Martijn J Schuemie

List of Publications by Year
in descending order

111 papers	6,377 citations	71102 41 h-index	79698 73 g-index
120 all docs	120 docs citations	120 times ranked	7916 citing authors

#	Article	IF	Citations
1	Observational Health Data Sciences and Informatics (OHDSI): Opportunities for Observational Researchers. Studies in Health Technology and Informatics, 2015, 216, 574-8.	0.3	533
2	Characterizing treatment pathways at scale using the OHDSI network. Proceedings of the National Academy of Sciences of the United States of America, 2016, 113, 7329-7336.	7.1	256
3	Overview of BioCreative II gene normalization. Genome Biology, 2008, 9, S3.	9.6	237
4	Comprehensive comparative effectiveness and safety of first-line antihypertensive drug classes: a systematic, multinational, large-scale analysis. Lancet, The, 2019, 394, 1816-1826.	13.7	228
5	Combining electronic healthcare databases in Europe to allow for largeâ€scale drug safety monitoring: the EUâ€ADR Project. Pharmacoepidemiology and Drug Safety, 2011, 20, 1-11.	1.9	222
6	Feasibility and utility of applications of the common data model to multiple, disparate observational health databases. Journal of the American Medical Informatics Association: JAMIA, 2015, 22, 553-564.	4.4	198
7	Non-steroidal anti-inflammatory drugs and risk of heart failure in four European countries: nested case-control study. BMJ, The, 2016, 354, i4857.	6.0	195
8	Comparative effectiveness of canagliflozin, SGLT2 inhibitors and nonâ€SGLT2 inhibitors on the risk of hospitalization for heart failure and amputation in patients with type 2 diabetes mellitus: A realâ€world metaâ€analysis of 4 observational databases (OBSERVEâ€4D). Diabetes, Obesity and Metabolism, 2018, 20, 2585-2597.	4.4	164
9	Interpreting observational studies: why empirical calibration is needed to correct pâ€values. Statistics in Medicine, 2014, 33, 209-218.	1.6	163
10	Evaluating the Impact of Database Heterogeneity on Observational Study Results. American Journal of Epidemiology, 2013, 178, 645-651.	3.4	149
11	Risk of Upper Gastrointestinal Bleeding From Different Drug Combinations. Gastroenterology, 2014, 147, 784-792.e9.	1.3	132
12	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, 2018, 25, 969-975.	4.4	131
13	Risk of hydroxychloroquine alone and in combination with azithromycin in the treatment of rheumatoid arthritis: a multinational, retrospective study. Lancet Rheumatology, The, 2020, 2, e698-e711.	3.9	117
14	A dictionary to identify small molecules and drugs in free text. Bioinformatics, 2009, 25, 2983-2991.	4.1	116
15	Evaluating large-scale propensity score performance through real-world and synthetic data experiments. International Journal of Epidemiology, 2018, 47, 2005-2014.	1.9	113
16	Association of Ticagrelor vs Clopidogrel With Net Adverse Clinical Events in Patients With Acute Coronary Syndrome Undergoing Percutaneous Coronary Intervention. JAMA - Journal of the American Medical Association, 2020, 324, 1640.	7.4	112
17	Defining a Reference Set to Support Methodological Research in Drug Safety. Drug Safety, 2013, 36, 33-47.	3.2	109
18	Risk of lower extremity amputations in people with type 2 diabetes mellitus treated with sodiumâ€glucose coâ€transporterâ€2 inhibitors in the USA: A retrospective cohort study. Diabetes, Obesity and Metabolism, 2018, 20, 582-589.	4.4	108

