[Artificial Intelligence Transforms the Future of Health Care]

- **Creator:** Noorbakhsh-Sabet, Nariman ; Zand, Ramin ; Zhang, Yanfei ; Abedi, Vida
- **Subjects:** Abridged Index Medicus ; Artificial intelligence (AI) ; Artificial Intelligence - trends ; Decision Support Systems, Clinical - trends ; Delivery of Health Care - methods ; Delivery of Health Care - trends ; Drug Discovery - trends ; Epidemics - prevention & control ; Forecasting ; General & Internal Medicine ; Humans ; Integrated health care systems ; Life Sciences & Biomedicine ; Machine Learning ; Medical informatics ; Medicine, General & Internal ; Precision medicine ; Science & Technology ; Translational Medical Research - trends
- **Is Part Of:** The American journal of medicine, 2019-07, Vol.132 (7), p.795-801
- **Description:** Life sciences researchers using artificial intelligence (AI) are under pressure to innovate faster than ever. Large, multilevel, and integrated data sets offer the promise of unlocking novel insights and accelerating breakthroughs. Although more data are available than ever, only a fraction is being curated, integrated, understood, and analyzed. AI focuses on how computers learn from data and mimic human thought processes. AI increases learning capacity and provides decision support system at scales that are transforming the future of health care. This article is a review of applications for machine learning in health care with a focus on clinical, translational, and public health applications with an overview of the important role of privacy, data sharing, and genetic information.
- **Publisher:** NEW YORK: Elsevier Inc
- **Language:** English
- **Identifier:** ISSN: 0002-9343; EISSN: 1555-7162; DOI: 10.1016/j.amjmed.2019.01.017; PMID: 30710543
- **Source:** Scopus; ScienceDirect Journals (5 years ago - present); Web of Science - Science Citation Index Expanded - 2019 from WEB OF SCIENCE; Alma/SFX Local Collection

[The future is now? Clinical and translational aspects of “Omics” technologies]

- **Creator:** D’Adamo, Gemma L ; Widdop, James T ; Giles, Edward M
- **Subjects:** Artificial intelligence ; Artificial Intelligence - trends ; Big Data ; COVID-19 ; COVID-19 - epidemiology ; Ethics ; Genomics ; Genomics - methods ; Genomics - trends ; Humans ; Learning algorithms ; machine learning ; Medical Oncology - methods ; Medical Oncology - trends ; Medical research ; Metabolomics - methods ; Metabolomics - trends ; microbiome ; Pandemics ; Precision Medicine - methods ; Precision Medicine - trends ; Proteomics ; Proteomics - methods ;

Proteomics - trends ; Time Factors ; translational immunology ; Translational Medical Research - methods ; Translational Medical Research - trends
- **Is Part Of:** Immunology and cell biology, 2021-02, Vol.99 (2), p.168-176
- **Description:** Big data has become a central part of medical research, as well as modern life generally. “Omics” technologies include genomics, proteomics, microbiomics and increasingly other omics. These have been driven by rapid advances in laboratory techniques and equipment. Crucially, improved information handling capabilities have allowed concepts such as artificial intelligence and machine learning to enter the research world. The COVID−19 pandemic has shown how quickly information can be generated and analyzed using such approaches, but also showed its limitations. This review will look at how “omics” has begun to be translated into clinical practice. While there appears almost limitless potential in using big data for “precision” or “personalized” medicine, the reality is that this remains largely aspirational. Oncology is the only field of medicine that is widely adopting such technologies, and even in this field uptake is irregular. There are practical and ethical reasons for this lack of translation of increasingly affordable techniques into the clinic. Undoubtedly, there will be increasing use of large data sets from traditional (e.g. tumor samples, patient genomics) and nontraditional (e.g. smartphone) sources. It is perhaps the greatest challenge of the health−care sector over the coming decade to integrate these resources in an effective, practical and ethical way. Omics technologies have begun to dominate medical research as well as other aspects of modern life. Clinical use of these technologies remains limited. This review explores how big data is starting to be used in clinical practice and the hurdles that it will face in implementation.
- **Publisher:** United States: Blackwell Science Ltd
- **Language:** English
- **Identifier:** ISSN: 0818-9641; EISSN: 1440-1711; DOI: 10.1111/imcb.12404; PMID: 32924178
- **Source:** Scopus; Wiley Online Library All Journals

