Accuracy of an Automated Knowledgebase for Identifying Adverse Drug Reactions

Erica A. Voss, MPH[1,2], Richard D. Boyce, PhD[3,2], Patrick B. Ryan, PhD[1,2,4], Peter R. Rijnbeek, PhD[5,2], Martijn Schuemie, PhD[1,2]

[1]Janssen Research and Development, Raritan, NJ, [2]Observational Health Data Sciences and Informatics (OHDSI), New York, NY,
[3]University of Pittsburgh, Pittsburgh, PA, [4]Columbia University, New York, NY, [5]Erasmus University Medical Center, Rotterdam, Netherlands

OBJECTIVES

To build a machine learning classifier using **Large-scale Adverse Effects Related to Treatment Evidence Standardization (LAERTES)** to identify:

- ❑ **Positive Controls:** drugs that cause certain outcomes.
- ❑ **Negative Controls:** drugs that lack evidence for causing certain outcomes.

A classifier addresses the immediate need of **automating the selection of positive and negative controls**.

Hypothesis: The information found in LAERTES can be used to classify drugs and health outcomes of interest (HOI) into positive or negative controls.

BACKGROUND

Safety assessment is time-consuming and resource-intensive, as evidence is often compiled from varying, often unstructured, and generally incomplete, systems which yield conflicting information that must be interpreted and reconciled with appropriate clinical and analytical expertise.

Drug safety evidence for adverse drug reaction (ADR) is produced across many disparate sources.

LAERTES provides a standardized structure to enable analyses across relevant ADR evidence.

MATERIALS & METHODS

Each piece of evidence listed in Table 1 was identified by a drug (RxNorm) and HOI (SNOMED) pair.

Table 1: Evidence that LAERTES Contains

Data Source Type	Data Source & Description
Literature	Medline MeSH Clinical Trials (1)
	Medline MeSH Case Reports (1)
	Medline MeSH Other (1)
	Medline SemMedDB Clinical Trials (2)
	Medline SemMedDB Case Reports (2)
	Medline SemMedDB Other (2)
Product Labels	European Product Label Adverse Drug Reactions (3)
	Structured Product Label Adverse Drug Reactions from SPLICER (4)
Spontaneous Reports	FAERS Report Count (5)
	FDA Adverse Event Reporting System (FAERS) Proportional Reporting Ratio (PRR) (5)

A regularized logistic regression model was fitted on manually created reference sets and predictive accuracy was estimated using cross-validation.

- ❑ Observational Medical Outcomes Partnership (OMOP) Reference Set (6).
- ❑ Exploring and Understanding Adverse Drug Reactions (EU-ADR) Reference Set (7).

To test the generalizability a 3rd reference set was used, the Arizona Center for Education and Research on Therapeutics (AZCERT) dataset (8).

Each reference set defined the ingredients/condition(s) of interest, however only concepts with enough evidence in LAERTES were considered (i.e. had at least one piece of evidence in each data source type grouping [Table 1]).

RESULTS

Table 2 provides the measured performance of models based on individual evidence types as well as the full model with all evidence types.

Table 2: AUC and 95% confidence interval for individual predictors and a regularized regression model using all predictors, using leave-pair-out cross-validation.

Column(s) in Model	AUC on OMOP Set	AUC on EU-ADR set
Medline Clinical Trial	0.57 (0.54-0.59)	0.64 (0.56-0.71)
Medline Case Reports	0.68 (0.63-0.72)	0.89 (0.82-0.95)
Medline Other	0.52 (0.50-0.54)	0.55 (0.50-0.60)
Medline SemMedDB Clinical Trial	0.58 (0.55-0.61)	0.57 (0.51-0.62)
Medline SemMedDB Case Reports	0.58 (0.55-0.61)	0.58 (0.51-0.65)
EU Product Labels	0.57 (0.54-0.60)	0.53 (0.49-0.56)
US Product Labels	0.81 (0.77-0.85)	0.77 (0.68-0.87)
FAERS *	0.75 (0.69-0.81)	0.67 (0.55-0.80)
FAERS PRR **	0.56 (0.50-0.63)	0.60 (0.47-0.73)
All Predictors	***0.87 (0.82-0.91)***	***0.93 (0.88-0.99)***

