



OHDSI

OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

Community Publications

Agarwal V, Podchiyska T, Banda JM, Goel V, Leung TI, Minty EP, . . . Shah NH. Learning statistical models of phenotypes using noisy labeled training data. *Journal of the American Medical Informatics Association : JAMIA*. Nov 2016;23(6):1166-1173.
10.1093/jamia/ocw028

Albers DJ, Elhadad N, Claassen J, Perotte R, Goldstein A, Hripcsak G. Estimating summary statistics for electronic health record laboratory data for use in high-throughput phenotyping algorithms. *Journal of biomedical informatics*. Feb 2018;78:87-101.
10.1016/j.jbi.2018.01.004

Ayvaz S, Horn J, Hassanzadeh O, Zhu Q, Stan J, Tatonetti NP, . . . Boyce RD. Toward a complete dataset of drug-drug interaction information from publicly available sources. *Journal of biomedical informatics*. Jun 2015;55:206-217.
10.1016/j.jbi.2015.04.006

Ayvaz S, Zhu Q, Hochheiser H, Brochhausen M, Horn J, Dumontier M, . . . Boyce RD. Drug-drug interaction data source survey and linking. *AMIA Joint Summits on Translational Science proceedings. AMIA Joint Summits on Translational Science*. 2014;2014:16.

Banda JM, Callahan A, Winnenburgh R, Strasberg HR, Cami A, Reis BY, . . . Shah NH. Feasibility of Prioritizing Drug-Drug-Event Associations Found in Electronic Health Records. *Drug safety*. Jan 2016;39(1):45-57.
10.1007/s40264-015-0352-2

Banda JM, Evans L, Vanguri RS, Tatonetti NP, Ryan PB, Shah NH. A curated and standardized adverse drug event resource to accelerate drug safety research. *Scientific data*. May 10 2016;3:160026.
10.1038/sdata.2016.26

Banda JM, Halpern Y, Sontag D, Shah NH. Electronic phenotyping with APHRODITE and the Observational Health Sciences and Informatics (OHDSI) data network. *AMIA Joint Summits on Translational Science proceedings. AMIA Joint Summits on Translational Science*. 2017;2017:48-57.

Bell C, Chakravarty A, Gruber S, Heckbert SR, Levenson M, Martin D, . . . Walker AM. Characteristics of study design and elements that may contribute to the success of electronic safety monitoring systems. *Pharmacoepidemiology and drug safety*. Nov 2014;23(11):1223-1225.
10.1002/pds.3712

Blaisure JC, Ceusters WM. Improving the 'Fitness for Purpose' of Common Data Models through Realism Based Ontology. *AMIA ... Annual Symposium proceedings. AMIA Symposium*. 2017;2017:440-447.

Boland MR, Karczewski KJ, Tatonetti NP. Ten Simple Rules to Enable Multi-site Collaborations through Data Sharing. *PLoS computational biology*. Jan 2017;13(1):e1005278.
10.1371/journal.pcbi.1005278

Boland MR, Parhi P, Li L, Miotto R, Carroll R, Iqbal U, . . . Tatonetti NP. Uncovering exposures responsible for birth season - disease effects: a global study. *Journal of the American Medical Informatics Association : JAMIA*. Sep 28 2017.
10.1093/jamia/ocx105

Boland MR, Polubriaginof F, Tatonetti NP. Development of A Machine Learning Algorithm to Classify Drugs Of Unknown Fetal Effect. *Scientific reports*. Oct 9 2017;7(1):12839.
10.1038/s41598-017-12943-x

Boland MR, Shahn Z, Madigan D, Hripcsak G, Tatonetti NP. Birth month affects lifetime disease risk: a phenome-wide method. *Journal of the American Medical Informatics Association : JAMIA*. Sep 2015;22(5):1042-1053.
10.1093/jamia/ocv046

Boland MR, Tatonetti NP. Are All Vaccines Created Equal? Using Electronic Health Records to Discover Vaccines Associated With Clinician-Coded Adverse Events. *AMIA Joint Summits on Translational Science proceedings. AMIA Joint Summits on Translational Science*. 2015;2015:196-200.