Strategic vision for improving human health at The Forefront of Genomics

- **Creator:** Green, Eric D ; Gunter, Chris ; Biesecker, Leslie G ; Di Francesco, Valentina ; Easter, Carla L ; Feingold, Elise A ; Felsenfeld, Adam L ; Kaufman, David J ; Ostrander, Elaine A ; Pavan, William J ; Phillippy, Adam M ; Wise, Anastasia L ; Dayal, Jyoti Gupta ; Kish, Britny J ; Mandich, Allison ; Wellington, Christopher R ; Wetterstrand, Kris A ; Bates, Sarah A ; Leja, Darryl ; Vasquez, Susan ; Gahl, William A ; Graham, Bettie J ; Kastner, Daniel L ; Liu, Paul ; Rodriguez, Laura Lyman ; Solomon, Benjamin D ; Bonham, Vence L ; Brody, Lawrence C ; Hutter, Carolyn M ; Manolio, Teri A

- **Subjects:** Biomedical Research - economics ; Biomedical Research - trends ; COVID-19 - genetics ; Genetic research ; Genome, Human - genetics ; Genomics ; Genomics - economics ; Genomics - trends ; Health aspects ; Human genome ; Humans ; National Human Genome Research Institute (U.S.) - economics ; Public Health - standards ; Research ; Social Change ; Translational Medical Research - economics ; Translational Medical Research - trends ; United States
- **Is Part Of:** Nature (London), 2020-10-29, Vol.586 (7831), p.683-692
- **Description:** Starting with the launch of the Human Genome Project three decades ago, and continuing after its completion in 2003, genomics has progressively come to have a central and catalytic role in basic and translational research. In addition, studies increasingly demonstrate how genomic information can be effectively used in clinical care. In the future, the anticipated advances in technology development, biological insights, and clinical applications (among others) will lead to more widespread integration of genomics into almost all areas of biomedical research, the adoption of genomics into mainstream medical and public-health practices, and an increasing relevance of genomics for everyday life. On behalf of the research community, the National Human Genome Research Institute recently completed a multi-year process of strategic engagement to identify future research priorities and opportunities in human genomics, with an emphasis on health applications. Here we describe the highest-priority elements envisioned for the cutting-edge of human genomics going forward-that is, at 'The Forefront of Genomics'.
- **Publisher:** England: Nature Publishing Group
- **Language:** English
- **Identifier:** ISSN: 0028-0836; EISSN: 1476-4687; DOI: 10.1038/s41586-020-2817-4; PMID: 33116284
- **Source:** Scopus; Nature Journals Online

Harnessing big ‘omics’ data and AI for drug discovery in hepatocellular carcinoma

- **Creator:** Chen, Bin ; Garmire, Lana ; Calvisi, Diego F ; Chua, Mei-Sze ; Kelley, Robin K ; Chen, Xin
- **Subjects:** Angiogenesis ; Animal models ; Animals ; Antineoplastic Agents - therapeutic use ; Artificial Intelligence ; Biomarkers ; Biomarkers - metabolism ; Carcinoma, Hepatocellular - drug therapy ; Carcinoma, Hepatocellular - genetics ; Cell culture ; Cell Line, Tumor ; Clinical trials ; Disease Models, Animal ; Drug discovery ; Drug Discovery - methods ; Gene expression ; Genomes ; Genomics - methods ; Hepatocellular carcinoma ; Humans ; Immune checkpoint ; Immunosuppressive agents ; Liver cancer ; Liver Neoplasms - drug therapy ; Liver Neoplasms - genetics ; Liver Neoplasms - secondary ; Machine Learning ; Molecular Targeted Therapy - methods ; Monoclonal antibodies ; Patients ; Pembrolizumab ;