#	Article	IF	Citations
19	Anni 2.0: a multipurpose text-mining tool for the life sciences. Genome Biology, 2008, 9, R96.	9.6	104
20	Identification of acute myocardial infarction from electronic healthcare records using different disease coding systems: a validation study in three European countries. BMJ Open, 2013, 3, e002862.	1.9	101
21	Comparison of Cardiovascular and Safety Outcomes of Chlorthalidone vs Hydrochlorothiazide to Treat Hypertension. JAMA Internal Medicine, 2020, 180, 542.	5.1	97
22	Reninâ€“angiotensin system blockers and susceptibility to COVID-19: an international, open science, cohort analysis. The Lancet Digital Health, 2021, 3, e98-e114.	12.3	94
23	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, 2018, 115, 2571-2577.	7.1	91
24	Deep phenotyping of 34,128 adult patients hospitalised with COVID-19 in an international network study. Nature Communications, 2020, 11, 5009.	12.8	86
25	Accuracy of an automated knowledge base for identifying drug adverse reactions. Journal of Biomedical Informatics, 2017, 66, 72-81.	4.3	85
26	Using Electronic Health Care Records for Drug Safety Signal Detection. Medical Care, 2012, 50, 890-897.	2.4	79
27	Chronic disease prevalence from Italian administrative databases in the VALORE project: a validation through comparison of population estimates with general practice databases and national survey. BMC Public Health, 2013, 13, 15.	2.9	79
28	Harmonization process for the identification of medical events in eight European healthcare databases: the experience from the EU-ADR project. Journal of the American Medical Informatics Association: JAMIA, 2013, 20, 184-192.	4.4	75
29	Word Sense Disambiguation in the Biomedical Domain: An Overview. Journal of Computational Biology, 2005, 12, 554-565.	1.6	74
30	Comparative First-Line Effectiveness and Safety of ACE (Angiotensin-Converting Enzyme) Inhibitors and Angiotensin Receptor Blockers: A Multinational Cohort Study. Hypertension, 2021, 78, 591-603.	2.7	63
31	Learning from electronic health records across multiple sites: A communication-efficient and privacy-preserving distributed algorithm. Journal of the American Medical Informatics Association: JAMIA, 2020, 27, 376-385.	4.4	61
32	Jane: suggesting journals, finding experts. Bioinformatics, 2008, 24, 727-728.	4.1	58
33	Empirical Performance of a New User Cohort Method: Lessons for Developing a Risk Identification and Analysis System. Drug Safety, 2013, 36, 59-72.	3.2	57
34	Association of Hemoglobin A_{1c} Levels With Use of Sulfonylureas, Dipeptidyl Peptidase 4 Inhibitors, and Thiazolidinediones in Patients With Type 2 Diabetes Treated With Metformin. JAMA Network Open, 2018, 1, e181755.	5.9	54
35	Guillain-BarrÃ© Syndrome and Adjuvanted Pandemic Influenza A (H1N1) 2009 Vaccines: A Multinational Self-Controlled Case Series in Europe. PLoS ONE, 2014, 9, e82222.	2.5	53
36	Incidence of diabetic ketoacidosis among patients with type 2 diabetes mellitus treated with SGLT2 inhibitors and other antihyperglycemic agents. Diabetes Research and Clinical Practice, 2017, 128, 83-90.	2.8	53

