Yong Chen

List of Publications by Year
in descending order

134 papers	2,326 citations	331670 21 h-index	302126 39 g-index
145 all docs	145 docs citations	145 times ranked	3180 citing authors

#	ARTICLE	IF	CITATIONS
1	Variation in US Hospital Mortality Rates for Patients Admitted With COVID-19 During the First 6 Months of the Pandemic. JAMA Internal Medicine, 2021, 181, 471.	5.1	197
2	Empirical Comparison of Publication Bias Tests in Meta-Analysis. Journal of General Internal Medicine, 2018, 33, 1260-1267.	2.6	184
3	A Meta-Analysis of the Associations Between the Nurse Work Environment in Hospitals and 4 Sets of Outcomes. Medical Care, 2019, 57, 353-361.	2.4	182
4	Electronic health records and polygenic risk scores for predicting disease risk. Nature Reviews Genetics, 2020, 21, 493-502.	16.3	78
5	A historical review of publication bias. Research Synthesis Methods, 2020, 11, 725-742.	8.7	69
6	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
7	Dexras1, a Small GTPase, Is Required for Glutamate-NMDA Neurotoxicity. Journal of Neuroscience, 2013, 33, 3582-3587.	3.6	60
8	Bivariate random effects models for meta-analysis of comparative studies with binary outcomes: Methods for the absolute risk difference and relative risk. Statistical Methods in Medical Research, 2012, 21, 621-633.	1.5	58
9	Lysosomal iron modulates NMDA receptor-mediated excitation via small GTPase, Dexras1. Molecular Brain, 2016, 9, 38.	2.6	47
10	Cytosine methylation predicts renal function decline in American Indians. Kidney International, 2018, 93, 1417-1431.	5.2	46
11	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
12	Enhancement of Stress Resilience Through Histone Deacetylase 6–Mediated Regulation of Glucocorticoid Receptor Chaperone Dynamics. Biological Psychiatry, 2015, 77, 345-355.	1.3	44
13	Association of Race/Ethnicity With Hospital Discharge Disposition After Elective Total Knee Arthroplasty. JAMA Network Open, 2019, 2, e1914259.	5.9	37
14	Diagnostic accuracy of the <scp>PLASMIC</scp> score in patients with suspected thrombotic thrombocytopenic purpura: A systematic review and <scp>meta‐analysis</scp>. Transfusion, 2020, 60, 2047-2057.	1.6	37
15	On the asymptotic behaviour of the pseudolikelihood ratio test statistic with boundary problems. Biometrika, 2010, 97, 603-620.	2.4	33
16	Use of Deep Learning to Analyze Social Media Discussions About the Human Papillomavirus Vaccine. JAMA Network Open, 2020, 3, e2022025.	5.9	32
17	SCOR: A secure international informatics infrastructure to investigate COVID-19. Journal of the American Medical Informatics Association: JAMIA, 2020, 27, 1721-1726.	4.4	31
18	Does biologic therapy impact the development of PsA among patients with psoriasis?. Annals of the Rheumatic Diseases, 2022, 81, 80-86.	0.9	29

#	ARTICLE	IF	CITATIONS
19	An alternative pseudolikelihood method for multivariate random-effects meta-analysis. Statistics in Medicine, 2015, 34, 361-380.	1.6	28
20	A Unification of Models for Meta-Analysis of Diagnostic Accuracy Studies without a Gold Standard. Biometrics, 2015, 71, 538-547.	1.4	27
21	A Bayesian hierarchical model for network meta-analysis of multiple diagnostic tests. Biostatistics, 2018, 19, 87-102.	1.5	24
22	Testing small study effects in multivariate meta-analysis. Biometrics, 2020, 76, 1240-1250.	1.4	24
23	Heterogeneity-aware and communication-efficient distributed statistical inference. Biometrika, 2022, 109, 67-83.	2.4	24
24	Bayesian analysis on meta-analysis of case-control studies accounting for within-study correlation. Statistical Methods in Medical Research, 2015, 24, 836-855.	1.5	23
