

Atul Butte

List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/5952749/publications.pdf>

Version: 2024-02-01

120
papers

19,316
citations

53794

45
h-index

21540

114
g-index

142
all docs

142
docs citations

142
times ranked

33119
citing authors

#	ARTICLE	IF	CITATIONS
1	xCell: digitally portraying the tissue cellular heterogeneity landscape. <i>Genome Biology</i> , 2017, 18, 220.	8.8	2,572
2	Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. <i>Nature Immunology</i> , 2019, 20, 163-172.	14.5	2,330
3	Scalable and accurate deep learning with electronic health records. <i>Npj Digital Medicine</i> , 2018, 1, 18.	10.9	1,440
4	Personal Omics Profiling Reveals Dynamic Molecular and Medical Phenotypes. <i>Cell</i> , 2012, 148, 1293-1307.	28.9	1,134
5	Systematic pan-cancer analysis of tumour purity. <i>Nature Communications</i> , 2015, 6, 8971.	12.8	937
6	Variation in the Human Immune System Is Largely Driven by Non-Heritable Influences. <i>Cell</i> , 2015, 160, 37-47.	28.9	828
7	ImmPort: disseminating data to the public for the future of immunology. <i>Immunologic Research</i> , 2014, 58, 234-239.	2.9	724
8	Discovery and Preclinical Validation of Drug Indications Using Compendia of Public Gene Expression Data. <i>Science Translational Medicine</i> , 2011, 3, 96ra77.	12.4	708
9	Clinical assessment incorporating a personal genome. <i>Lancet, The</i> , 2010, 375, 1525-1535.	13.7	637
10	Computational Repositioning of the Anticonvulsant Topiramate for Inflammatory Bowel Disease. <i>Science Translational Medicine</i> , 2011, 3, 96ra76.	12.4	534
11	ImmPort, toward repurposing of open access immunological assay data for translational and clinical research. <i>Scientific Data</i> , 2018, 5, 180015.	5.3	529
12	An Environment-Wide Association Study (EWAS) on Type 2 Diabetes Mellitus. <i>PLoS ONE</i> , 2010, 5, e10746.	2.5	470
13	Comprehensive analysis of normal adjacent to tumor transcriptomes. <i>Nature Communications</i> , 2017, 8, 1077.	12.8	394
14	Systematic survey reveals general applicability of "guilt-by-association" within gene coexpression networks. <i>BMC Bioinformatics</i> , 2005, 6, 227.	2.6	370
15	A Drug Repositioning Approach Identifies Tricyclic Antidepressants as Inhibitors of Small Cell Lung Cancer and Other Neuroendocrine Tumors. <i>Cancer Discovery</i> , 2013, 3, 1364-1377.	9.4	366
16	A longitudinal big data approach for precision health. <i>Nature Medicine</i> , 2019, 25, 792-804.	30.7	329
17	Quantifying the relationship between co-expression, co-regulation and gene function. <i>BMC Bioinformatics</i> , 2004, 5, 18.	2.6	313
18	Minimum information about clinical artificial intelligence modeling: the MI-CLAIM checklist. <i>Nature Medicine</i> , 2020, 26, 1320-1324.	30.7	262

