: In each country, the translation validation was conducted with the author of the IRIS-QOL, using either the standard forward/backward methodology or the adjusted process with reviews by the author, a clinician, and cognitive interviews with 12 patients. The basis for the validation was the list of concepts (LOC) elaborated by the author. For each country, the history of the grid translation process was analyzed. References/remarks of the LOC and author’s interventions were counted for each item across countries. RESULTS: Eleven items were found challenging (items 2, 3, 7, 9, 16, 18, 25, 28, 31, 32). Four of them needed the most to refer to the LOC: i.e. item 3 “I am bothered by how much time I spend on the toilet” (100% of the countries), item 18 “I feel I get less done…” (53%), item 25 “I feel sluggish…” (76%), and item 29 “A common complaint of people…” (20%). For the remaining difficulties with item 3 were conceptual (“on the toilet” not to be confused with “in the toilet”), and cultural (e.g., lack of toilets in India, or people squatting, not sitting). Examples of challenges will be provided. CONCLUSIONS: The development of the LOC and the MAIC translation process and enabled the conceptual harmonization of the translations.

RESEARCH ON METHODS – Statistical Methods

PM120 EVALUATION OF MATCHING-ADJUSTED INDIRECT COMPARISON IMPLEMENTED BY A REASAMPLING METHOD
Wang J1, Oborn D2, Chirla C1, Zheng Q1
1BTI Health Solutions, Research Triangle Park, NC, USA, 2BTI Health Solutions, Research Triangle Park, NC, USA
OBJECTIVES: Matching-adjusted indirect comparison (MAIC) has been proposed as a new approach to indirect treatment comparisons (ITC) in the situation where both individual-level data are available from one study, but only summary data are available from another study. This study evaluated the performance of the MAIC method proposed by Malangone and Sherman (2011) which is implemented by a resampling (bootstrapping) technique. METHODS: Two patient-level data sets, similar to two clinical trials, were generated: the first with treatments A and placebo, and the second with treatments B and placebo. Other variables included in both data sets were survival time, censoring indicator, and two baseline categorical variables. In both data sets, interactions between baseline characteristics and treatments were incorporated such that differential treatment effects across baseline strata were present. The SAS programs illustrated in Malangone and Sherman were adopted for the MAIC analysis. When MAIC was applied in which individual-level data were available from the first data set and individual-level data were available from the second data set. Subsequently, the roles of two data sets were switched and the MAIC analysis was applied once again. RESULTS: Using MAIC, when the first data set provided summary statistics, the hazard ratio (HR) (95% confidence interval [CII] for A versus placebo was 0.283 (0.246-0.325); the HR (95% CI) for B versus placebo was 0.586 (0.466, 0.740). When the second data set provided summary data, the HR (95% CI) for A versus placebo was 0.489 (0.390-0.619) and for B versus placebo was 0.237 (0.205-0.273). The two comparisons produced opposite significant inferences. CONCLUSIONS: The method proposed by Malangone and Sherman is an improvement upon the MAIC field, but results could be misleading under some circumstances. Therefore, the conditions under which this method is suitable should be explored further.

PM121 A SYSTEMATIC REVIEW OF THE METHODOLOGICAL QUALITY OF NETWORK META-ANALYSES
Chirila C1, Naci H2, Wouters O2, Pyo J1, Gunjal S3, Kennedy R1, Hoey M4, Wann A5, Neumann PJ
1Tufts Medical Center, Boston, MA, USA; 2London School of Economics, London, UK; 3Precision Health Economics, Austin, TX, USA; 4Royal Victoria Hospital, Belfast, UK; 5The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA, 2Center for the Evaluation of Value and Risk in Health, Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, MA, USA
OBJECTIVES: To evaluate the methodological quality of published network meta-analyses (NMA). METHODS: We performed a systematic review of the NMA literature (through July 2014). For NMAs including pharmacologicals we assessed general study characteristics, study reproducibility and transparency, methods, and reporting of findings. We compared NMAs published in higher impact factor journals with those published in lower impact factor journals, NMAs published before January 1st, 2013, with those published after that date, and studies receiving financial support from industry with those receiving financial support from non-profit institutions or that did not receive support. RESULTS: The systematic literature search identified 318 NMAs meeting our inclusion criteria. Forty-eight percent were published in high impact factor journals, 40% were published before January 1st, 2013, with those published after that date, and studies receiving financial support from industry with those receiving financial support from non-profit institutions or that did not receive support (68%). We found notable inconsistencies among NMAs. Eighty percent reported search terms and journals, 71% mentioned sufficient data to conduct the analysis, and 61% characterized their analysis within specific clinical trials, and 61% the network diagram. Seventy percent reported a risk of bias assessment of included clinical trials, 56% a sensitivity analysis, and 40% an assessment of model fit. Among NMAs with a closed loop, 37% assessed the consistency of direct and indirect evidence; four percent of NMAs presented the complete matrix of head-to-head treatment comparisons. For Bayesian NMAs, 41% reported the probability of each treatment being best, 31% reported efficacy ranking, and 16% reported or referenced the model code. NMAs published in higher impact journals and those that did not receive financial support from industry performed better across our assessment criteria. We did not find substantial differences between NMAs published before January 1st, 2013, with studies published after that date. CONCLUSIONS: There is substantial variation in the NMA literature. Consensus among NMA guidelines is required to improve methodological quality, consistency, and transparency of study conduct and reporting.