25	Phospholipase C Beta 1: a Candidate Signature Gene for Proneural Subtype High-Grade Glioma. Molecular Neurobiology, 2016, 53, 6511-6525.	4.0	23
26	A composite likelihood method for bivariate meta-analysis in diagnostic systematic reviews. Statistical Methods in Medical Research, 2017, 26, 914-930.	1.5	23
27	Leveraging deep learning to understand health beliefs about the Human Papillomavirus Vaccine from social media. Npj Digital Medicine, 2019, 2, 27.	10.9	22
28	Extracting postmarketing adverse events from safety reports in the vaccine adverse event reporting system (VAERS) using deep learning. Journal of the American Medical Informatics Association: JAMIA, 2021, 28, 1393-1400.	4.4	22
29	Ideas for how informaticians can get involved with COVID-19 research. BioData Mining, 2020, 13, 3.	4.0	20
30	A Bayesian latent class approach for EHR-based phenotyping. Statistics in Medicine, 2019, 38, 74-87.	1.6	19
31	Risk Factors for Diagnosis of Psoriatic Arthritis, Psoriasis, Rheumatoid Arthritis, and Ankylosing Spondylitis: A Set of Parallel Case-control Studies. Journal of Rheumatology, 2022, 49, 53-59.	2.0	19
32	An Empirical Study for Impacts of Measurement Errors on EHR based Association Studies. AMIA … Annual Symposium proceedings, 2016, 2016, 1764-1773.	0.2	19
33	Bayesian hierarchical methods for meta-analysis combining randomized-controlled and single-arm studies. Statistical Methods in Medical Research, 2019, 28, 1293-1310.	1.5	18
34	Regression analysis of longitudinal data with irregular and informative observation times. Biostatistics, 2015, 16, 727-739.	1.5	17
35	Neuronal Activity-Induced Sterol Regulatory Element Binding Protein-1 (SREBP1) is Disrupted in Dysbindin-Null Mice—Potential Link to Cognitive Impairment in Schizophrenia. Molecular Neurobiology, 2017, 54, 1699-1709.	4.0	17
36	Integration of genetic and clinical information to improve imputation of data missing from electronic health records. Journal of the American Medical Informatics Association: JAMIA, 2019, 26, 1056-1063.	4.4	17

#	Article	IF	Citations
37	Inflation of type I error rates due to differential misclassification in EHR‐derived outcomes: Empirical illustration using breast cancer recurrence. Pharmacoepidemiology and Drug Safety, 2019, 28, 264-268.	1.9	17
38	Meta‐analysis of studies with bivariate binary outcomes: a marginal beta‐binomial model approach. Statistics in Medicine, 2016, 35, 21-40.	1.6	16
39	Difference Between Users and Nonusers of a Patient Portal in Health Behaviors and Outcomes: Retrospective Cohort Study. Journal of Medical Internet Research, 2019, 21, e13146.	4.3	16
40	Semiparametric Tests for Identifying Differentially Methylated Loci With Case–Control Designs Using Illumina Arrays. Genetic Epidemiology, 2014, 38, 42-50.	1.3	15
41	PIE: A prior knowledge guided integrated likelihood estimation method for bias reduction in association studies using electronic health records data. Journal of the American Medical Informatics Association: JAMIA, 2018, 25, 345-352.	4.4	15
42	Prediction of 30-day pediatric unplanned hospitalizations using the Johns Hopkins Adjusted Clinical Groups risk adjustment system. PLoS ONE, 2019, 14, e0221233.	2.5	15
43	Predictive P-score for treatment ranking in Bayesian network meta-analysis. BMC Medical Research Methodology, 2021, 21, 213.	3.1	15
44	A hybrid model for combining case-control and cohort studies in systematic reviews of diagnostic tests. Journal of the Royal Statistical Society Series C: Applied Statistics, 2015, 64, 469-489.	1.0	14
45	Visualized Emotion Ontology: a model for representing visual cues of emotions. BMC Medical Informatics and Decision Making, 2018, 18, 64.	3.0	14