#	ARTICLE	IF	CITATIONS
19	Mutations in NGLY1 cause an inherited disorder of the endoplasmic reticulum-associated degradation pathway. <i>Genetics in Medicine</i> , 2014, 16, 751-758.	2.4	191
20	A call for deep-learning healthcare. <i>Nature Medicine</i> , 2019, 25, 14-15.	30.7	161
21	Cross-tissue Analysis of Gene and Protein Expression in Normal and Cancer Tissues. <i>Scientific Reports</i> , 2016, 6, 24799.	3.3	155
22	Autoimmune Disease Classification by Inverse Association with SNP Alleles. <i>PLoS Genetics</i> , 2009, 5, e1000792.	3.5	151
23	Reversal of cancer gene expression correlates with drug efficacy and reveals therapeutic targets. <i>Nature Communications</i> , 2017, 8, 16022.	12.8	151
24	Leveraging big data to transform target selection and drug discovery. <i>Clinical Pharmacology and Therapeutics</i> , 2016, 99, 285-297.	4.7	147
25	Evidence for benefit of statins to modify cognitive decline and risk in Alzheimer's disease. <i>Alzheimer's Research and Therapy</i> , 2017, 9, 10.	6.2	145
26	Assessment of a Deep Learning Model Based on Electronic Health Record Data to Forecast Clinical Outcomes in Patients With Rheumatoid Arthritis. <i>JAMA Network Open</i> , 2019, 2, e190606.	5.9	135
27	A Meta-analysis of Lung Cancer Gene Expression Identifies <i>PTK7</i> as a Survival Gene in Lung Adenocarcinoma. <i>Cancer Research</i> , 2014, 74, 2892-2902.	0.9	131
28	Robust meta-analysis of gene expression using the elastic net. <i>Nucleic Acids Research</i> , 2015, 43, e79-e79.	14.5	124
29	A Nutrient-Wide Association Study on Blood Pressure. <i>Circulation</i> , 2012, 126, 2456-2464.	1.6	122
30	Prototype of running clinical trials in an untrustworthy environment using blockchain. <i>Nature Communications</i> , 2019, 10, 917.	12.8	114
31	Comprehensive transcriptomic analysis of cell lines as models of primary tumors across 22 tumor types. <i>Nature Communications</i> , 2019, 10, 3574.	12.8	111
32	Disease signatures are robust across tissues and experiments. <i>Molecular Systems Biology</i> , 2009, 5, 307.	7.2	101
33	Cross-Species Functional Analysis of Cancer-Associated Fibroblasts Identifies a Critical Role for CLCF1 and IL-6 in Non-Small Cell Lung Cancer <i>In Vivo</i> . <i>Cancer Research</i> , 2012, 72, 5744-5756.	0.9	96
34	PDX-MI: Minimal Information for Patient-Derived Tumor Xenograft Models. <i>Cancer Research</i> , 2017, 77, e62-e66.	0.9	92
35	Opportunities and challenges in using real-world data for health care. <i>Journal of Clinical Investigation</i> , 2020, 130, 565-574.	8.2	88
36	Widespread parainflammation in human cancer. <i>Genome Biology</i> , 2016, 17, 145.	8.8	87

#	ARTICLE	IF	CITATIONS
37	Characteristics and challenges of the clinical pipeline of digital therapeutics. <i>Npj Digital Medicine</i> , 2020, 3, 159.	10.9	81
38	ZeitZeiger: supervised learning for high-dimensional data from an oscillatory system. <i>Nucleic Acids Research</i> , 2016, 44, e80-e80.	14.5	76
39	Better medicine through machine learning: What's real, and what's artificial?. <i>PLoS Medicine</i> , 2018, 15, e1002721.	8.4	73
40	Identification of antiviral antihistamines for COVID-19 repurposing. <i>Biochemical and Biophysical Research Communications</i> , 2021, 538, 173-179.	2.1	73
41	Differentially Expressed RNA from Public Microarray Data Identifies Serum Protein Biomarkers for Cross-Organ Transplant Rejection and Other Conditions. <i>PLoS Computational Biology</i> , 2010, 6, e1000940.	3.2	72
42	Drug Discovery in a Multidimensional World: Systems, Patterns, and Networks. <i>Journal of Cardiovascular Translational Research</i> , 2010, 3, 438-447.	2.4	59
43	Integrating multiple "omics" analyses identifies serological protein biomarkers for preeclampsia. <i>BMC Medicine</i> , 2013, 11, 236.	5.5	58
44	Relating hepatocellular carcinoma tumor samples and cell lines using gene expression data in translational research. <i>BMC Medical Genomics</i> , 2015, 8, S5.	1.5	56
45	Integrating biomedical research and electronic health records to create knowledge-based biologically meaningful machine-readable embeddings. <i>Nature Communications</i> , 2019, 10, 3045.	12.8	54
46	Use of electronic health records to support a public health response to the COVID-19 pandemic in the United States: a perspective from 15 academic medical centers. <i>Journal of the American Medical Informatics Association: JAMIA</i> , 2021, 28, 393-401.	4.4	54
47	Peptidomic Identification of Serum Peptides Diagnosing Preeclampsia. <i>PLoS ONE</i> , 2013, 8, e65571.	2.5	52
48	MetaCyto: A Tool for Automated Meta-analysis of Mass and Flow Cytometry Data. <i>Cell Reports</i> , 2018, 24, 1377-1388.	6.4	52
49	Combined inhibition of atypical PKC and histone deacetylase 1 is cooperative in basal cell carcinoma treatment. <i>JCI Insight</i> , 2017, 2, .	5.0	49
50	ProfileChaser: searching microarray repositories based on genome-wide patterns of differential expression. <i>Bioinformatics</i> , 2011, 27, 3317-3318.	4.1	47
51	Relating Chemical Structure to Cellular Response: An Integrative Analysis of Gene Expression, Bioactivity, and Structural Data Across 11,000 Compounds. <i>CPT: Pharmacometrics and Systems Pharmacology</i> , 2015, 4, 576-584.	2.5	47
52	Corticosteroid use is not associated with improved outcomes in acute exacerbation of IPF. <i>Respirology</i> , 2020, 25, 629-635.	2.3	47
53	Meta-Analysis of Vaginal Microbiome Data Provides New Insights Into Preterm Birth. <i>Frontiers in Microbiology</i> , 2020, 11, 476.	3.5	47
54	Precision annotation of digital samples in NCBI's gene expression omnibus. <i>Scientific Data</i> , 2017, 4, 170125.	5.3	44