Precision Medicine - methods ; Targeted cancer therapy ; Translational Medical Research - methods

- **Is Part Of:** Nature reviews. Gastroenterology & hepatology, 2020-04-01, Vol.17 (4), p.238-251
- **Description:** Hepatocellular carcinoma (HCC) is the most common form of primary adult liver cancer. After nearly a decade with sorafenib as the only approved treatment, multiple new agents have demonstrated efficacy in clinical trials, including the targeted therapies regorafenib, lenvatinib and cabozantinib, the anti-angiogenic antibody ramucirumab, and the immune checkpoint inhibitors nivolumab and pembrolizumab. Although these agents offer new promise to patients with HCC, the optimal choice and sequence of therapies remains unknown and without established biomarkers, and many patients do not respond to treatment. The advances and the decreasing costs of molecular measurement technologies enable profiling of HCC molecular features (such as genome, transcriptome, proteome and metabolome) at different levels, including bulk tissues, animal models and single cells. The release of such data sets to the public enhances the ability to search for information from these legacy studies and provides the opportunity to leverage them to understand HCC mechanisms, rationally develop new therapeutics and identify candidate biomarkers of treatment response. Here, we provide a comprehensive review of public data sets related to HCC and discuss how emerging artificial intelligence methods can be applied to identify new targets and drugs as well as to guide therapeutic choices for improved HCC treatment.
- **Publisher:** England: Nature Publishing Group
- **Language:** English
- **Identifier:** ISSN: 1759-5045; EISSN: 1759-5053; DOI: 10.1038/s41575-019-0240-9; PMID: 31900465
- **Source:** Gale Academic OneFile; Biological Science Collection; Gale OneFile: Health and Medicine; Scopus; Gale Health and Wellness; Gale OneFile: Nursing and Allied Health

China Brain Project: Basic Neuroscience, Brain Diseases, and Brain-Inspired Computing

- **Creator:** Poo, Mu-ming ; Du, Jiu-lin ; Ip, Nancy Y ; Xiong, Zhi-Qi ; Xu, Bo ; Tan, Tieniu
- **Subjects:** Animals ; artificial intelligence ; Behavioral Research - organization & administration ; Brain diseases ; Brain Diseases - diagnosis ; Brain Diseases - therapy ; brain disorders ; brain-inspired computing ; China ; Chinese medicine ; Cognition - physiology ; human cognition ; Humans ; Life Sciences & Biomedicine ; Man-Machine Systems ; Neural Pathways - physiology ; neurodegenerative diseases ; Neurosciences ; Neurosciences & Neurology ; Neurosciences - organization &

administration ; non-human primates ; Primates ; robotics ; Science & Technology ; Translational Medical Research - organization & administration
- **Is Part Of:** Neuron (Cambridge, Mass.), 2016-11-02, Vol.92 (3), p.591-596
- **Description:** The China Brain Project covers both basic research on neural mechanisms underlying cognition and translational research for the diagnosis and intervention of brain diseases as well as for brain-inspired intelligence technology. We discuss some emerging themes, with emphasis on unique aspects. The China Brain Project covers both basic research on neural mechanisms underlying cognition and translational research for the diagnosis and intervention of brain diseases as well as for brain-inspired intelligence technology.
- **Publisher:** CAMBRIDGE: Elsevier Inc
- **Language:** English
- **Identifier:** ISSN: 0896-6273; EISSN: 1097-4199; DOI: 10.1016/j.neuron.2016.10.050; PMID: 27809999
- **Source:** Biological Science Collection; Scopus; Cell Press Free Archives; ScienceDirect Journals (5 years ago - present); Web of Science - Science Citation Index Expanded - 2016 from WEB OF SCIENCE

