of healthcare, delivering enhanced healthcare value and achieving better solutions for particular situations.

PNS261
HOW DOES MACHINE-LEARNING COMPARE TO AN INCOMING MEDICAL STUDENT IN EXTRACTING OUTCOMES RESULTS FROM ABSTRACTS?

Michelson M,[1] Ross M,[1] Minton S[2]
[1]Evid Science, El Segundo, CA, USA, [2]Inferlink, El Segundo, CA, USA

Objectives: Literature analysis could benefit from machine-learning (ML) methods that parse medical text to extract reported results. Therefore, we benchmark one ML approach against an incoming medical school student in the same task, as a comparative study. ***Methods:*** We ran an ML algorithm, using Named Entity Recognizers, against a corpus of PubMed abstracts from Randomized Controlled Trials. We then sampled 102 resulting sentences, limiting to those with at least two interventions compared to one another. The ML parsed the disease, end-points, interventions studied, and the reported numerator and denominator (e.g., the sentence, "The total response rate was 75.0% (6/8) in non-transplantation group and 37.5% (3/8) in transplantation group, respectively." yields two results. The first has 6 as a numerator, 8 as a denominator, "total response rate" as the end-point, and "non-transplantation" as the intervention. The second result has 6, 8, "total response rate", and "transplantation" for its attributes.) We then trained the student on the same task and gave him the 102 sentences to process. ***Results:*** The student and the ML algorithm yielded the same end-points for 97/102 sentences (95%). We then compared triples of (numerator, denominator, intervention) across the sentences. The ML and student match on 87/102 (85%) of sentences (e.g., both pairs matched for each sentence 85% of the time between the ML and student). Accounting for errors associated with only the numerator and denominator (e.g., the interventions match), this number improves to 94/102 sentences (92%). These errors are mostly associated with incorrectly assigning one intervention's quantities to the other. ***Conclusions:*** Although preliminary, we demonstrate that ML can be almost as effective as an incoming medical student in the task of turning written text into structured results.

PNS262
GLOBAL TRENDS IN REGULATORY LEGISLATION FOR POST MARKETING SURVEILLANCE (PMS) REAL-WORLD EVIDENCE (RWE) STUDIES AND THE SECONDARY USE OF HEALTH DATA IN PMS STUDIES.

Smulders M,[1] Forstner M,[2] Bunoza A,[3] Valverde D[4]
[1]PrimeVigilance, Waltham, MA, USA, [2]PrimeVigilance, Guildford, United Kingdom, [3]PrimeVigilance, Zagreb, Croatia, [4]Techtrials, Vinhedo, SP, Brazil

Objectives: Describe regulatory legislation for PMS RWE studies, and the secondary use of health data including administrative, electronic health record (EHR) and claims databases in PMS studies in Europe and seven AsiaPacific and Latin American countries. ***Methods:*** We reviewed local country regulatory websites and the European Network of Centers for Pharmacoepidemiology and Pharmacovigilance (ENCePP) website. We also searched for relevant publications in PubMed from January 2015 to December 2018. ***Results:*** In Europe, 18/28 countries have specific legislation for the secondary use of health data. The number of data sources available rose from 83 in 2017 to 134 in 2019 (61% increase), supporting about 140 PASS in 2017 and 193 in 2019. The majority of regulatory decision making is based on only 34 databases in 13 EU countries. Similar to many EU countries, Brazil, India, Japan, mainland China, Mexico, South Korea, and Taiwan, request PMS RWE. In Asia and Mexico, the trend is to request primary data collection of exposed patient cohorts with prospective follow-up, also in Japan, South Korea, and Taiwan, where comprehensive national databases are available. In Brazil, there has been an increase in the use of structured databases to generate RWE for PMS in 2018. ***Conclusions:*** Regulators have been and will likely increasingly accept or request secondary use of health data in PMS studies to observe drug effects in routine practice conditions and larger and more diverse populations, provided the studies use fit for purpose databases, assessed by experts in the field who use correct methodologies and validation where needed. Existing comprehensive and representative EHR/claims databases available in Brazil, Japan, South Korea, and Taiwan, are expected to be used more in PMS studies. We eagerly anticipate regulatory secondary use of health data guidance documents in Asia, several EU countries and Latin America.

PNS263
FREQUENCIES OF CONSECUTIVELY MISSED DAILY DOSES OF ESCITALOPRAM

Kaufman AS,[1] Morrison A,[1] Stauffer ME[2]
[1]ScribCo, Effort, PA, USA, [2]Scientific Editing Solutions, Walworth, WI, USA