#	ARTICLE	IF	CITATIONS
55	Differential Phasing between Circadian Clocks in the Brain and Peripheral Organs in Humans. <i>Journal of Biological Rhythms</i> , 2016, 31, 588-597.	2.6	42
56	Time for NIH to lead on data sharing. <i>Science</i> , 2020, 367, 1308-1309.	12.6	42
57	The 10,000 Immunomes Project: Building a Resource for Human Immunology. <i>Cell Reports</i> , 2018, 25, 513-522.e3.	6.4	40
58	A robust and interpretable end-to-end deep learning model for cytometry data. <i>Proceedings of the National Academy of Sciences of the United States of America</i> , 2020, 117, 21373-21380.	7.1	40
59	Opening clinical trial data: are the voluntary data-sharing portals enough?. <i>BMC Medicine</i> , 2015, 13, 280.	5.5	38
60	Protected Health Information filter (Philter): accurately and securely de-identifying free-text clinical notes. <i>Npj Digital Medicine</i> , 2020, 3, 57.	10.9	38
61	A genome-wide association study identifies only two ancestry specific variants associated with spontaneous preterm birth. <i>Scientific Reports</i> , 2018, 8, 226.	3.3	37
62	Heterogeneity in HIV and cellular transcription profiles in cell line models of latent and productive infection: implications for HIV latency. <i>Retrovirology</i> , 2019, 16, 32.	2.0	35
63	<i>in silico</i> and <i>in vitro</i> drug screening identifies new therapeutic approaches for Ewing sarcoma. <i>Oncotarget</i> , 2017, 8, 4079-4095.	1.8	34
64	Postoperative delirium prediction using machine learning models and preoperative electronic health record data. <i>BMC Anesthesiology</i> , 2022, 22, 8.	1.8	30
65	Application of Machine Learning for Cytometry Data. <i>Frontiers in Immunology</i> , 2021, 12, 787574.	4.8	30
66	Likelihood ratios for genome medicine. <i>Genome Medicine</i> , 2010, 2, 30.	8.2	27
67	Constraints on Biological Mechanism from Disease Comorbidity Using Electronic Medical Records and Database of Genetic Variants. <i>PLoS Computational Biology</i> , 2016, 12, e1004885.	3.2	27
68	Protein Microarrays Discover Angiotensinogen and PRKRIP1 as Novel Targets for Autoantibodies in Chronic Renal Disease. <i>Molecular and Cellular Proteomics</i> , 2011, 10, M110.000497.	3.8	26
69	Robust prediction of clinical outcomes using cytometry data. <i>Bioinformatics</i> , 2019, 35, 1197-1203.	4.1	25
70	Comparing Ethnicity-Specific Reference Intervals for Clinical Laboratory Tests from EHR Data. <i>Journal of Applied Laboratory Medicine</i> , 2018, 3, 366-377.	1.3	24
71	Machine Learning Methods in Health Economics and Outcomes Research—The PALISADE Checklist: A Good Practices Report of an ISPOR Task Force. <i>Value in Health</i> , 2022, 25, 1063-1080.	0.3	24
72	Reanalysis of the Rituximab in ANCA-Associated Vasculitis trial identifies granulocyte subsets as a novel early marker of successful treatment. <i>Arthritis Research and Therapy</i> , 2015, 17, 262.	3.5	23

