containing at least one product and one adverse event keyword were collected, de-identified, and standardized using a vernacular to MedDRA dictionary. Posts were classified as resembling an adverse event report (Proto-AEs) or simply discussing a product (Mention). **RESULTS:** There were a total of 1,410,819 posts categorized as Proto-AEs, 265,838 (19%) from Facebook and 1,144,981 (81%) from Twitter. The top 10 products accounted for 940,666 (67%) of the total Proto-AEs in Facebook and Twitter combined. The top 25 accounted for 1,180,040 (84%), the top 50 for 1,285,836 (91%), and the top 100 for 1,245,010 (95%) of the total Proto-AEs. The top 10 products (diphenhydramine, flu vaccine, dextroamphetamine, codeine, morphine, ibuprofen, alprazolam, acetaminophen, oxycodone, and zolpidem) were comprised of six controlled substances, three over-the-counter (OTC) products, and one class of vaccine. Of the top 50 products, controlled substances accounted for 32%, OTC products for 24%, and vaccines for 10%. **CONCLUSIONS:** Review of publically available data over the past two years from two popular social media sites, Facebook and Twitter, offers a high number of potential adverse events (Proto-AEs) for further evaluation. Social listening may be potentially valuable as a supplement to traditional pharmacovigilance practices, particularly for controlled substances, over-the-counter products, and vaccines. These initial findings warrant more research and a closer inspection as to the nature of these posts.

PRM55

HUNTING FOR RANDOMISED CONTROLLED TRIALS (RCTS): A COMPARISON OF SEARCH FILTERS DESIGNED TO IDENTIFY RCTS

Steeves S, Buchanan-Hughes AM, Evans J, Leonard S

Costello Medical Consulting Ltd, Cambridge, UK

OBJECTIVES: There are a number of search filters designed to identify studies with particular study designs in electronic databases. This study compared three filters for identifying randomised controlled trials (RCTs). **METHODS:** Searches were conducted on 15thJune 2015 in the Ovid MEDLINE and MEDLINE In-process databases using The Cochrane Highly Sensitive Search Strategy for Identifying Randomized Trials in MEDLINE, the SIGN Randomised Controlled Trials MEDLINE filter and the BMJ MEDLINE Randomised Controlled Trial Strategy. Differences were explored by reviewing samples of records uniquely identified by each filter. For comparison, a sample of articles returned by all three filters was also reviewed. To estimate the sensitivity of each filter, the detection of 39 publications of RCTs included in a randomly-selected Cochrane Collaboration systematic literature review (SLR) was tested. **RESULTS:** 476,551 records were identified by all three filters. From a sample of 384 records, 230 were RCTs and 18 were SLRs, meta-analyses or pooled analyses of RCTs. 1,000,716 records were uniquely identified by the Cochrane filter; of 400 records sampled , 0 were RCTs and 3 were SLRs or meta-analyses of RCTs. 500,127 articles were uniquely identified by the SIGN filter; of 386 records sampled, 8 were RCTs and 1 was a meta-analysis of RCTs. 84,938 records were uniquely identified by the BMJ filter; of 400 records sampled, 6 were RCTs and 5 were SLRs or meta-analyses of RCTs. 39/39, 38/39 and 37/39 of the Cochrane review publications were identified by the Cochrane, SIGN and BMJ filters, respectively. The publication missed by the SIGN filter was not the same as the 2 missed by the BMJ filter. **CONCLUSIONS:** All filters failed to identify at least some RCTs, SLRs or meta-analyses of RCTs. Differences between filters, including the publications uniquely identified by each, should be considered when selecting filters for use in literature reviews.