Boland MR, Tatonetti NP. In Search of 'Birth Month Genes': Using Existing Data Repositories to Locate Genes Underlying Birth Month-Disease Relationships. *AMIA Joint Summits on Translational Science proceedings. AMIA Joint Summits on Translational Science*. 2016;2016:189-198.

Boland MR, Tatonetti NP. Investigation of 7-dehydrocholesterol reductase pathway to elucidate off-target prenatal effects of pharmaceuticals: a systematic review. *The pharmacogenomics journal*. Oct 2016;16(5):411-429.
10.1038/tpj.2016.48

Boland MR, Tatonetti NP, Hripcsak G. Development and validation of a classification approach for extracting severity automatically from electronic health records. *Journal of biomedical semantics*. 2015;6:14.
10.1186/s13326-015-0010-8

Boyce RD, Handler SM, Karp JF, Perera S, Reynolds CF, 3rd. Preparing Nursing Home Data from Multiple Sites for Clinical Research - A Case Study Using Observational Health Data Sciences and Informatics. *EGEMS (Washington, DC)*. 2016;4(1):1252.
10.13063/2327-9214.1252

Boyce RD, Ryan PB, Noren GN, Schuemie MJ, Reich C, Duke J, Tatonetti NP, Trifirò G, Harpaz R, Overhage JM, Hartzema AG, Khayter M, Voss EA, Lambert CG, Huser V, Dumontier M. Bridging islands of information to establish an integrated knowledge base of drugs and health outcomes of interest. *Drug safety*. Aug 2014;37(8):557-567.
10.1007/s40264-014-0189-0

Butler A, Wei W, Yuan C, Kang T, Si Y, Weng C. The Data Gap in the EHR for Clinical Research Eligibility Screening. *AMIA Joint Summits on Translational Science proceedings. AMIA Joint Summits on Translational Science*. 2018;2017:320-329.

Callahan TJ, Bauck AE, Bertoch D, Brown J, Khare R, Ryan PB, . . . Kahn MG. A Comparison of Data Quality Assessment Checks in Six Data Sharing Networks. *EGEMS (Washington, DC)*. Jun 12 2017;5(1):8.
10.5334/egems.223

Ceusters W, Blaisure J. A Realism-Based View on Counts in OMOP's Common Data Model. *Studies in health technology and informatics*. 2017;237:55-62.

Chakrabarti S, Sen A, Huser V, Hrubby GW, Rusanov A, Albers DJ, Weng C. An Interoperable Similarity-based Cohort Identification Method Using the OMOP Common Data Model version 5.0. *Journal of healthcare informatics research*. Jun 2017;1(1):1-18.
10.1007/s41666-017-0005-6

Chiu PH, Hripcsak G. EHR-based phenotyping: Bulk learning and evaluation. *Journal of biomedical informatics*. Jun 2017;70:35-51.
10.1016/j.jbi.2017.04.009

de Bie S, Coloma PM, Ferrajolo C, Verhamme KM, Trifiro G, Schuemie MJ, . . . Sturkenboom MC. The role of electronic healthcare record databases in paediatric drug safety surveillance: a retrospective cohort study. *British journal of clinical pharmacology*. Aug 2015;80(2):304-314.
10.1111/bcp.12610

Defalco FJ, Ryan PB, Soledad Cepeda M. Applying standardized drug terminologies to observational healthcare databases: a case study on opioid exposure. *Health services & outcomes research methodology*. Mar 2013;13(1):58-67.
10.1007/s10742-012-0102-1