46	Why Is the Electronic Health Record So Challenging for Research and Clinical Care?. Methods of Information in Medicine, 2021, 60, 032-048.	1.2	13
47	ODAL: A one-shot distributed algorithm to perform logistic regressions on electronic health records data from multiple clinical sites. , 2018, , .		13
48	A hybrid Bayesian hierarchical model combining cohort and case–control studies for meta-analysis of diagnostic tests: Accounting for partial verification bias. Statistical Methods in Medical Research, 2016, 25, 3015-3037.	1.5	12
49	pETM: a penalized Exponential Tilt Model for analysis of correlated high-dimensional DNA methylation data. Bioinformatics, 2017, 33, 1765-1772.	4.1	12
50	A Semiparametric Model for VQTL Mapping. Biometrics, 2017, 73, 571-581.	1.4	12
51	A signal detection method for temporal variation of adverse effect with vaccine adverse event reporting system data. BMC Medical Informatics and Decision Making, 2017, 17, 76.	3.0	12
52	The performance of rapid plasma reagin (RPR) titer in HIV-negative general paresis after neurosyphilis therapy. BMC Infectious Diseases, 2018, 18, 144.	2.9	12
53	Bayesian Mixed Treatment Comparisons Meta-Analysis for Correlated Outcomes Subject to Reporting Bias. Journal of the Royal Statistical Society Series C: Applied Statistics, 2018, 67, 127-144.	1.0	12
54	Serum antinuclear antibodies associate with worse prognosis in AQP4‐positive neuromyelitis optica spectrum disorder. Brain and Behavior, 2021, 11, e01865.	2.2	12

#	ARTICLE	IF	CITATIONS
55	Maximum likelihood estimation and EM algorithm of Copas-like selection model for publication bias correction. Biostatistics, 2017, 18, 495-504.	1.5	11
56	An improved method for bivariate meta‐analysis when within‐study correlations are unknown. Research Synthesis Methods, 2018, 9, 73-88.	8.7	11
57	Privacy-preserving harmonization via distributed ComBat. NeuroImage, 2022, 248, 118822.	4.2	11
58	Controversy and Debate: Questionable utility of the relative risk in clinical research: Paper 2: Is the Odds Ratio "portable" in meta-analysis? Time to consider bivariate generalized linear mixed model. Journal of Clinical Epidemiology, 2022, 142, 280-287.	5.0	10
59	Controversy and Debate : Questionable utility of the relative risk in clinical research: Paper 4 :Odds Ratios are far from "portable"â€" A call to use realistic models for effect variation in meta-analysis. Journal of Clinical Epidemiology, 2022, 142, 294-304.	5.0	10
60	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
61	A regression framework to uncover pleiotropy in large-scale electronic health record data. Journal of the American Medical Informatics Association: JAMIA, 2019, 26, 1083-1090.	4.4	9
62	Blood pressure and body fat percent in women with NMOSD. Brain and Behavior, 2019, 9, e01350.	2.2	9
63	An augmented estimation procedure for EHR-based association studies accounting for differential misclassification. Journal of the American Medical Informatics Association: JAMIA, 2020, 27, 244-253.	4.4	9
64	Reducing Bias Due to Outcome Misclassification for Epidemiologic Studies Using EHR-derived Probabilistic Phenotypes. Epidemiology, 2020, Publish Ahead of Print, 542-550.	2.7	9
65	Predictors of postinfectious inflammatory response syndrome in HIV-negative immunocompetent cryptococcal meningitis. Journal of Neurology, Neurosurgery and Psychiatry, 2021, 92, 680-681.	1.9	9
66	An efficient and accurate distributed learning algorithm for modeling multi-site zero-inflated count outcomes. Scientific Reports, 2021, 11, 19647.	3.3	9
67	Analysis of Dual Combination Therapies Used in Treatment of Hypertension in a Multinational Cohort. JAMA Network Open, 2022, 5, e223877.	5.9	9