Omics biomarker identification pipeline for translational medicine

- **Creator:** Bravo-Merodio, Laura ; Williams, John A ; Gkoutos, Georgios V ; Acharjee, Animesh
- **Subjects:** Algorithms ; Analysis ; Area Under Curve ; Biomarker ; Biomarkers - analysis ; Feature selection ; Genomics ; Humans ; Life Sciences & Biomedicine ; Lipids - analysis ; Machine learning ; Medicine, Research & Experimental ; Omics ; Regularization ; Research ; Research & Experimental Medicine ; Science & Technology ; Scientists ; Transcriptome - genetics ; Translational Medical Research ; Translational medicine
- **Is Part Of:** Journal of translational medicine, 2019-05-14, Vol.17 (1), p.155-155
- **Description:** BackgroundTranslational medicine (TM) is an emerging domain that aims to facilitate medical or biological advances efficiently from the scientist to the clinician. Central to the TM vision is to narrow the gap between basic science and applied science in terms of time, cost and early diagnosis of the disease state. Biomarker identification is one of the main challenges within TM. The identification of disease biomarkers from -omics data will not only help the stratification of diverse patient cohorts but will also provide early diagnostic information which could improve patient management and potentially prevent adverse outcomes. However, biomarker identification needs to be robust and reproducible. Hence a robust unbiased computational framework that can help clinicians identify those biomarkers is necessary.MethodsWe developed a pipeline (workflow) that includes two different

supervised classification techniques based on regularization methods to identify biomarkers from -omics or other high dimension clinical datasets. The pipeline includes several important steps such as quality control and stability of selected biomarkers. The process takes input files (outcome and independent variables or -omics data) and pre-processes (normalization, missing values) them. After a random division of samples into training and test sets, Least Absolute Shrinkage and Selection Operator and Elastic Net feature selection methods are applied to identify the most important features representing potential biomarker candidates. The penalization parameters are optimised using 10-fold cross validation and the process undergoes 100 iterations and a combinatorial analysis to select the best performing multivariate model. An empirical unbiased assessment of their quality as biomarkers for clinical use is performed through a Receiver Operating Characteristic curve and its Area Under the Curve analysis on both permuted and real data for 1000 different randomized training and test sets. We validated this pipeline against previously published biomarkers.ResultsWe applied this pipeline to three different datasets with previously published biomarkers: lipidomics data by Acharjee et al. (Metabolomics 13:25, 2017) and transcriptomics data by Rajamani and Bhasin (Genome Med 8:38, 2016) and Millset al. (Blood 114:1063-1072, 2009). Our results demonstrate that our method was able to identify both previously published biomarkers as well as new variables that add value to the published results.ConclusionsWe developed a robust pipeline to identify clinically relevant biomarkers that can be applied to different -omics datasets. Such identification reveals potentially novel drug targets and can be used as a part of a machine-learning based patient stratification framework in the translational medicine settings.

- **Publisher:** LONDON: BMC
- **Language:** English
- **Identifier:** ISSN: 1479-5876; EISSN: 1479-5876; DOI: 10.1186/s12967-019-1912-5; PMID: 31088492
- **Source:** Gale Academic OneFile; Gale OneFile: Health and Medicine; Scopus; Gale OneFile: Nursing and Allied Health; Publicly Available Content Database; BioMedCentral Open Access; Academic Search Ultimate; Web of Science - Science Citation Index Expanded - 2019 from WEB OF SCIENCE; PubMed Central; Alma/SFX Local Collection; DOAJ Directory of Open Access Journals - Not for CDI Discovery

Looking beyond the hype: Applied AI and machine learning in translational medicine

- **Creator:** Toh, Tzen S ; Dondelinger, Frank ; Wang, Dennis
- **Subjects:** Animals ; Artificial Intelligence ; Drug Discovery ; General & Internal Medicine ; Genomic medicine ; Genomics - methods ; Humans ; Imaging ; Life Sciences & Biomedicine ; Machine Learning ; Medicine, General & Internal ;

Medicine, Research & Experimental ; Molecular Imaging ; Precision Medicine ; Reproducibility of Results ; Research & Experimental Medicine ; Science & Technology ; Translational Medical Research - methods ; Translational medicine