- Duke JD, Ryan PB, Suchard MA, Hripcsak G, Jin P, Reich C, . . . Schuemie MJ. Risk of angioedema associated with levetiracetam compared with phenytoin: Findings of the observational health data sciences and informatics research network. *Epilepsia*. Aug 2017;58(8):e101-e106.
10.1111/epi.13828
- FitzHenry F, Resnic FS, Robbins SL, Denton J, Nookala L, Meeker D, . . . Matheny ME. Creating a Common Data Model for Comparative Effectiveness with the Observational Medical Outcomes Partnership. *Applied clinical informatics*. 2015;6(3):536-547.
10.4338/aci-2014-12-cr-0121
- Garza M, Del Fiol G, Tenenbaum J, Walden A, Zozus MN. Evaluating common data models for use with a longitudinal community registry. *Journal of biomedical informatics*. Dec 2016;64:333-341.
10.1016/j.jbi.2016.10.016
- Gini R, Schuemie M, Brown J, Ryan P, Vacchi E, Coppola M, . . . Sturkenboom M. Data Extraction and Management in Networks of Observational Health Care Databases for Scientific Research: A Comparison of EU-ADR, OMOP, Mini-Sentinel and MATRICE Strategies. *EGEMS (Washington, DC)*. 2016;4(1):1189.
10.13063/2327-9214.1189
- Gini R, Schuemie MJ, Mazzaglia G, Lapi F, Francesconi P, Pasqua A, . . . Klazinga N. Automatic identification of type 2 diabetes, hypertension, ischaemic heart disease, heart failure and their levels of severity from Italian General Practitioners' electronic medical records: a validation study. *BMJ open*. Dec 9 2016;6(12):e012413.
10.1136/bmjopen-2016-012413
- Gini R, Schuemie MJ, Pasqua A, Carlini E, Profili F, Cricelli I, . . . Klazinga N. Monitoring compliance with standards of care for chronic diseases using healthcare administrative databases in Italy: Strengths and limitations. *PLoS one*. 2017;12(12):e0188377.
10.1371/journal.pone.0188377
- Harpaz R, Callahan A, Tamang S, Low Y, Odgers D, Finlayson S, . . . Shah NH. Text mining for adverse drug events: the promise, challenges, and state of the art. *Drug safety*. Oct 2014;37(10):777-790.
10.1007/s40264-014-0218-z
- Harpaz R, DuMochel W, Shah NH. Big Data and Adverse Drug Reaction Detection. *Clinical pharmacology and therapeutics*. Mar 2016;99(3):268-270.
10.1002/cpt.302
- Harpaz R, DuMouchel W, Schuemie M, Bodenreider O, Friedman C, Horvitz E, . . . Shah NH. Toward multimodal signal detection of adverse drug reactions. *Journal of biomedical informatics*. Dec 2017;76:41-49.
10.1016/j.jbi.2017.10.013
- Harpaz R, Odgers D, Gaskin G, DuMouchel W, Winnenburger R, Bodenreider O, . . . Shah NH. A time-indexed reference standard of adverse drug reactions. *Scientific data*. Nov 11 2014;1:140043.
10.1038/sdata.2014.43
- He Z, Ryan P, Hoxha J, Wang S, Carini S, Sim I, Weng C. Multivariate analysis of the population representativeness of related clinical studies. *Journal of biomedical informatics*. Apr 2016;60:66-76.
10.1016/j.jbi.2016.01.007
- Hripcsak G, Duke JD, Shah NH, Reich CG, Huser V, Schuemie MJ, . . . Ryan PB. Observational Health Data Sciences and Informatics (OHDSI): Opportunities for Observational Researchers. *Studies in health technology and informatics*. 2015;216:574-578.
- Hripcsak G, Ryan PB, Duke JD, Shah NH, Park RW, Huser V, . . . Madigan D. Characterizing treatment pathways at scale using the OHDSI network. *Proceedings of the National Academy of Sciences of the United States of America*. Jul 5 2016;113(27):7329-7336.
10.1073/pnas.1510502113
- Huser V, DeFalco FJ, Schuemie M, Ryan PB, Shang N, Velez M, . . . Bailey C. Multisite Evaluation of a Data Quality Tool for Patient-Level Clinical Data Sets. *EGEMS (Washington, DC)*. 2016;4(1):1239.
10.13063/2327-9214.1239

Huser V, Kahn MG, Brown JS, Gouripeddi R. Methods for examining data quality in healthcare integrated data repositories. *Pacific Symposium on Biocomputing*. Pacific Symposium on Biocomputing. 2018;23:628-633.