PM122 APPLICAION OF DIFFERENCE-IN-DIFFERENCE METHODOLOGY IN COMPARATIVE EFFECTIVENESS RESEARCH WITH UNBALANCED GROUPS
Li Y1, Zou H2, Ko J1, Hawker K1, Nazareth T1, Arcona S3, Sassean R3
1Novartis Pharmaceuticals, East Hanover, NJ, USA; 2Knoll Consulting Inc. for Novartis Pharmaceuticals, Frumark Park, NJ, USA
OBJECTIVES: Propensity score matching (PSM) and Inverse Probability of Treatment Weighting (IPTW) are analytical methods used in comparative effective research (CER) to estimate treatment effect. However, these methods may not be applicable in studies with small sample sizes and unbalanced comparison groups. In such cases, where there is an unbalanced design due to unmeasured factors, difference-in-difference (DD) can be applied to estimate treatment effects. This research aimed to apply DD to a small sample of multiple sclerosis (MS) patients with unbalanced comparison groups. METHODS: A retrospective study was conducted using MarketScan Commercial Claims and Encounters and Medicare Supplemental and Coordination of Benefits Databases to compare MS patients who switched from glatiramer acetate (GA) to fingolimod (FTY) with another group another maintaining continuous use of GA. IPTW and DD were applied to balance the comparison groups; PSM could not be implemented due to imbalance between groups. Using DD, the treatment effect (i.e. % of patients with relapse in the GA vs. FTY group) minus (difference in relapses in the GA--F TY group) minus (difference in relapses in the GA-only group). RESULTS: IPTW was first employed to balance the two groups (GA--F TY group). Despite implementation of IPTW, the comparison groups could not be balanced on multiple factors such as patient demographics and clinical characteristics; therefore, DD was utilized to evaluate treatment effect on MS relapse rate. CONCLUSIONS: DD is an effective methodological when allows for estimating treatment effects from populations with unbalanced comparison groups.