- **Is Part Of:** EBioMedicine, 2019-09, Vol.47, p.607-615
- **Description:** Big data problems are becoming more prevalent for laboratory scientists who look to make clinical impact. A large part of this is due to increased computing power, in parallel with new technologies for high quality data generation. Both new and old techniques of artificial intelligence (AI) and machine learning (ML) can now help increase the success of translational studies in three areas: drug discovery, imaging, and genomic medicine. However, ML technologies do not come without their limitations and shortcomings. Current technical limitations and other limitations including governance, reproducibility, and interpretation will be discussed in this article. Overcoming these limitations will enable ML methods to be more powerful for discovery and reduce ambiguity within translational medicine, allowing data-informed decision-making to deliver the next generation of diagnostics and therapeutics to patients quicker, at lowered costs, and at scale.
- **Publisher:** AMSTERDAM: Elsevier B.V
- **Language:** English
- **Identifier:** ISSN: 2352-3964; EISSN: 2352-3964; DOI: 10.1016/j.ebiom.2019.08.027; PMID: 31466916
- **Source:** Scopus; ScienceDirect Journals (Transactional Access); Web of Science - Science Citation Index Expanded - 2019 from WEB OF SCIENCE; Elsevier:ScienceDirect:Open Access

Development of AI-based pathology biomarkers in gastrointestinal and liver cancer

- **Creator:** Kather, Jakob N ; Calderaro, Julien
- **Is Part Of:** Nature reviews. Gastroenterology & hepatology, 2020-07-03, Vol.17 (10), p.591-592
- **Language:** English
- **Identifier:** ISSN: 1759-5045; EISSN: 1759-5053; DOI: 10.1038/s41575-020-0343-3
- **Source:** Gale Academic OneFile; Biological Science Collection; Gale OneFile: Health and Medicine; Gale Health and Wellness; Gale OneFile: Nursing and Allied Health

Bridging a translational gap: Using machine learning to improve the prediction of PTSD

- **Creator:** Karstoft, Karen-Inge ; Galatzer-Levy, Isaac R ; Statnikov, Alexander ; Li, Zhiguo ; Shalev, Arieh Y ; Ankri, Yael ; Freedman, Sara ; Addesky, Rhonda ; Israeli-Shalev, Yossi ; Gilad, Moran ; Roitman, Pablo
- **Subjects:** Adaptation, Psychological - physiology ; Adult ; Algorithms ; Artificial Intelligence ; Early Diagnosis ; Early prediction ; Female ; Humans ; Life Sciences & Biomedicine ; Machine learning ; Male ; Markov boundary feature selection ; Middle Aged ; Post-traumatic stress disorder ; Posttraumatic Stress Disorder (PTSD) ; Prognosis ; Psychiatry ; Research ; Risk Assessment ; Risk Factors ; ROC Curve ; Science & Technology ; Stress Disorders, Post-Traumatic - diagnosis ; Stress Disorders, Post-Traumatic - etiology ; Stress Disorders, Post-Traumatic - physiopathology ; Stress Disorders, Post-Traumatic - prevention & control ; Support vector machines ; Translational Medical Research ; Wounds and Injuries - complications ; Wounds and Injuries - psychology
- **Is Part Of:** BMC psychiatry, 2015, Vol.15 (1), p.30-30
- **Description:** Background: Predicting Posttraumatic Stress Disorder (PTSD) is a pre-requisite for targeted prevention. Current research has identified group-level risk-indicators, many of which (e.g., head trauma, receiving opiates) concern but a subset of survivors. Identifying interchangeable sets of risk indicators may increase the efficiency of early risk assessment. The study goal is to use supervised machine learning (ML) to uncover interchangeable, maximally predictive combinations of early risk indicators. Methods: Data variables (features) reflecting event characteristics, emergency department (ED) records and early symptoms were collected in 957 trauma survivors within ten days of ED admission, and used to predict PTSD symptom trajectories during the following fifteen months. A Target Information Equivalence Algorithm (TIE*) identified all minimal sets of features (Markov Boundaries; MBs) that maximized the prediction of a non-remitting PTSD symptom trajectory when integrated in a support vector machine (SVM). The predictive accuracy of each set of predictors was evaluated in a repeated 10-fold cross-validation and expressed as average area under the Receiver Operating Characteristics curve (AUC) for all validation trials. Results: The average number of MBs per cross validation was 800. MBs' mean AUC was 0.75 (95% range: 0.67-0.80). The average number of features per MB was 18 (range: 12-32) with 13 features present in over 75% of the sets. Conclusions: Our findings support the hypothesized existence of multiple and interchangeable sets of risk indicators that equally and exhaustively predict non-remitting PTSD. ML's ability to increase prediction versatility is a promising step towards developing algorithmic, knowledge-based, personalized prediction of post-traumatic psychopathology.
- **Publisher:** LONDON: BMC
- **Language:** English
- **Identifier:** ISSN: 1471-244X; EISSN: 1471-244X; DOI: 10.1186/s12888-015-0399-8; PMID: 25886446