Jiang G, Kiefer R, Prud'hommeaux E, Solbrig HR. Building Interoperable FHIR-Based Vocabulary Mapping Services: A Case Study of OHDSI Vocabularies and Mappings. *Studies in health technology and informatics*. 2017;245:1327.

Jiang G, Kiefer RC, Sharma DK, Prud'hommeaux E, Solbrig HR. A Consensus-Based Approach for Harmonizing the OHDSI Common Data Model with HL7 FHIR. *Studies in health technology and informatics*. 2017;245:887-891.

Kahn MG, Brown JS, Chun AT, Davidson BN, Meeker D, Ryan PB, . . . Zozus MN. Transparent reporting of data quality in distributed data networks. *EGEMS (Washington, DC)*. 2015;3(1):1052.
10.13063/2327-9214.1052

Kang T, Zhang S, Tang Y, Hrubby GW, Rusanov A, Elhadad N, Weng C. EliiE: An open-source information extraction system for clinical trial eligibility criteria. *Journal of the American Medical Informatics Association : JAMIA*. Nov 1 2017;24(6):1062-1071.
10.1093/jamia/ocx019

Kim H, Choi J, Jang I, Quach J, Ohno-Machado L. Feasibility of Representing Data from Published Nursing Research Using the OMOP Common Data Model. *AMIA ... Annual Symposium proceedings. AMIA Symposium*. 2016;2016:715-723.

Knowledge Base workgroup of the Observational Health Data Sciences and Informatics (OHDSI) collaborative. Large-scale adverse effects related to treatment evidence standardization (LAERTES): an open scalable system for linking pharmacovigilance evidence sources with clinical data. *Journal of biomedical semantics*. Mar 7 2017;8(1):11.
10.1186/s13326-017-0115-3

Lambert CG, Mazurie AJ, Lauve NR, Hurwitz NG, Young SS, Obenchain RL, Hengartner NW, Perkins DJ, Tohen M, Kerner B. Hypothyroidism risk compared among nine common bipolar disorder therapies in a large US cohort. *Bipolar Disord*. 2016;18(3):247-260.
10.1111/bdi.12391

Lee S, Choi J, Kim HS, Kim GJ, Lee KH, Park CH, . . . Kim JH. Standard-based comprehensive detection of adverse drug reaction signals from nursing statements and laboratory results in electronic health records. *Journal of the American Medical Informatics Association : JAMIA*. Jul 1 2017;24(4):697-708.
10.1093/jamia/ocw168

Li L, Boland MR, Miotto R, Tatonetti NP, Dudley JT. Replicating Cardiovascular Condition-Birth Month Associations. *Scientific reports*. Sep 14 2016;6:33166.
10.1038/srep33166

Li Y, Ryan PB, Wei Y, Friedman C. A Method to Combine Signals from Spontaneous Reporting Systems and Observational Healthcare Data to Detect Adverse Drug Reactions. *Drug safety*. Oct 2015;38(10):895-908.
10.1007/s40264-015-0314-8

Lorberbaum T, Sampson KJ, Woosley RL, Kass RS, Tatonetti NP. An Integrative Data Science Pipeline to Identify Novel Drug Interactions that Prolong the QT Interval. *Drug safety*. May 2016;39(5):433-441.
10.1007/s40264-016-0393-1

Maier C, Lang L, Storf H, Vormstein P, Bieber R, Bernarding J, . . . Sedlmayr M. Towards Implementation of OMOP in a German University Hospital Consortium. *Applied clinical informatics*. Jan 2018;9(1):54-61.
10.1055/s-0037-1617452

Makadia R, Ryan PB. Transforming the Premier Perspective Hospital Database into the Observational Medical Outcomes Partnership (OMOP) Common Data Model. *EGEMS (Washington, DC)*. 2014;2(1):1110.
10.13063/2327-9214.1110