68	DLMM as a lossless one-shot algorithm for collaborative multi-site distributed linear mixed models. Nature Communications, 2022, 13, 1678.	12.8	9
69	ODACH: a one-shot distributed algorithm for Cox model with heterogeneous multi-center data. Scientific Reports, 2022, 12, 6627.	3.3	9
70	Distributed Quasi-Poisson regression algorithm for modeling multi-site count outcomes in distributed data networks. Journal of Biomedical Informatics, 2022, 131, 104097.	4.3	9
71	Federated Multi-view Learning for Private Medical Data Integration and Analysis. ACM Transactions on Intelligent Systems and Technology, 2022, 13, 1-23.	4.5	9
72	Analysis of Individual Differences in Vaccine Pharmacovigilance Using VAERS Data and MedDRA System Organ Classes: A Use Case Study With Trivalent Influenza Vaccine. Biomedical Informatics Insights, 2017, 9, 117822261770062.	4.6	8

#	Article	IF	Citations
73	The Galaxy Plot: A New Visualization Tool for Bivariate Meta-Analysis Studies. American Journal of Epidemiology, 2020, 189, 861-869.	3.4	8
74	Studying pediatric health outcomes with electronic health records using Bayesian clustering and trajectory analysis. Journal of Biomedical Informatics, 2021, 113, 103654.	4.3	8
75	Accounting for publication bias using a bivariate trim and fill meta-analysis procedure. Statistics in Medicine, 2022, 41, 3466-3478.	1.6	8
76	Distributed learning for heterogeneous clinical data with application to integrating COVID-19 data across 230 sites. Npj Digital Medicine, 2022, 5, .	10.9	8
77	Inference for correlated effect sizes using multiple univariate meta-analyses. Statistics in Medicine, 2016, 35, 1405-1422.	1.6	7
78	PLEMT: A Novel Pseudolikelihood-Based EM Test for Homogeneity in Generalized Exponential Tilt Mixture Models. Journal of the American Statistical Association, 2017, 112, 1393-1404.	3.1	7
79	A simple and robust method for multivariate meta-analysis of diagnostic test accuracy. Statistics in Medicine, 2017, 36, 105-121.	1.6	7
80	Comparing drug safety of hepatitis C therapies using post-market data. BMC Medical Informatics and Decision Making, 2019, 19, 147.	3.0	7
81	Determining the Association Between End-of-Life Care Resources and Patient Outcomes in Pennsylvania ICUs*. Critical Care Medicine, 2019, 47, 1591-1598.	0.9	7
82	Regression analysis of longitudinal data with outcome-dependent sampling and informative censoring. Scandinavian Journal of Statistics, 2019, 46, 831-847.	1.4	7
83	Risk of Persistent Opioid Use following Major Surgery in Matched Samples of Patients with and without Cancer. Cancer Epidemiology Biomarkers and Prevention, 2020, 29, 2126-2133.	2.5	7
84	Prediction of post-vaccination Guillain-Barré syndrome using data from a passive surveillance system. Pharmacoepidemiology and Drug Safety, 2021, 30, 602-609.	1.9	7
85	ODAL: A one-shot distributed algorithm to perform logistic regressions on electronic health records data from multiple clinical sites. Pacific Symposium on Biocomputing Pacific Symposium on Biocomputing, 2019, 24, 30-41.	0.7	7
86	Robust-ODAL: Learning from heterogeneous health systems without sharing patient-level data. Pacific Symposium on Biocomputing Pacific Symposium on Biocomputing, 2020, 25, 695-706.	0.7	7
87	Methodological challenges in spatial and contextual exposome-health studies. Critical Reviews in Environmental Science and Technology, 2023, 53, 827-846.	12.8	7
88	A Class of Pseudolikelihood Ratio Tests for Homogeneity in Exponential Tilt Mixture Models. Scandinavian Journal of Statistics, 2015, 42, 504-517.	1.4	6