PRM56

INDIRECT TREATMENT COMPARISON (ITC) TO DEMONSTRATE THE UTILITY OF ONCOLITBANK: AN ONCOLOGY LITERATURE REGISTRY

Gala S, Shah A, Nanavaty M, Mwamburi M

Market Access Solutions LLC, Raritan, NJ, USA

OBJECTIVES: OncoLitBank is a registry of published oncology trials and Health Technology Assessments aimed to help provide a platform for secondary data analytics in the field of oncology. To demonstrate the utility and functionality of OncoLitBank, we conducted a basic indirect treatment comparison (ITC) between gemcitabine/nab-paclitaxel (GemNpac) and gemcitabine/capecitabine(GemCap) combinations which were each tested against gemcitabine but not against each other for treatment of metastatic pancreatic cancer. **METHODS:** Using OncoLitBank, data for metastatic pancreatic cancer were filtered for treatment-comparator arms of interest by using built-in interactive features. Two phase III randomized controlled trials (RCTs) comparing GemCap vs Gem, while 1 RCT comparing GemNpac vs Gem were included for the ITC. Data on overall response rate (ORR), 1-year survival, overall survival (OS), progression-free survival (PFS), and grade 3-4 adverse events (AEs) were pooled for the 2 RCTs comparing GemCap vs Gem using RevMan 5.0. Pooled risk ratios (RRs) and mean differences (MD) for the 2 studies were derived for dichotomous and continuous outcome variables respectively and compared to the single RCT that evaluated GemNpac vs Gem to derive the ITC RRs and MDs using the Canadian Agency for Drugs and Technologies in Health (CADTH) ITC application which employs Bucher et al. (1997) method. **RESULTS:** The use of OncoLitBank was successful and eliminated the need to conduct a new systematic review to perform the ITC leading to a quick turn-around of tasks at hand. Results demonstrated that GemNpac was not superior to GemCap in ORR, 1-year survival, OS, PFS and grade 3-4 AEs, as no significant differences was detected. **CONCLUSIONS:** OncoLitBank provides users with a robust data platform that can be easily used for systematic reviews, conduct meta-analyses through direct orindirect comparisons, inform economic models, landscape analyses, value dossiers, create target product profiles and value development plans.

PRM57

UTILITY AND METRICS OF NATURAL LANGUAGE PROCESSING ON IDENTIFYING PATIENTS FOR PHARMACOEPIDEMIOLOGIC STUDIES

Bailey N[1], Wilson A[1], Kamauu A[1], DuVall S[2]

[1]*Anolinx, Murray, UT, USA,* [2]*VA Salt Lake City Health Care System, Salt Lake City, UT, USA*

OBJECTIVES: Electronic medical records (EMR) are increasingly utilized in clinical practice and research, allowing for more efficient availability of rich patient records. However, most use of EMR is limited to coded, structured, and administrative data, while the vast majority of patient information (e.g. disease subtype, severity, medical device usage, etc.) is tied up in narrative clinical notes. The challenge remains in accessing the information in these patient notes. Historically this has been done via timely and costly manual chart review, but as the amount of EHR data increases exponentially, manual chart review becomes impractical and impossible. Advancements in Natural Language Processing (NLP) have demonstrated promising results in combining the capture of additional clinical note information with the efficiency of modern informatics. The objective of this study is to demonstrate the relevancy and utility of NLP to extract health data from EMR in real-world observational studies. **METHODS:** We conducted a systematic review and meta analysis of performance metrics for five (5) NLP-driven projects involving oncology, inflammation and medical devices ,which had similar protocols and objectives. We assessed and validated the accuracy of NLP algorithms, as well as heterogeneity of accuracy between studies using random effects meta-analysis (represented by I2 value). **RESULTS:** A total of 382,523 patients were identified using NLP among the 5 studies. Accuracy among the studies ranged from 95.2% to 100% (95% CI: 95.1%, 100%), with an I2 value of 95.9% (95% CI: 92.9%, 97.7%). **CONCLUSIONS:** NLP provide a unique opportunity to extract meaningful information from patient-level narrative clinical notes in EMR data sources with high degree of accuracy. This provides additional rich sources of data from narrative clinical notes, that are otherwise not easily available, to support epidemiology and other real-world observational studies.