#	ARTICLE	IF	CITATIONS
73	Embedding electronic health records onto a knowledge network recognizes prodromal features of multiple sclerosis and predicts diagnosis. <i>Journal of the American Medical Informatics Association: JAMIA</i> , 2022, 29, 424-434.	4.4	22
74	Translational Bioinformatics: Data-driven Drug Discovery and Development. <i>Clinical Pharmacology and Therapeutics</i> , 2012, 91, 949-952.	4.7	20
75	Whole-genome sequencing of Atacama skeleton shows novel mutations linked with dysplasia. <i>Genome Research</i> , 2018, 28, 423-431.	5.5	19
76	Androgen-deprivation therapy and SARS-CoV-2 in men with prostate cancer: findings from the University of California Health System registry. <i>Annals of Oncology</i> , 2021, 32, 678-679.	1.2	19
77	Predicting Inpatient Medication Orders From Electronic Health Record Data. <i>Clinical Pharmacology and Therapeutics</i> , 2020, 108, 145-154.	4.7	18
78	Human splice factors contribute to latent HIV infection in primary cell models and blood CD4+ T cells from ART-treated individuals. <i>PLoS Pathogens</i> , 2020, 16, e1009060.	4.7	18
79	Computational prediction and experimental validation associating FABP-1 and pancreatic adenocarcinoma with diabetes. <i>BMC Gastroenterology</i> , 2011, 11, 5.	2.0	16
80	Big data opens a window onto wellness. <i>Nature Biotechnology</i> , 2017, 35, 720-721.	17.5	16
81	Open challenges in developing digital therapeutics in the United States. , 2022, 1, e0000008.		16
82	Novel integration of hospital electronic medical records and gene expression measurements to identify genetic markers of maturation. <i>Pacific Symposium on Biocomputing Pacific Symposium on Biocomputing</i> , 2008, , 243-54.	0.7	15
83	Heterogeneity of Diabetes: Î²-Cells, Phenotypes, and Precision Medicine: Proceedings of an International Symposium of the Canadian Institutes of Health Researchâ€™s Institute of Nutrition, Metabolism and Diabetes and the U.S. National Institutes of Healthâ€™s National Institute of Diabetes and Digestive and Kidney Diseases. <i>Diabetes Care</i> , 2022, 45, 3-22.	8.6	14
84	Comparison of automated and human assignment of MeSH terms on publicly-available molecular datasets. <i>Journal of Biomedical Informatics</i> , 2011, 44, S39-S43.	4.3	13
85	Making it personal: translational bioinformatics. <i>Journal of the American Medical Informatics Association: JAMIA</i> , 2013, 20, 595-596.	4.4	13
86	A patient-level data meta-analysis of standard-of-care treatments from eight prostate cancer clinical trials. <i>Scientific Data</i> , 2016, 3, 160027.	5.3	12
87	Age- and Sex-Associated Variations in the Sensitivity of Serological Tests Among Individuals Infected With SARS-CoV-2. <i>JAMA Network Open</i> , 2021, 4, e210337.	5.9	12
88	Deep learning from multiple experts improves identification of amyloid neuropathologies. <i>Acta Neuropathologica Communications</i> , 2022, 10, 66.	5.2	12
89	Risky Business: Meeting the Structural Needs of Transdisciplinary Science. <i>Journal of Pediatrics</i> , 2017, 191, 255-258.	1.8	11
90	ROMOP: a light-weight R package for interfacing with OMOP-formatted electronic health record data. <i>JAMIA Open</i> , 2019, 2, 10-14.	2.0	11