#	Article	IF	Citations
37	Improving reproducibility by using high-throughput observational studies with empirical calibration. Philosophical Transactions Series A, Mathematical, Physical, and Engineering Sciences, 2018, 376, 20170356.	3.4	53
38	Validation study in four health-care databases: upper gastrointestinal bleeding misclassification affects precision but not magnitude of drug-related upper gastrointestinal bleeding risk. Journal of Clinical Epidemiology, 2014, 67, 921-931.	5.0	49
39	Useful Interplay Between Spontaneous ADR Reports and Electronic Healthcare Records in Signal Detection. Drug Safety, 2015, 38, 1201-1210.	3.2	49
40	A Systematic Statistical Approach to Evaluating Evidence from Observational Studies. Annual Review of Statistics and Its Application, 2014, 1, 11-39.	7.0	46
41	Risk of acute myocardial infarction during use of individual NSAIDs: A nested case-control study from the SOS project. PLoS ONE, 2018, 13, e0204746.	2.5	46
42	Learning from local to global: An efficient distributed algorithm for modeling time-to-event data. Journal of the American Medical Informatics Association: JAMIA, 2020, 27, 1028-1036.	4.4	46
43	Robust empirical calibration of pâ€values using observational data. Statistics in Medicine, 2016, 35, 3883-3888.	1.6	43
44	Data Extraction And Management In Networks Of Observational Health Care Databases For Scientific Research: A Comparison Among EU-ADR, OMOP, Mini-Sentinel And MATRICE Strategies. EGEMS (Washington, DC), 2017, 4, 2.	2.0	43
45	A plea to stop using the caseâ€control design in retrospective database studies. Statistics in Medicine, 2019, 38, 4199-4208.	1.6	42
46	Replication of the OMOP Experiment in Europe: Evaluating Methods for Risk Identification in Electronic Health Record Databases. Drug Safety, 2013, 36, 159-169.	3.2	41
47	Drug-Induced Acute Myocardial Infarction: Identifying â€˜Prime Suspectsâ€™ from Electronic Healthcare Records-Based Surveillance System. PLoS ONE, 2013, 8, e72148.	2.5	41
48	Novel Protein-Protein Interactions Inferred from Literature Context. PLoS ONE, 2009, 4, e7894.	2.5	41
49	Improving sensitivity of machine learning methods for automated case identification from free-text electronic medical records. BMC Medical Informatics and Decision Making, 2013, 13, 30.	3.0	40
50	Risk of ischemic stroke and the use of individual non-steroidal anti-inflammatory drugs: A multi-country European database study within the SOS Project. PLoS ONE, 2018, 13, e0203362.	2.5	38
51	Risk of angioedema associated with levetiracetam compared with phenytoin: Findings of the observational health data sciences and informatics research network. Epilepsia, 2017, 58, e101-e106.	5.1	37
52	Multisite Evaluation of a Data Quality Tool for Patient-Level Clinical Datasets. EGEMS (Washington,) Tj ETQq0 0 0 rgBT /Overlock 10 Tf 5	2.0	37
53	Literature-based concept profiles for gene annotation: The issue of weighting. International Journal of Medical Informatics, 2008, 77, 354-362.	3.3	35
54	Uncovering exposures responsible for birth season â€“ disease effects: a global study. Journal of the American Medical Informatics Association: JAMIA, 2018, 25, 275-288.	4.4	33

#	Article	IF	Citations
55	ContextD: an algorithm to identify contextual properties of medical terms in a Dutch clinical corpus. BMC Bioinformatics, 2014, 15, 373.	2.6	32
56	How Confident Are We About Observational Findings in Health Care: A Benchmark Study. , 2020, 2, .		32
57	Automating classification of freeâ€text electronic health records for epidemiological studies. Pharmacoepidemiology and Drug Safety, 2012, 21, 651-658.	1.9	31
58	Using real-world healthcare data for pharmacovigilance signal detection â€“ the experience of the EU-ADR project. Expert Review of Clinical Pharmacology, 2015, 8, 95-102.	3.1	31
59	Evaluating performance of electronic healthcare records and spontaneous reporting data in drug safety signal detection. International Journal of Clinical Pharmacy, 2015, 37, 94-104.	2.1	31
60	Principles of Large-scale Evidence Generation and Evaluation across a Network of Databases (LEGEND). Journal of the American Medical Informatics Association: JAMIA, 2020, 27, 1331-1337.	4.4	31
61	Evaluation of techniques for increasing recall in a dictionary approach to gene and protein name identification. Journal of Biomedical Informatics, 2007, 40, 316-324.	4.3	30
62	Empirical Performance of the Caseâ€“Control Method: Lessons for Developing a Risk Identification and Analysis System. Drug Safety, 2013, 36, 73-82.	3.2	28
63	Does design matter? Systematic evaluation of the impact of analytical choices on effect estimates in observational studies. Therapeutic Advances in Drug Safety, 2013, 4, 53-62.	2.4	27
64	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, 2016, 6, e012413.	1.9	26
65	Signal Detection of Potentially Drug-Induced Acute Liver Injury in Children Using a Multi-Country Healthcare Database Network. Drug Safety, 2014, 37, 99-108.	3.2	25
66	The role of electronic healthcare record databases in paediatric drug safety surveillance: a retrospective cohort study. British Journal of Clinical Pharmacology, 2015, 80, 304-314.	2.4	25
67	Rewriting and suppressing UMLS terms for improved biomedical term identification. Journal of Biomedical Semantics, 2010, 1, 5.	1.6	24
68	Can Italian Healthcare Administrative Databases Be Used to Compare Regions with Respect to Compliance with Standards of Care for Chronic Diseases?. PLoS ONE, 2014, 9, e95419.	2.5	24
69	Applying a common data model to Asian databases for multinational pharmacoepidemiologic studies: opportunities and challenges. Clinical Epidemiology, 2018, Volume 10, 875-885.	3.0	24
70	Comparative safety and effectiveness of alendronate versus raloxifene in women with osteoporosis. Scientific Reports, 2020, 10, 11115.	3.3	23
71	The Implicitome: A Resource for Rationalizing Gene-Disease Associations. PLoS ONE, 2016, 11, e0149621.	2.5	22
72	Drawing Reproducible Conclusions from Observational Clinical Data with OHDSI. Yearbook of Medical Informatics, 2021, 30, 283-289.	1.0	21