Matcho A, Ryan P, Fife D, Reich C. Fidelity assessment of a clinical practice research datalink conversion to the OMOP common data model. *Drug safety*. Nov 2014;37(11):945-959.
10.1007/s40264-014-0214-3

Moskovitch R, Polubriaginof F, Weiss A, Ryan P, Tatonetti N. Procedure prediction from symbolic Electronic Health Records via time intervals analytics. *Journal of biomedical informatics*. Nov 2017;75:70-82.
10.1016/j.jbi.2017.07.018

Natarajan S, Bangera V, Khot T, Picado J, Wazalwar A, Costa VS, . . . Caldwell M. Markov Logic Networks for Adverse Drug Event Extraction from Text. *Knowledge and information systems*. May 2017;51(2):435-457.
10.1007/s10115-016-0980-6

Natsiavas P, Boyce RD, Jaulent MC, Koutkias V. OpenPVSignal: Advancing Information Search, Sharing and Reuse on Pharmacovigilance Signals via FAIR Principles and Semantic Web Technologies. *Frontiers in pharmacology*. 2018;9:609.
10.3389/fphar.2018.00609

Nestsiarovich A, Mazurie AJ, Hurwitz NG, Kerner B, Nelson SJ, Crisanti AS, Tohen M, Krall RL, Perkins DJ, Lambert CG. Comprehensive comparison of monotherapies for psychiatric hospitalization risk in bipolar disorders. *Bipolar Disord*. June 2018.
10.1111/bdi.12665

Nissim N, Boland MR, Tatonetti NP, Elovici Y, Hripcsak G, Shahar Y, Moskovitch R. Improving condition severity classification with an efficient active learning based framework. *Journal of biomedical informatics*. Jun 2016;61:44-54.
10.1016/j.jbi.2016.03.016

Nissim N, Shahar Y, Elovici Y, Hripcsak G, Moskovitch R. Inter-labeler and intra-labeler variability of condition severity classification models using active and passive learning methods. *Artificial intelligence in medicine*. Sep 2017;81:12-32.
10.1016/j.artmed.2017.03.003

Ogunyemi OI, Meeker D, Kim HE, Ashish N, Farzaneh S, Boxwala A. Identifying appropriate reference data models for comparative effectiveness research (CER) studies based on data from clinical information systems. *Medical care*. Aug 2013;51(8 Suppl 3):S45-52.
10.1097/MLR.0b013e31829b1e0b

Ohno-Machado L, Agha Z, Bell DS, Dahm L, Day ME, Doctor JN, . . . Nebeker JR. pSCANNER: patient-centered Scalable National Network for Effectiveness Research. *Journal of the American Medical Informatics Association : JAMIA*. Jul-Aug 2014;21(4):621-626.
10.1136/amiajnl-2014-002751

Ong TC, Kahn MG, Kwan BM, Yamashita T, Brandt E, Hosokawa P, . . . Schilling LM. Dynamic-ETL: a hybrid approach for health data extraction, transformation and loading. *BMC medical informatics and decision making*. Sep 13 2017;17(1):134.
10.1186/s12911-017-0532-3

Overhage JM, Ryan PB, Reich CG, Hartzema AG, Stang PE. Validation of a common data model for active safety surveillance research. *Journal of the American Medical Informatics Association : JAMIA*. Jan-Feb 2012;19(1):54-60.
10.1136/amiajnl-2011-000376

Overhage JM, Ryan PB, Schuemie MJ, Stang PE. Authors' reply to Hennessy and Leonard's comment on "Desideratum for evidence-based epidemiology". *Drug safety*. Jan 2015;38(1):105-107.
10.1007/s40264-014-0254-8

Pacaci A, Gonul S, Sinaci AA, Yuksel M, Laleci Erturkmen GB. A Semantic Transformation Methodology for the Secondary Use of Observational Healthcare Data in Postmarketing Safety Studies. *Frontiers in pharmacology*. 2018;9:435.
10.3389/fphar.2018.00435