- **Source:** Gale Academic OneFile; Gale OneFile: Health and Medicine; Scopus; Medline Complete; Gale OneFile: Nursing and Allied Health; Publicly Available Content Database; Gale OneFile: Psychology; BioMedCentral Open Access; Academic Search Ultimate; Web of Science - Social Sciences Citation Index - 2015  ; Web of Science - Science Citation Index Expanded - 2015  ; PubMed Central; Alma/SFX Local Collection; DOAJ Directory of Open Access Journals - Not for CDI Discovery

Inter-species pathway perturbation prediction via data-driven detection of functional homology

- **Creator:** Hafemeister, Christoph ; Romero, Roberto ; Bilal, Erhan ; Meyer, Pablo ; Norel, Raquel ; Rhrissorrakrai, Kahn ; Bonneau, Richard ; Tarca, Adi L
- **Subjects:** Algorithms ; Animals ; Artificial Intelligence ; Bronchi - cytology ; Bronchi - metabolism ; Cells, Cultured ; Cytokines - metabolism ; Databases, Factual ; Epithelial Cells - cytology ; Epithelial Cells - metabolism ; Gene Expression Profiling - methods ; Gene Expression Regulation ; Humans ; Improver Challenge Special Issue; Species Translation Challenge ; Models, Animal ; Oligonucleotide Array Sequence Analysis ; Phosphoproteins - metabolism ; Phosphorylation ; Rats ; Signal Transduction ; Software ; Species Specificity ; Systems Biology - methods ; Translational Medical Research
- **Is Part Of:** Bioinformatics, 2015-02-15, Vol.31 (4), p.501-508
- **Description:** Experiments in animal models are often conducted to infer how humans will respond to stimuli by assuming that the same biological pathways will be affected in both organisms. The limitations of this assumption were tested in the IMPROVER Species Translation Challenge, where 52 stimuli were applied to both human and rat cells and perturbed pathways were identified. In the Inter-species Pathway Perturbation Prediction sub-challenge, multiple teams proposed methods to use rat transcription data from 26 stimuli to predict human gene set and pathway activity under the same perturbations. Submissions were evaluated using three performance metrics on data from the remaining 26 stimuli. We present two approaches, ranked second in this challenge, that do not rely on sequence-based orthology between rat and human genes to translate pathway perturbation state but instead identify transcriptional response orthologs across a set of training conditions. The translation from rat to human accomplished by these so-called direct methods is not dependent on the particular analysis method used to identify perturbed gene sets. In contrast, machine learning-based methods require performing a pathway analysis initially and then mapping the pathway activity between organisms. Unlike most machine learning approaches, direct methods can be used to predict the activation of a human pathway for a new (test) stimuli, even when that pathway was never

activated by a training stimuli. Gene expression data are available from ArrayExpress (accession E-MTAB-2091), while software implementations are available from http://bioinformaticsprb.med.wayne.edu?p=50 and http://goo.gl/hJny3h. christoph.hafemeister@nyu.edu or atarca@med.wayne.edu. Supplementary data are available at Bioinformatics online.