#	Article	IF	Citations
73	In silico discovery and experimental validation of new proteinâ€“protein interactions. Proteomics, 2011, 11, 843-853.	2.2	20
74	Identifying Cases of Type 2 Diabetes in Heterogeneous Data Sources: Strategy from the EMIF Project. PLoS ONE, 2016, 11, e0160648.	2.5	20
75	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.	3.5	20
76	Comprehensive Comparative Effectiveness and Safety of First-Line Î²-Blocker Monotherapy in Hypertensive Patients. Hypertension, 2021, 77, 1528-1538.	2.7	20
77	Literature-aided interpretation of gene expression data with the weighted global test. Briefings in Bioinformatics, 2011, 12, 518-529.	6.5	19
78	Desideratum for Evidence Based Epidemiology. Drug Safety, 2013, 36, 5-14.	3.2	19
79	Large-scale evidence generation and evaluation across a network of databases (LEGEND): assessing validity using hypertension as a case study. Journal of the American Medical Informatics Association: JAMIA, 2020, 27, 1268-1277.	4.4	19
80	Comparison of First-Line Dual Combination Treatments in Hypertension: Real-World Evidence from Multinational Heterogeneous Cohorts. Korean Circulation Journal, 2020, 50, 52.	1.9	19
81	Channeling in the Use of Nonprescription Paracetamol and Ibuprofen in an Electronic Medical Records Database: Evidence and Implications. Drug Safety, 2017, 40, 1279-1292.	3.2	18
82	A standardized analytics pipeline for reliable and rapid development and validation of prediction models using observational health data. Computer Methods and Programs in Biomedicine, 2021, 211, 106394.	4.7	18
83	Assignment of protein function and discovery of novel nucleolar proteins based on automatic analysis of MEDLINE. Proteomics, 2007, 7, 921-931.	2.2	16
84	Incidence, prevalence and prescription patterns of antipsychotic medications use in Asia and US: A cross-nation comparison with common data model. Journal of Psychiatric Research, 2020, 131, 77-84.	3.1	16
85	Bias, Precision and Timeliness of Historical (Background) Rate Comparison Methods for Vaccine Safety Monitoring: An Empirical Multi-Database Analysis. Frontiers in Pharmacology, 2021, 12, 773875.	3.5	13
86	Variation in Choice of Study Design: Findings from the Epidemiology Design Decision Inventory and Evaluation (EDDIE) Survey. Drug Safety, 2013, 36, 15-25.	3.2	12
87	Atypical Antipsychotics and the Risk of Falls and Fractures Among Older Adults. Journal of Clinical Psychopharmacology, 2017, 37, 162-168.	1.4	11
88	dPQL: a lossless distributed algorithm for generalized linear mixed model with application to privacy-preserving hospital profiling. Journal of the American Medical Informatics Association: JAMIA, 2022, 29, 1366-1371.	4.4	10
89	Atypical Antipsychotics and the Risks of Acute Kidney Injury and Related Outcomes Among Older Adults: A Replication Analysis and an Evaluation of Adapted Confounding Control Strategies. Drugs and Aging, 2017, 34, 211-219.	2.7	9
90	Applied comparison of largeâ€scale propensity score matching and cardinality matching for causal inference in observational research. BMC Medical Research Methodology, 2021, 21, 109.	3.1	9