Panaccio MP, Cummins G, Wentworth C, Lanes S, Reynolds SL, Reynolds MW, . . . Koren A. A common data model to assess cardiovascular hospitalization and mortality in atrial fibrillation patients using administrative claims and medical records. *Clinical epidemiology*. 2015;7:77-90.
10.2147/clep.s64936

Park RW. Sharing Clinical Big Data While Protecting Confidentiality and Security: Observational Health Data Sciences and Informatics. *Healthcare informatics research*. Jan 2017;23(1):1-3.
10.4258/hir.2017.23.1.1

Patadia VK, Coloma P, Schuemie MJ, Herings R, Gini R, Mazzaglia G, . . . Trifiro G. Using real-world healthcare data for pharmacovigilance signal detection - the experience of the EU-ADR project. *Expert review of clinical pharmacology*. Jan 2015;8(1):95-102.

10.1586/17512433.2015.992878

Patadia VK, Schuemie MJ, Coloma PM, Herings R, van der Lei J, Sturkenboom M, Trifiro G. Can Electronic Health Records Databases Complement Spontaneous Reporting System Databases? A Historical-Reconstruction of the Association of Rofecoxib and Acute Myocardial Infarction. *Frontiers in pharmacology*. 2018;9:594.

10.3389/fphar.2018.00594

Polubriaginof FCG, Vanguri R, Quinnes K, Belbin GM, Yahi A, Salmasian H, . . . Tatonetti NP. Disease Heritability Inferred from Familial Relationships Reported in Medical Records. *Cell*. Jun 14 2018;173(7):1692-1704.e1611.

10.1016/j.cell.2018.04.032

Reps JM, Aickelin U, Hubbard RB. Refining adverse drug reaction signals by incorporating interaction variables identified using emergent pattern mining. *Computers in biology and medicine*. Feb 1 2016;69:61-70.

10.1016/j.combiomed.2015.11.014

Reps JM, Schuemie MJ, Suchard MA, Ryan PB, Rijnbeek PR. Design and implementation of a standardized framework to generate and evaluate patient-level prediction models using observational healthcare data. *Journal of the American Medical Informatics Association : JAMIA*. Apr 27 2018.

10.1093/jamia/ocy032

Rijnbeek PR. Converting to a common data model: what is lost in translation? : Commentary on "fidelity assessment of a clinical practice research datalink conversion to the OMOP common data model". *Drug safety*. Nov 2014;37(11):893-896.

10.1007/s40264-014-0221-4

Roberto G, Leal I, Sattar N, Loomis AK, Avillach P, Egger P, . . . Gini R. Identifying Cases of Type 2 Diabetes in Heterogeneous Data Sources: Strategy from the EMIF Project. *PloS one*. 2016;11(8):e0160648.

10.1371/journal.pone.0160648

Rosenbloom ST, Carroll RJ, Warner JL, Matheny ME, Denny JC. Representing Knowledge Consistently Across Health Systems. *Yearbook of medical informatics*. 01 Aug 2017;26(1):139-147.

Ryan PB, Buse JB, Schuemie MJ, DeFalco F, Yuan Z, Stang PE, . . . Norman R. Comparative effectiveness of canagliflozin, SGLT2 inhibitors and non-SGLT2 inhibitors on the risk of hospitalization for heart failure and amputation in patients with type 2 diabetes mellitus: A real-world meta-analysis of 4 observational databases (OBSERVE-4D). *Diabetes, Obesity and Metabolism*. 2018;0(0).

doi:10.1111/dom.13424

Samwald M, Xu H, Blagec K, Empey PE, Malone DC, Ahmed SM, . . . Boyce RD. Incidence of Exposure of Patients in the United States to Multiple Drugs for Which Pharmacogenomic Guidelines Are Available. *PloS one*. 2016;11(10):e0164972.

10.1371/journal.pone.0164972

Schuemie MJ, Gini R, Coloma PM, Straatman H, Herings RM, Pedersen L, . . . Sturkenboom MC. Replication of the OMOP experiment in Europe: evaluating methods for risk identification in electronic health record databases. *Drug safety*. Oct 2013;36 Suppl 1:S159-169.