89	Plasma Homocysteine Level Is Associated with the Expanded Disability Status Scale in Neuromyelitis Optica Spectrum Disorder. NeuroImmunoModulation, 2019, 26, 258-264.	1.8	6
90	Elevated Plasma Homocysteine Levels in Anti-N-methyl-D-aspartate Receptor Encephalitis. Frontiers in Neurology, 2019, 10, 464.	2.4	6

#	ARTICLE	IF	CITATIONS
91	Impact of Hospitalization and Medication Switching on Post–discharge Adherence to Oral Anticoagulants in Patients With Atrial Fibrillation. Pharmacotherapy, 2020, 40, 1022-1035.	2.6	6
92	National Survey of Hospitalists' Experiences with Incidental Pulmonary Nodules. Journal of Hospital Medicine, 2019, 14, 353-356.	1.4	6
93	Copas-like selection model to correct publication bias in systematic review of diagnostic test studies. Statistical Methods in Medical Research, 2019, 28, 2912-2923.	1.5	5
94	Global identifiability of latent class models with applications to diagnostic test accuracy studies: A Gröbner basis approach. Biometrics, 2020, 76, 98-108.	1.4	5
95	A cost-effective chart review sampling design to account for phenotyping error in electronic health records (EHR) data. Journal of the American Medical Informatics Association: JAMIA, 2021, 29, 52-61.	4.4	5
96	Using logic regression to characterize extreme heat exposures and their health associations: a time-series study of emergency department visits in Atlanta. BMC Medical Research Methodology, 2021, 21, 87.	3.1	5
97	Small-study effects: current practice and challenges for future research. Statistics and Its Interface, 2020, 13, 475-484.	0.3	5
98	Robust-ODAL: Learning from heterogeneous health systems without sharing patient-level data. , 2019, , .		5
99	The association of prescription opioid use with suicide attempts: An analysis of statewide medical claims data. PLoS ONE, 2022, 17, e0269809.	2.5	5
100	Implementing optimal allocation in clinical trials with multiple endpoints. Journal of Statistical Planning and Inference, 2017, 182, 88-99.	0.6	4
101	A conditional composite likelihood ratio test with boundary constraints. Biometrika, 2018, 105, 225-232.	2.4	4
102	Embracing study heterogeneity for finding genetic interactions in large–scale research consortia. Genetic Epidemiology, 2020, 44, 52-66.	1.3	4
103	A fast score test for generalized mixture models. Biometrics, 2020, 76, 811-820.	1.4	4
104	How Computational Experiments Can Improve Our Understanding of the Genetic Architecture of Common Human Diseases. Artificial Life, 2020, 26, 23-37.	1.3	4
105	Comparing the power of family-based association tests for sequence data with applications in the GAW18 simulated data. BMC Proceedings, 2014, 8, S27.	1.6	3
106	Rapid network meta-analysis using data from Food and Drug Administration approval packages is feasible but with limitations. Journal of Clinical Epidemiology, 2019, 114, 84-94.	5.0	3
107	Investigating safety profiles of human papillomavirus vaccine across group differences using VAERS data and MedDRA. PeerJ, 2019, 7, e7490.	2.0	3
108	Automated discovery of test statistics using genetic programming. Genetic Programming and Evolvable Machines, 2019, 20, 127-137.	2.2	3

#	Article	IF	Citations
109	EMBRACE: An EM‐based bias reduction approach through Copas‐model estimation for quantifying the evidence of selective publishing in network meta‐analysis. Biometrics, 2022, 78, 754-765.	1.4	3
110	Monitoring vaccine safety by studying temporal variation of adverse events using vaccine adverse event reporting system. Annals of Applied Statistics, 2021, 15, .	1.1	3
111	Estimating the reference interval from a fixed effects <scp>meta‐analysis</scp>. Research Synthesis Methods, 2021, 12, 630-640.	8.7	3