PM123 LOW-MOLECULAR-WEIGHT HEPARIN TREATMENT OF DEEP-VEIN THROMBOSIS: A NETWORK META-ANALYSIS
Diaz JP1, Soto Molina H2, Marques M1, Escobar Juárez Y1
1Universidad Nacional Autónoma de México, México D.F., Mexico, 2HS Estudios FarmacovíaAmicos, Mexico City, Mexico. 3Universidad Autónoma Metropolitana, México D.F., Mexico
OBJECTIVES: It is estimated that up to 400,000 persons in Mexico are hospitalized yearly for deep-vein thrombosis (DVT). DVT is the presence of a blood clot (thrombus) in the deep veins of the body. The main objective was to indirectly compare efizcaparin and tinzaparin in efficacy and safety for the treatment of deep-vein thrombosis using a network meta-analysis (NMA). METHODS: To identify suitable studies for a systematic review of treat- ment options, 47 studies were included. A search was conducted in MEDLINE, Pubmed, EMBASE, CENTRAL (all via the Cochrane Library), Emboid, HTA, for relevant studies recorded between 1994 to April 2014. Only randomized controlled trials assessed in long- and short-term studies were included. We had to report the proportion of patients having recurrence of DVT (efficacy) and the proportion of patients having major bleeding (safety). Titles and abstracts were screened, data were extracted and risk of bias assessment was undertaken. Bayesian NMA was used to compare the different interventions. RESULTS: Four studies, assessing four low molecular weight heparins (LMWH), were judged to be sufficiently comparable for inclusion in the NMA. For the proportion of patients having recurrence of DVT or major bleeding, enoxaparin 1 mg/kg twice daily, tinzaparin 175 IU/kg once daily and nadroparin 100 IU/kg twice daily had a higher probability of being more effective and safer than unfractioned heparin. None of the LMWHs demonstrated a significant superiority over each other in terms of efficacy and safety; therefore, the group of LMWHs is suitable for a further cost minimization analysis and reference price implementation. CONCLUSIONS: We found no evidence of differences between tinzaparin, nadroparin and enoxaparin for recurrence of DVT and major bleeding. Tinzaparin may be preferred by clini- cians because it is always given once daily.
by Signorovitch (WS) vs. Entropy Balancing (EB). WS is based on propensity score estimation in AGR trials while EB relies on a maximization scheme. **RESULTS:** Simulation show the optimal weighting method is to match on covariates against the AGR treatment and control arms separately. In addition, re-balancing on prognostic variables between the IPD arms using EB is better than AGR, which is not reported. For example, assessment mean difference between AGR and IPD (GC) of one, six predictive variables in AGR and IPD and three prognostic variables in IPD, the Bucher method gives a biased estimate of 0.34 (95% confidence interval: 0.5, 1.08). WS balancing gives 0.99 (0.32 – 1.73), balancing on each arm separately gives 0.99 (0.30-1.68) while re-balancing using EB gives 0.99 (0.47-1.52). Also, simulations demonstrate that including placebo response into the weighting in addition to baseline covariates can provide biased results and is not recommended. **CONCLUSIONS:** MAIC can be improved if weighting is performed on each arm separately together with re-balancing of the IPD on the prognostic variables not reported.

**PM125**

**PSYCHOTROPIC PHARMACOTHERAPY ASSOCIATED WITH QT PROLONGATION AMONG VETERANS WITH POSTTRAUMATIC STRESS DISORDER**

**Stock EM**, Zeber JE, McNeill CJ, Banchs JE, Copeland LA

1Center for Applied Health Research - Central Texas Veterans Health Care System jointly with Baylor Scott & White Health, Temple, TX, USA, 2Baylor Scott & White Health, Temple, TX, USA

**OBJECTIVES:** In 2012, the FDA issued Drug Safety Communications on several drugs associated with QT prolongation and fatal ventricular arrhythmias. Among these was tipisetron, a selective serotonin reuptake inhibitor (SSRI) commonly used to treat post-traumatic stress disorder (PTSD). We examined whether SSRI use, known to prolong QTc, is associated with drug-related QT prolongation in patients with severe mental illnesses. This study explores psychotropic drugs associated with QT prolongation among Veterans with PTSD using the COAPT database. **RESULTS:** In total, we identified 176 Veterans diagnosed with QT prolongation. Cases were matched 1:1 on age, gender, visit date and setting, respectively. Occurrence in PTSD was assessed drug-related QT prolongation in patients with severe mental illnesses. **CONCLUSIONS:** This study explored psychotropic drugs associated with QT prolongation among Veterans with PTSD using the COAPT database. MAIC can be improved if weighting is performed on each arm separately together with re-balancing of the IPD on the prognostic variables not reported.

**PM127**

**NUMERICAL ESTIMATION OF STATISTICAL POWER AND MINIMUM NUMBER OF STUDIES NEEDED FOR A NETWORK META-ANALYSIS**


1GS, Phoenicia, PA, USA, 2ClasinoSmithKline, King of Prussia, PA, USA, 3GSK, London, UK, 4GSK, Phoenix, AZ, USA, 5GSK, Waltham, MA, USA