- **Publisher:** England: Oxford University Press
- **Language:** English
- **Identifier:** ISSN: 1367-4803; EISSN: 1460-2059; EISSN: 1367-4811; DOI: 10.1093/bioinformatics/btu570; PMID: 25150249
- **Source:** Scopus; Medline Complete; Alma/SFX Local Collection

Big data in IBD: big progress for clinical practice

- **Creator:** Seyed Tabib, Nasim Sadat ; Madgwick, Matthew ; Sudhakar, Padhmanand ; Verstockt, Bram ; Korcsmaros, Tamas ; Vermeire, Séverine
- **Subjects:** 1240 ; 1506 ; 2312 ; Algorithms ; Artificial intelligence ; Big Data ; Biology ; Clinical medicine ; Crohn's disease ; Datasets ; Disease management ; Environmental factors ; Gastrointestinal Microbiome ; Gene expression ; Gene Expression Profiling ; Genomics ; Humans ; IBD ; Image Interpretation, Computer-Assisted ; Immune system ; Inflammatory Bowel Diseases - diagnosis ; Inflammatory Bowel Diseases - drug therapy ; Inflammatory Bowel Diseases - genetics ; Inflammatory Bowel Diseases - metabolism ; Inflammatory diseases ; Integration ; Intestinal microflora ; Learning algorithms ; Machine Learning ; Metagenomics ; Microbiomes ; Next-generation sequencing ; Pathogenesis ; Patients ; Precision Medicine ; Principal components analysis ; Prognosis ; Proteins ; Proteomics ; Recent Advances in Clinical Practice ; Research ; Researchers ; Risk Assessment ; Translational Medical Research ; ulcerative colitis
- **Is Part Of:** Gut, 2020-08, Vol.69 (8), p.1520-1532
- **Description:** IBD is a complex multifactorial inflammatory disease of the gut driven by extrinsic and intrinsic factors, including host genetics, the immune system, environmental factors and the gut microbiome. Technological advancements such as next-generation sequencing, high-throughput omics data generation and molecular networks have catalysed IBD research. The advent of artificial intelligence, in particular, machine learning, and systems biology has opened the avenue for the efficient integration and interpretation of big datasets for discovering clinically translatable knowledge. In this narrative review, we discuss how big data integration and machine learning have been applied to translational IBD research. Approaches such as machine learning may enable patient stratification, prediction of disease progression and therapy responses for fine-tuning treatment options with positive

impacts on cost, health and safety. We also outline the challenges and opportunities presented by machine learning and big data in clinical IBD research.
- **Publisher:** England: BMJ Publishing Group LTD
- **Language:** English
- **Identifier:** ISSN: 0017-5749; EISSN: 1468-3288; DOI: 10.1136/gutjnl-2019-320065; PMID: 32111636
- **Source:** BMJ Open Access Journals; Biological Science Collection; Scopus; Alma/SFX Local Collection