112	A meta‐analytic framework for detection of genetic interactions. Genetic Epidemiology, 2016, 40, 534-543.	1.3	2
113	On meta‐and mega‐analyses for gene–environment interactions. Genetic Epidemiology, 2017, 41, 876-886.	1.3	2
114	Two‐sample test for correlated data under outcome‐dependent sampling with an application to self‐reported weight loss data. Statistics in Medicine, 2019, 38, 4999-5009.	1.6	2
115	Semiparametric modelling and estimation of covariate‐adjusted dependence between bivariate recurrent events. Biometrics, 2020, 76, 1229-1239.	1.4	2
116	Lossless integration of multiple electronic health records for identifying pleiotropy using summary statistics. Nature Communications, 2021, 12, 168.	12.8	2
117	Comparing the Human Papillomavirus Vaccination Opinions Trends from Different Twitter User Groups with a Machine Learning Based System and Semiparametric Nonlinear Regression. Studies in Health Technology and Informatics, 2017, 245, 1218.	0.3	2
118	SAT: a Surrogate-Assisted Two-wave case boosting sampling method, with application to EHR-based association studies. Journal of the American Medical Informatics Association: JAMIA, 2022, 29, 918-927.	4.4	2
119	A frailty model for recurrent events during alternating restraint and non-restraint time periods. Statistics in Medicine, 2017, 36, 643-654.	1.6	1
120	On specification tests for composite likelihood inference. Biometrika, 2020, 107, 907-917.	2.4	1
121	Evaluation of Phenotyping Errors on Polygenic Risk Score Predictions. , 2020, , .		1
122	Does hospitalization for thromboembolism improve oral anticoagulant adherence in patients with atrial fibrillation?. Journal of the American Pharmacists Association: JAPhA, 2020, 60, 986-992.e2.	1.5	1
123	Identifying Clinical Risk Factors for Opioid Use Disorder using a Distributed Algorithm to Combine Real-World Data from a Large Clinical Data Research Network. AMIA ... Annual Symposium proceedings, 2020, 2020, 1220-1229.	0.2	1
124	Leverage Real-world Longitudinal Data in Large Clinical Research Networks for Alzheimer's Disease and Related Dementia (ADRD). AMIA ... Annual Symposium proceedings, 2020, 2020, 393-401.	0.2	1
125	A pseudolikelihood approach for assessing genetic association in case–control studies with unmeasured population structure. Statistical Methods in Medical Research, 2020, 29, 3153-3165.	1.5	0
126	Back Cover Image. Genetic Epidemiology, 2020, 44, ii.	1.3	0

#	Article	IF	Citations
127	Accounting for post‑randomization variables in meta‑analysis: A joint meta‑regression approach. Biometrics, 2023, 79, 358-367.	1.4	0
128	PSB 2019 Workshop on Text Mining and Visualization for Precision Medicine. , 2018, , .		0
129	Inherited Thrombophilia and the Risk of Arterial Ischemic Stroke: A Systematic Review and Meta-Analysis. Blood, 2018, 132, 2518-2518.	1.4	0
130	Identification of Rare Adverse Events with Year-varying Reporting Rates for FLU4 Vaccine in VAERS. AMIA ... Annual Symposium proceedings, 2018, 2018, 1544-1551.	0.2	0
131	Automated discovery of test statistics using genetic programming. Genetic Programming and Evolvable Machines, 2019, 20, 127-137.	2.2	0
132	Extending Hui‑Walter framework to correlated outcomes with application to diagnosis tests of an eye disease among premature infants. Statistics in Medicine, 2022, 41, 433-448.	1.6	0
133	A Bayesian Network to Predict the Risk of Post Influenza Vaccination Guillain-Barré Syndrome: Development and Validation Study. JMIR Public Health and Surveillance, 2022, 8, e25658.	2.6	0
134	Multisite learning of high-dimensional heterogeneous data with applications to opioid use disorder study of 15,000 patients across 5 clinical sites. Scientific Reports, 2022, 12, .	3.3	0