Objectives: The therapeutic effect of a drug will be lost when there are occurrences of consecutively missed doses that exceed the duration of the drug's effect. We previously reported that runs of ≥3 and ≥4 consecutively missed doses of oral antiretroviral medications occurred more frequently than predicted by random chance. We hypothesize that this non-random pattern of consecutively missed doses might be a general phenomenon, regardless of the therapeutic area. In the case of escitalopram, a selective serotonin reuptake inhibitor indicated for treatment of anxiety and depression, therapeutic serum levels are expected to be maintained if there are no occurrences of ≥3 consecutively missed daily doses. Here, we determine the frequencies of ≥2, ≥3, and ≥4 consecutively missed doses of oral escitalopram 20 mg once daily, using published Medication Event Monitoring System data (*Int Clin Psychopharmacol* [2012] 27;291-297). ***Methods:*** We recorded occurrences of ≥2, ≥3, or ≥4 consecutively missed doses in 75 patients with ≥30 days of data, for a total of 398 30-day periods. The observed distributions of the frequencies of consecutively missed doses as a proportion of the number of missed days of dosing per 30-day period were compared to theoretical random distributions of misses using the Wilcoxon signed-rank test. ***Results:*** The overall frequencies of occurrences of ≥2, ≥3, and ≥4 consecutively missed doses were 0.44, 0.28, and 0.20, respectively. The probabilities that the observed distributions of misses were as expected from random chance were P=0.51, P<0.01, and P<0.01, respectively. ***Conclusions:*** Observed runs of ≥3 and ≥4 consecutively missed doses—which are expected to result in serum levels of escitalopram falling below the therapeutic range—occurred more frequently than expected if missed doses were randomly distributed. The hypothesis, that this non-random pattern of consecutively missed doses is a general phenomenon across therapeutic areas, was not disproven.

PNS264
AUTOMATED GENERATION OF EVIDENCE-GAP MAPS FROM MEDICAL LITERATURE

Ross M,[1] Michelson M,[1] Tee Qiao Ying A,[1] Ashish N,[2] Minton S[2]
[1]Evid Science, El Segundo, CA, USA, [2]Inferlink, El Segundo, CA, USA

Objectives: Literature analysis could benefit from machine-learning (ML) methods that parse medical text to extract reported results. We demonstrate this by automating the creation of comparative effectiveness Evidence-Gap Maps. ***Methods:*** We ran an ML algorithm, using Text Classifiers and Named Entity Recognizers, against a corpus of PubMed abstracts. The algorithm identified the study type (e.g., Randomized Control Trial), and parsed diseases and conditions, interventions, group-sizes and fractional outcomes from the abstract text (e.g., the phrase, "RA clinical remission was 6 of 8 for infliximab" yields {condition: rheumatoid arthritis; intervention: infliximab; outcome: clinical remission; outcome-count: 6; group-size: 8}). Text-spans identifying conditions, interventions, and outcomes were normalized across documents using clustering and ontology alignment. To construct an Evidence-Gap Map for a given disease or condition, the N most common interventions and outcomes are determined from the extracted data (N=10). The counts of abstracts pairing each intervention and outcome are then visualized as a bubble map using standard plotting software. To validate, we generated maps for a random sample of conditions with >=100 PubMed abstracts. Spurious normalizations (e.g., "primary outcome") were manually identified and removed. Condition-intervention-outcome triplets were then manually converted into PubMed search queries (e.g., "rheumatoid arthritis" infliximab "clinical remission"), and the results were tallied for gap and non-gap triplets. ***Results:*** The initial sample showed an average of 0.34 and 4.68 papers for gap and non-gap triplets. Removal of spurious normalizations brought the averages to 0 and 4.7. Search results for those triplets returned an average of 3.2 and 25 results. ***Conclusions:*** We demonstrated that machine learning methods can be used to create automated Evidence-Gap Maps which could potentially help identify comparative effectiveness evidence gaps in medical literature.

PNS265
AUTOMATED DISCOVERY OF COMPARATIVE EFFECTIVENESS HYPOTHESES FROM MEDICAL LITERATURE.

Ross M,[1] Michelson M,[1] Tee Qiao Ying A,[1] Ashish N,[2] Minton S[2]
[1]Evid Science, El Segundo, CA, USA, [2]Inferlink, El Segundo, CA, USA

Objectives: Literature analysis could benefit from machine-learning (ML) methods that parse medical text to extract reported results. We demonstrate this by creating an algorithm for automated discovery of Comparative Effectiveness Hypotheses for medical interventions. ***Methods:*** We ran an ML algorithm, using Text Classifiers and Named Entity Recognizers, against a corpus of PubMed abstracts. The algorithm identified the study type (e.g., Randomized Control Trial), and parsed diseases and conditions, interventions, group-sizes and fractional outcomes from the abstract text (e.g., the phrase, "RA clinical remission was 6 of 8 for infliximab" yields {condition: rheumatoid arthritis; intervention: infliximab; outcome: clinical remission; outcome-count: 6; group-size: 8}). Text-spans identifying conditions, interventions, and outcomes were normalized across documents using clustering and ontology alignment. The extracted data was used to construct a table containing Relative Risk (RR) values with 95% CI. For each intervention, paired conditions and outcomes are found, and the RR values are calculated from the fractional outcome results of the intervention and its competitors, aggregated across studies. Over-/Under-Perform Hypotheses are those condition/outcome pairs for which RR=1.0 falls outside the 95% CI. Statistics were calculated over a random sample of interventions with >= 100 published abstracts. ***Results:*** The average number of generated hypotheses for each intervention was 4.0, 3.6, and 10.2 for over-perform, under-perform, and no-significant-difference. Average counts of disease conditions for each intervention were 3.5, 3.2 and 7.5, respectively. ***Conclusions:*** We demonstrated that AI can be used to assist in discovery of comparative effectiveness hypotheses from medical literature. Generated hypotheses spanned a range of conditions and outcomes for each intervention.