10.1007/s40264-013-0109-8

Schuemie MJ, Hripcsak G, Ryan PB, Madigan D, Suchard MA. Robust empirical calibration of p-values using observational data. *Statistics in medicine*. Sep 30 2016;35(22):3883-3888.

10.1002/sim.6977

Schuemie MJ, Hripcsak G, Ryan PB, Madigan D, Suchard MA. Empirical confidence interval calibration for population-level effect estimation studies in observational healthcare data. *Proceedings of the National Academy of Sciences of the United States of America*. Mar 13 2018;115(11):2571-2577.

10.1073/pnas.1708282114

Schuemie MJ, Ryan PB, DuMouchel W, Suchard MA, Madigan D. Interpreting observational studies: why empirical calibration is needed to correct p-values. *Statistics in medicine*. Jan 30 2014;33(2):209-218.

10.1002/sim.5925

Schuemie MJ, Ryan PB, Suchard MA, Shahn Z, Madigan D. Discussion: An estimate of the science-wise false discovery rate and application to the top medical literature. *Biostatistics (Oxford, England)*. Jan 2014;15(1):36-39; discussion 39-45.

10.1093/biostatistics/kxt037

Schuemie MJ, Trifiro G, Coloma PM, Ryan PB, Madigan D. Detecting adverse drug reactions following long-term exposure in longitudinal observational data: The exposure-adjusted self-controlled case series. *Statistical methods in medical research*. Dec 2016;25(6):2577-2592.

10.1177/0962280214527531

Sen A, Chakrabarti S, Goldstein A, Wang S, Ryan PB, Weng C. GIST 2.0: A scalable multi-trait metric for quantifying population representativeness of individual clinical studies. *Journal of biomedical informatics*. Oct 2016;63:325-336.

10.1016/j.jbi.2016.09.003

Sen A, Ryan PB, Goldstein A, Chakrabarti S, Wang S, Koski E, Weng C. Correlating eligibility criteria generalizability and adverse events using Big Data for patients and clinical trials. *Annals of the New York Academy of Sciences*. Jan 2017;1387(1):34-43.

10.1111/nyas.13195

Shaddox TR, Ryan PB, Schuemie MJ, Madigan D, Suchard MA. Hierarchical Models for Multiple, Rare Outcomes Using Massive Observational Healthcare Databases. *Statistical analysis and data mining*. Aug 2016;9(4):260-268.

10.1002/sam.11324

Shang N, Weng C, Hripcsak G. A conceptual framework for evaluating data suitability for observational studies. *Journal of the American Medical Informatics Association : JAMIA*. Sep 8 2017.

10.1093/jamia/ocx095

Si Y, Weng C. An OMOP CDM-Based Relational Database of Clinical Research Eligibility Criteria. *Studies in health technology and informatics*. 2017;245:950-954.

Tian Y, Schuemie MJ, Suchard MA. Evaluating large-scale propensity score performance through real-world and synthetic data experiments. *International journal of epidemiology*. Jun 22 2018.

10.1093/ije/dyy120

Vilar S, Friedman C, Hripcsak G. Detection of drug-drug interactions through data mining studies using clinical sources, scientific literature and social media. *Briefings in bioinformatics*. Feb 17 2017.

10.1093/bib/bbx010

Vilar S, Lorberbaum T, Hripcsak G, Tatonetti NP. Improving Detection of Arrhythmia Drug-Drug Interactions in Pharmacovigilance Data through the Implementation of Similarity-Based Modeling. *PLoS one*. 2015;10(6):e0129974.

10.1371/journal.pone.0129974

Vilar S, Ryan PB, Madigan D, Stang PE, Schuemie MJ, Friedman C, . . . Hripcsak G. Similarity-based modeling applied to signal detection in pharmacovigilance. *CPT: pharmacometrics & systems pharmacology*. Sep 24 2014;3:e137.

10.1038/psp.2014.35

Voss EA, Boyce RD, Ryan PB, van der Lei J, Rijnbeek PR, Schuemie MJ. Accuracy of an automated knowledge base for identifying drug adverse reactions. *Journal of biomedical informatics*. Feb 2017;66:72-81.