PRM58

METHOTDOLOGICAL DIFFICULTIES OF COMPLIANCE ANALYSES BASED ON REAL-WORLD DATA

Andriska P, Komáromi T, Frigyesy R, Salfer B

Healthware Consulting Ltd., Budapest, Hungary

OBJECTIVES: Regarding compliance analysis numerous ratios can be found in international scientific literature with simpler or more complex methodology. In our analysis we tend to reveal, that choosing an adequate ratio is not sufficient itself, it is essential to know the difficulties and pitfalls of the data management and methodology to the objective assessment of the chosen ratio. The chief aim of our study to demonstrate factors in course of practical examples, which may substantially influence the results and the right conclusions, if these factors are modified. **METHODS:** The analysis is based on prescription refilling's data of database of the Hungarian Health Fund in the field of the following indications: diabetes, COPD, oncology. From the ratios available in scientific literature, the PDC (Proportion of Days Covered) was chosen. The following aspects were considered as influencing factors: patient inclusion criteria (index date, time frame, criteria of refillings); DDD (WHO, SPC or real-world dosage to DOT); Gap (period without medication supply). A basic setting was established to calculate PDC ratio, then after changing each above specified parameters one by one (ceteris paribus), the ratio was recalculated. **RESULTS:** The PDC ratio shows huge variability recalculated by the different values of each parameters. Even more than 20% difference can be observed after modifying the gap (strict 1-day or permissive 30-day), or applying the SPC dosage instead of WHO DDD. In course of modifying the patient inclusion criteria both patient numbers and the ratio also show significant differences. **CONCLUSIONS:** Based on the results it may be concluded, that no general best practice can be observed, all settings have both advantages and limitations. It may be worth choosing the key parameters considering the specialties of each indications in order to draw conclusions as correct as possible with the focus of the original aim of the study.

PRM59

REVIEW OF COMORBIDITY MEASURES TO PREDICT ECONOMIC OUTCOMES IN REAL-LIFE DATABASE STUDIES

Guelfucci F[1], Bessou A[2], Aballea S[1]

[1]*Creativ-Ceutical, Paris, France,* [2]*National School For Statistics and Information Analysis (ENSAI), BRUZ, France*

OBJECTIVES: Generic comorbidity measures developed to predict mortality and/or healthcare costs are often used as adjustment covariates in observational studies comparing health expenditures between different therapeutic strategies. The objective of this review is to identify available measures, and assess their performance for prediction of economic outcomes using large longitudinal patient databases. **METHODS:** We conducted a comprehensive literature search in MEDLINE, until April 2015. All methodological papers describing a new comorbidity measure or assessing their ability to predict economic outcomes using administrative data or electronic medical records were selected. We search for additional studies through references lists of selected articles. We extracted information on the conditions for using each index and predictive performance. **RESULTS:** 323 abstracts were identified during the search in MEDLINE and 25 full papers were reviewed. Eleven comorbidity measures and seven comparative studies were found. Four comorbidity measures were single cumulative weighted scores: two were diagnosis-based indexes developed using large administrative health databases and two others were medication-based indexes developed using pharmacy data. Two comorbidity measurement systems consisted in classifying patients in mutually exclusive groups defined based on diagnosis and clinical or economical characteristics. Others were simple counts of diseases. All measures were based on a list of diseases pre-selected by clinicians, except for the Ambulatory Clinical Groups System (ACG). Five measures were adapted for use with ICD-9-CM and ICD-10 classifications. Hierarchical Cost Groups (HCC-CMS) and Quality and Outcome Framework (QOF) showed the highest predictive ability in three comparative studies. ACG was the best predictor in one study and the second one in three other analyses. **CONCLUSIONS:** HCC-CMS and QOF were reported to have the best predictive performance. However most comparative studies included a limited number of comorbidity measures.