#	Article	IF	Citations
91	DLMM as a lossless one-shot algorithm for collaborative multi-site distributed linear mixed models. Nature Communications, 2022, 13, 1678.	12.8	9
92	Current Approaches to Vaccine Safety Using Observational Data: A Rationale for the EUMAEUS (Evaluating Use of Methods for Adverse Events Under Surveillance-for Vaccines) Study Design. Frontiers in Pharmacology, 2022, 13, 837632.	3.5	8
93	Large-scale evidence generation and evaluation across a network of databases for type 2 diabetes mellitus (LEGEND-T2DM): a protocol for a series of multinational, real-world comparative cardiovascular effectiveness and safety studies. BMJ Open, 2022, 12, e057977.	1.9	8
94	GeneE: Gene and protein query expansion with disambiguation. Bioinformatics, 2010, 26, 147-148.	4.1	7
95	Monitoring compliance with standards of care for chronic diseases using healthcare administrative databases in Italy: Strengths and limitations. PLoS ONE, 2017, 12, e0188377.	2.5	7
96	Channeling Bias in the Analysis of Risk of Myocardial Infarction, Stroke, Gastrointestinal Bleeding, and Acute Renal Failure with the Use of Paracetamol Compared with Ibuprofen. Drug Safety, 2020, 43, 927-942.	3.2	7
97	Empirical assessment of caseâ€based methods for drug safety alert identification in the French National Healthcare System database (SNDS): Methodology of the ALCAPONE project. Pharmacoepidemiology and Drug Safety, 2020, 29, 993-1000.	1.9	6
98	Empirical assessment of caseâ€based methods for identification of drugs associated with upper gastrointestinal bleeding in the French National Healthcare System database (<scp>SNDS</scp>). Pharmacoepidemiology and Drug Safety, 2020, 29, 890-903.	1.9	5
99	Combining cox regressions across a heterogeneous distributed research network facing small and zero counts. Statistical Methods in Medical Research, 2021, , 096228022110605.	1.5	5
100	Discussion: An estimate of the science-wise false discovery rate and application to the top medical literature. Biostatistics, 2014, 15, 36-39.	1.5	4
101	Generating and evaluating a propensity model using textual features from electronic medical records. PLoS ONE, 2019, 14, e0212999.	2.5	4
102	Using the Data Quality Dashboard to Improve the EHDEN Network. Applied Sciences (Switzerland), 2021, 11, 11920.	2.5	4
103	Application of a Common Data Model (CDM) to rank the paediatric user and prescription prevalence of 15 different drug classes in South Korea, Hong Kong, Taiwan, Japan and Australia: an observational, descriptive study. BMJ Open, 2020, 10, e032426.	1.9	3
104	Acute pancreatitis risk in type 2 diabetes patients treated with canagliflozin versus other antihyperglycemic agents: an observational claims database study. Current Medical Research and Opinion, 2020, 36, 1117-1124.	1.9	3
105	Empirical assessment of caseâ€based methods for identification of drugs associated with acute liver injury in the French National Healthcare System database (<scp>SNDS</scp>). Pharmacoepidemiology and Drug Safety, 2021, 30, 320-333.	1.9	3
106	Quantifying bias in epidemiologic studies evaluating the association between acetaminophen use and cancer. Regulatory Toxicology and Pharmacology, 2021, 120, 104866.	2.7	3
107	A systematic assessment of the epidemiologic literature regarding an association between acetaminophen exposure and cancer. Regulatory Toxicology and Pharmacology, 2021, 127, 105043.	2.7	3
108	Characterizing Anchoring Bias in Vaccine Comparator Selection Due to Health Care Utilization With COVID-19 and Influenza: Observational Cohort Study. JMIR Public Health and Surveillance, 2022, 8, e33099.	2.6	2

#	Article	IF	Citations
109	Hip Fracture Risk After Treatment with Tramadol or Codeine: An Observational Study. Drug Safety, 2022, 45, 791-807.	3.2	2
110	Comment on â€œHow pharmacoepidemiology networks can manage distributed analyses to improve replicability and transparency and minimize biasâ€. Pharmacoepidemiology and Drug Safety, 2019, 28, 1032-1033.	1.9	1
111	Chlorthalidone and Hydrochlorothiazide for Treatment of Patients With Hypertensionâ€”Reply. JAMA Internal Medicine, 2020, 180, 1133.	5.1	1