10.1016/j.jbi.2016.12.005

Voss EA, Ma Q, Ryan PB. The impact of standardizing the definition of visits on the consistency of multi-database observational health research. *BMC medical research methodology*. Mar 8 2015;15:13.

10.1186/s12874-015-0001-6

Voss EA, Makadia R, Matcho A, Ma Q, Knoll C, Schuemie M, . . . Ryan PB. Feasibility and utility of applications of the common data model to multiple, disparate observational health databases. *Journal of the American Medical Informatics Association : JAMIA*. May 2015;22(3):553-564.

10.1093/jamia/ocu023

- Voss EA, Ryan PB, Stang PE, Hough D, Alphs L. Switching from risperidone long-acting injectable to paliperidone long-acting injectable or oral antipsychotics: analysis of a Medicaid claims database. *International clinical psychopharmacology*. May 2015;30(3):151-157.
10.1097/yic.0000000000000068
- Weinstein RB, Schuemie MJ, Ryan PB, Stang PE. Seasonality in acute liver injury? Findings in two health care claims databases. *Drug, healthcare and patient safety*. 2016;8:39-48.
10.2147/dhps.s95399
- Weng C, Li Y, Ryan P, Zhang Y, Liu F, Gao J, . . . Hripcsak G. A distribution-based method for assessing the differences between clinical trial target populations and patient populations in electronic health records. *Applied clinical informatics*. 2014;5(2):463-479.
10.4338/aci-2013-12-ra-0105
- Xu Y, Zhou X, Suehs BT, Hartzema AG, Kahn MG, Moride Y, . . . Bate A. A Comparative Assessment of Observational Medical Outcomes Partnership and Mini-Sentinel Common Data Models and Analytics: Implications for Active Drug Safety Surveillance. *Drug safety*. Aug 2015;38(8):749-765.
10.1007/s40264-015-0297-5
- Yang Y, Zhou X, Gao S, Lin H, Xie Y, Feng Y, . . . Zhan S. Evaluation of Electronic Healthcare Databases for Post-Marketing Drug Safety Surveillance and Pharmacoepidemiology in China. *Drug Safety*. 01 Jan 2018;41(1):125-137.
- Yoon D, Ahn EK, Park MY, Cho SY, Ryan P, Schuemie MJ, . . . Park RW. Conversion and Data Quality Assessment of Electronic Health Record Data at a Korean Tertiary Teaching Hospital to a Common Data Model for Distributed Network Research. *Healthcare informatics research*. Jan 2016;22(1):54-58.
10.4258/hir.2016.22.1.54
- Yoon D, Schuemie MJ, Kim JH, Kim DK, Park MY, Ahn EK, . . . Park RW. A normalization method for combination of laboratory test results from different electronic healthcare databases in a distributed research network. *Pharmacoepidemiology and drug safety*. Mar 2016;25(3):307-316.
10.1002/pds.3893
- You SC, Lee S, Cho SY, Park H, Jung S, Cho J, . . . Park RW. Conversion of National Health Insurance Service-National Sample Cohort (NHIS-NSC) Database into Observational Medical Outcomes Partnership-Common Data Model (OMOP-CDM). *Studies in health technology and informatics*. 2017;245:467-470.
- Yuan Z, Voss EA, DeFalco FJ, Pan G, Ryan PB, Yannicelli D, Nessel C. Risk Prediction for Ischemic Stroke and Transient Ischemic Attack in Patients Without Atrial Fibrillation: A Retrospective Cohort Study. *Journal of stroke and cerebrovascular diseases : the official journal of National Stroke Association*. Aug 2017;26(8):1721-1731.
10.1016/j.jstrokecerebrovasdis.2017.03.036
- Zhou X, Murugesan S, Bhullar H, Liu Q, Cai B, Wentworth C, Bate A. An evaluation of the THIN database in the OMOP Common Data Model for active drug safety surveillance. *Drug safety*. Feb 2013;36(2):119-134.
10.1007/s40264-012-0009-3