**OBJECTIVES:** Network meta-analysis (NMA) is increasingly used to inform reimbursement decisions and comparative effectiveness. Nevertheless, there is limited understanding of the assessment of statistical power, particularly when there is a small number of studies available. We propose an empirical calculation for statistical power using a simulation approach. **METHODS:** Simulation data were generated in minimum of 3 to maximum of 15 studies per network, using varying effect sizes and standard deviations under the assumptions of normality and exchangeability. The number of studies was varied using 3 target-specific treatment effects (i,j) came from a common distribution with mean (di) and variance (x2). The common distribution is usually chosen to be a normal distribution, so that the results of simulations to the gold standard, defined by available data from each study and Thorlund’s method. **RESULTS:** The proposed method was successful in estimating the statistical power and the number of studies needed, using simulation data from an NMA, compared to the gold standard. The estimations for statistical power/study of studies needed by different effect sizes, standard errors, and noise levels were compared with Thorlund’s method showing more accurate to estimate statistical power for NMA. Our method offers flexibility and can be implemented with Normal, Binomial, and Poisson distributions. Further, it can be applied to random effect models. **CONCLUSIONS:** While it is difficult to derive a mathematical formula for estimating statistical power and the number of studies needed in NMA because of its random nature, we propose a simulation-based empirical method using simulations allows for estimation of these quantities. The proposed method will be useful for researchers designing NMA to inform decision makers.

**PM128**

**ECONOMICS OF DIABETES MELLITUS: THEORY AND EVIDENCE FOR BRAZILIAN DATA IN 2008**

**Balcão I Jr, West R, Jacinto P**

Universidade Federal do Rio Grande da Sul, Porto Alegre, Brazil, 1pharcos, PORTO ALEGRE, Brazil

**OBJECTIVES:** to measure the DM social cost based in earnings losses of Brazilian workers due to data in 2008 using data from National Survey of Households (PNAD/IBGE). Diabetes Mellitus (DM) is characterized by the high level of blood glucose. Ministry of Health data estimated that Brazil had about 10 million DM cases in 2010, being the fourth main cause of death. WHO estimated the prevalence of DM in Brazil is about 5.2 million adults. **METHODS:** To evaluate the impact of DM on productivity, we used two approaches and the potential model to measure the participation in work force and a two-stage Heckman model to measure worked hours and productivity. Each model is estimated separately for both gender individuals, with and without disease, according three distinct definitions for DM: Restrict, Broad and Comorbidities. To capture the counterfactual effect, the model was calculated for ill and healthy individuals. The difference of both values exhibited the losses, which were aggregate to the whole population and the total cost was calculated. **RESULTS:** According each criterion, respectively, DM reduced the participation in the labor market in 0.97%, 4.60% and 7.06% for men and 0.14%, 4.79% and 6.44% for women, while reduced, respectively 1.51%, 6.40% and 9.15% in productivity. **CONCLUSIONS:** DM generates significant losses in income of Brazilian workers, especially in relation to their participation in the labor market, since affects both of gender. The results indicate that public policies should be directed to disease diagnosis and prevention, since the development of comorbidities amplifies the effect of losses.

**PM129**

**CLUSTER ANALYSIS OF HEALTHCARE COSTS PATTERNS IN END STAGE RENAL FAILURE PATIENTS WHO INITIATED HEMODIALYSIS**

Liao M1, L1Y, Kinafand F2, Ohi B1, Arcona S1

1KMR Consulting Inc., Florham Park, NJ, USA, 2Novartis Pharmaceuticals, East Hanover, NJ, USA, 3Rutgers University, Piscataway, NJ, USA

**OBJECTIVES:** Cluster analysis (CA) is a widely used statistical technique that helps reveal classifications of entities with similar characteristics in large data sets. However, little is known about whether it can be applied to healthcare data claims with highly skewed cost information. This study applied different clustering methods to changes in all-cause cost data from a group of patients with end stage renal disease (ESRD) who initiated hemodialysis (HD). **METHODS:** A retrospective, cross-sectional, observational study was conducted using the MarketScan Commercial Claims database. Patients aged ≥18 years with ≥2 ESRD diagnoses who initiated HD between 2008 and 2010 were included. The k-means CA method and hierarchical CA with a complete-linkage rule were applied to all-cause cost data (12-month pre-HD and follow-up periods (12-month post-HD) to identify clusters. Demographic, clinical, and cost information were extracted from both periods, and a total of 5,445 patients were enrolled. **CONCLUSIONS:** Meaningful all-cause cost clusters were generated using K-means and hierarchical CA with either flexible beta or Ward’s methods. Based on cluster sample sizes and changes of cost patterns, the K-means CA method and 4 clusters were selected. Cluster sizes at average costs range from $13,146 to $40,624, having very high costs (n=13), high and increasing costs (n=4,155), or very high costs reduced to high cost (n=89). Relatively stable costs after starting HD were associated with higher stable scores on comorbidity index from the pre- and post-HD periods, while increasing costs were associated with more sharply increasing comorbidity.