#	ARTICLE	IF	CITATIONS
91	CovidCounties is an interactive real time tracker of the COVID19 pandemic at the level of US counties. Scientific Data, 2020, 7, 405.	5.3	11
92	Tracing diagnosis trajectories over millions of patients reveal an unexpected risk in schizophrenia. Scientific Data, 2019, 6, 201.	5.3	10
93	Assessment of Postdonation Outcomes in US Living Kidney Donors Using Publicly Available Data Sets. JAMA Network Open, 2019, 2, e191851.	5.9	10
94	Closing the Evidence Gap in Interstitial Lung Disease. The Promise of Real-World Data. American Journal of Respiratory and Critical Care Medicine, 2019, 199, 1061-1065.	5.6	10
95	Knowledge Network Embedding of Transcriptomic Data from Spaceflown Mice Uncovers Signs and Symptoms Associated with Terrestrial Diseases. Life, 2021, 11, 42.	2.4	10
96	Big Data in Nephrology. Nature Reviews Nephrology, 2021, 17, 676-687.	9.6	10
97	Accuracy of medical billing data against the electronic health record in the measurement of colorectal cancer screening rates. BMJ Open Quality, 2020, 9, e000856.	1.1	9
98	Quantifying Variation in Treatment Utilization for Type 2 Diabetes Across Five Major University of California Health Systems. Diabetes Care, 2021, 44, 908-914.	8.6	9
99	Enabling precision medicine in neonatology, an integrated repository for preterm birth research. Scientific Data, 2018, 5, 180219.	5.3	9
100	Opal: an implementation science tool for machine learning clinical decision support in anesthesia. Journal of Clinical Monitoring and Computing, 2022, 36, 1367-1377.	1.6	9
101	Trans-channel fluorescence learning improves high-content screening for Alzheimer's disease therapeutics. Nature Machine Intelligence, 2022, 4, 583-595.	16.0	9
102	RImmPort: an R/Bioconductor package that enables ready-for-analysis immunology research data. Bioinformatics, 2017, 33, 1101-1103.	4.1	8
103	Explanatory Model of Dry Eye Disease Using Health and Nutrition Examinations: Machine Learning and Network-Based Factor Analysis From a National Survey. JMIR Medical Informatics, 2020, 8, e16153.	2.6	8
104	Integrating Clinical Phenotype and Gene Expression Data to Prioritize Novel Drug Uses. CPT: Pharmacometrics and Systems Pharmacology, 2016, 5, 599-607.	2.5	7
105	Open data informatics and data repurposing for IBD. Nature Reviews Gastroenterology and Hepatology, 2018, 15, 715-716.	17.8	7
106	Utility of routinely collected electronic health records data to support effectiveness evaluations in inflammatory bowel disease: a pilot study of tofacitinib. BMJ Health and Care Informatics, 2021, 28, e100337.	3.0	7
107	Development and Validation of a Deep Learning Strategy for Automated View Classification of Pediatric Focused Assessment With Sonography for Trauma. Journal of Ultrasound in Medicine, 2022, 41, 1915-1924.	1.7	7
108	Systematic identification of ACE2 expression modulators reveals cardiomyopathy as a risk factor for mortality in COVID-19 patients. Genome Biology, 2022, 23, 15.	8.8	7

#	ARTICLE	IF	CITATIONS
109	The Atacama skeleton. <i>Genome Research</i> , 2018, 28, 607-608.	5.5	6
110	Trials and Tribulations—11 Reasons Why We Need to Promote Clinical Trials Data Sharing. <i>JAMA Network Open</i> , 2021, 4, e2035043.	5.9	6
111	Rethinking PICO in the Machine Learning Era: ML-PICO. <i>Applied Clinical Informatics</i> , 2021, 12, 407-416.	1.7	6
112	Impact of Different Approaches to Preparing Notes for Analysis With Natural Language Processing on the Performance of Prediction Models in Intensive Care. , 2021, 3, e0450.		5
113	It takes a genome to understand a village: Population scale precision medicine. <i>Proceedings of the National Academy of Sciences of the United States of America</i> , 2016, 113, 12344-12346.	7.1	4
114	A pilot study showing a stronger H1N1 influenza vaccination response during pregnancy in women who subsequently deliver preterm. <i>Journal of Reproductive Immunology</i> , 2019, 132, 16-20.	1.9	3
115	Opportunities and Challenges in Democratizing Immunology Datasets. <i>Frontiers in Immunology</i> , 2021, 12, 647536.	4.8	2
116	Five-year pediatric use of a digital wearable fitness device: lessons from a pilot case study. <i>JAMIA Open</i> , 2021, 4, ooab054.	2.0	2
117	Abstract LB-006: Oncology model fidelity scores. , 2017, , .		2
118	MOLECULAR BIOINFORMATICS FOR DISEASE – SESSION INTRODUCTION. , 2008, , .		0
119	MOLECULAR BIOINFORMATICS FOR DISEASE: PROTEIN INTERACTIONS AND PHENOMICS – Session Introduction. , 2007, , .		0
120	Synergy Between Kit Ligand (KL) and IL-4 In Mast Cells Is Mediated by Cross-Receptor Interactions In Lipid Rafts.. <i>Blood</i> , 2010, 116, 1564-1564.	1.4	0