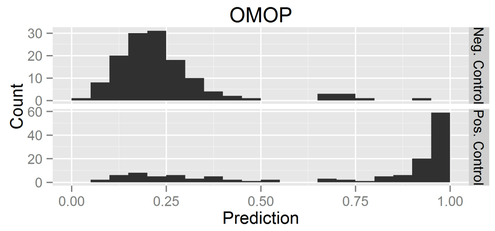
OMOP: Observational Medical Outcomes Partnership, EU-ADR: Exploring and Understanding Adverse Drug Reactions, AUC: area under the curve, LBCI: lower bound 95% confidence interval, UPCI: upper bound 95% confidence interval, FAERS: FDA Adverse Event Reporting System, PRR: proportional reporting ratio

* natural logs were taken to scale predictor ** geometric mean was used to scale predictor

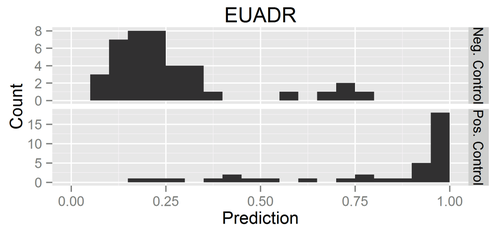
Figure 1 shows histograms of the predicted probability as well as the AUCs of the model for the positive and negative controls in all three reference sets.

The plots suggest that the predicted probabilities produced by the algorithm were generally useful for segregating positive and negative controls.

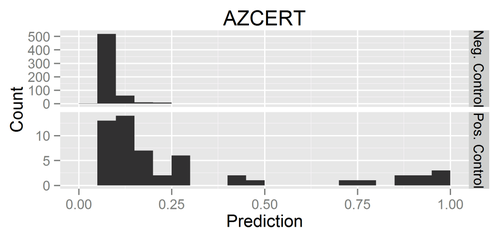
Figure 1. Histograms of predicted probabilities with AUCs for positive/negative controls in the various reference sets, using the model trained on both OMOP and EU-ADR ref. set.



AUC = 0.87 (0.83-0.92)



AUC = 0.95 (0.90-0.99)



AUC = 0.86 (0.79-0.93)

5% Worst Case Misclassification: AUC = 0.55

5% Best Case Misclassification: AUC = 0.92

OMOP: Observational Medical Outcomes Partnership, EU-ADR: Exploring and Understanding Adverse Drug Reactions, AUC: area under the curve

CONCLUSIONS

We found that LAERTES was relatively predictive of the reference sets, particularly when all the predictors with enough data were utilized.

LAERTES may have potential in identifying drug outcome pairs that warrant further evaluation, however this would need to be studied in the future.

LAERTES holds promise of being a powerful tool in selecting positive and negative controls in a faster manner than traditional manual efforts would allow.

CONFLICT OF INTEREST STATEMENT

EAV, RDB, PBR, PRR, and MS have no potential conflicts of interest to declare. EAV, PBR, and MS are full time employees of Janssen R&D, a unit of Johnson and Johnson. The work on this study was part of their employment. They also hold pension rights from the company and own stock and stock options.

REFERENCES

1. Avillach, P., et al., *Design and validation of an automated method to detect known adverse drug reactions in MEDLINE: a contribution from the EU-ADR project.* J Am Med Inform Assoc, 2013. **20**(3): p. 446-52.
2. Kilicoglu, H., et al., *Constructing a semantic predication gold standard from the biomedical literature.* BMC Bioinformatics, 2011. **12**: p. 486.
3. *Pharmacoepidemiological Research on Outcomes of Therapeutics by a European Consortium (PROTECT): Adverse Drug Reactions Database*. [web page] 2015.05.07]; Available from: http://www.imi-protect.eu/adverseDrugReactions.shtml.
4. Duke, J., J. Friedlin, and X. Li, *Consistency in the safety labeling of bioequivalent medications.* Pharmacoepidemiol Drug Saf, 2013. **22**(3): p. 294-301.
5. Tatonetti, N.P., et al., *Data-driven prediction of drug effects and interactions.* Sci Transl Med, 2012. **4**(125): p. 125ra31.
6. Ryan PB, Schuemie MJ, Welebob E, Duke J, Valentine S, Hartzema AG. Defining a reference set to support methodological research in drug safety. Drug safety. 2013 Oct;36 Suppl 1:S33-47. PubMed PMID: 24166222. Epub 2013/11/06. eng.
7. Coloma PM, Avillach P, Salvo F, Schuemie MJ, Ferrajolo C, Pariente A, et al. A reference standard for evaluation of methods for drug safety signal detection using electronic healthcare record databases. Drug safety. 2013 Jan;36(1):13-23. PubMed PMID: 23315292. Epub 2013/01/15. eng.
8. Woosley RL, Romero K. Assessing cardiovascular drug safety for clinical decision-making. Nat Rev Cardiol. 2013 06//print;10(6):330-7.