Translational Bioinformatics: Linking the Molecular World to the Clinical World

- **Creator:** Altman, R B
- **Subjects:** Abridged Index Medicus ; Artificial Intelligence ; Biological and medical sciences ; Computational biology ; Computational Biology - trends ; Drug Therapy - trends ; Genetic Privacy - ethics ; Genomics ; Humans ; Life Sciences & Biomedicine ; Medical sciences ; Methods ; Pathology, Molecular ; Pharmacology & Pharmacy ; Pharmacology - trends ; Pharmacology. Drug treatments ; Prognosis ; Public Health ; PubMed ; Science & Technology ; Systems Biology ; Translational Medical Research - trends ; United States ; Usage
- **Is Part Of:** Clinical pharmacology and therapeutics, 2012-06, Vol.91 (6), p.994-1000
- **Description:** Translational bioinformatics represents the union of translational medicine and bioinformatics. Translational medicine moves basic biological discoveries from the research bench into the patient−care setting and uses clinical observations to inform basic biology. It focuses on patient care, including the creation of new diagnostics, prognostics, prevention strategies, and therapies based on biological discoveries. Bioinformatics involves algorithms to represent, store, and analyze basic biological data, including DNA sequence, RNA expression, and protein and small−molecule abundance within cells. Translational bioinformatics spans these two fields; it involves the development of algorithms to analyze basic molecular and cellular data with an explicit goal of affecting clinical care. Clinical Pharmacology & Therapeutics (2012); 91 6, 994–1000. doi:10.1038/clpt.2012.49
- **Publisher:** NEW YORK: Blackwell Publishing Ltd
- **Language:** English
- **Identifier:** ISSN: 0009-9236; EISSN: 1532-6535; DOI: 10.1038/clpt.2012.49; PMID: 22549287; CODEN: CLPTAT
- **Source:** Scopus; Web of Science - Science Citation Index Expanded - 2012 from WEB OF SCIENCE; Wiley Online Library All Journals

Radiomics to predict outcomes and abscopal response of patients with cancer treated with immunotherapy combined with radiotherapy using a validated signature of CD8 cells

- **Creator:** Sun, Roger ; Sundahl, Nora ; Hecht, Markus ; Putz, Florian ; Lancia, Andrea ; Rouyar, Angela ; Milic, Marina ; Carré, Alexandre ; Battistella, Enzo ; Alvarez Andres, Emilie ; Niyoteka, Stéphane ; Romano, Edouard ; Louvel, Guillaume ; Durand-Labrunie, Jérôme ; Bockel, Sophie ; Bahleda, Rastilav ; Robert, Charlotte ; Boutros, Celine ; Vakalopoulou, Maria ; Paragios, Nikos ; Frey, Benjamin ; Soria, Jean-Charles ; Massard, Christophe ; Ferté, Charles ; Fietkau, Rainer ; Ost, Piet ; Gaipl, Udo ; Deutsch, Eric
- **Subjects:** Artificial Intelligence ; Cancer ; Computer Science ; Life Sciences ; radioimmunotherapy ; translational medical research ; tumor biomarkers ; tumor microenvironment
- **Is Part Of:** Journal for immunotherapy of cancer, 2020-11-13, Vol.8 (2), p.e001429
- **Description:** BackgroundCombining radiotherapy (RT) with immuno-oncology (IO) therapy (IORT) may enhance IO-induced antitumor response. Quantitative imaging biomarkers can be used to provide prognosis, predict tumor response in a non-invasive fashion and improve patient selection for IORT. A biologically inspired CD8 T-cells-associated radiomics signature has been developed on previous cohorts. We evaluated here whether this CD8 radiomic signature is associated with lesion response, whether it may help to assess disease spatial heterogeneity for predicting outcomes of patients treated with IORT. We also evaluated differences between irradiated and non-irradiated lesions.MethodsClinical data from patients with advanced solid tumors in six independent clinical studies of IORT were investigated. Immunotherapy consisted of 4 different drugs (antiprogrammed death-ligand 1 or anticytotoxic T-lymphocyte-associated protein 4 in monotherapy). Most patients received stereotactic RT to one lesion. Irradiated and non-irradiated lesions were delineated from baseline and the first evaluation CT scans. Radiomic features were extracted from contrast-enhanced CT images and the CD8 radiomics signature was applied. A responding lesion was defined by a decrease in lesion size of at least 30%. Dispersion metrices of the radiomics signature were estimated to evaluate the impact of tumor heterogeneity in patient's response.ResultsA total of 94 patients involving multiple lesions (100 irradiated and 189 non-irradiated lesions) were considered for a statistical interpretation. Lesions with high CD8 radiomics score at baseline were associated with significantly higher tumor response (area under the receiving operating characteristic curve (AUC)=0.63, p=0.0020). Entropy of the radiomics scores distribution on all lesions was shown to be associated with progression-free survival (HR=1.67, p=0.040), out-of-field abscopal response (AUC=0.70, p=0.014) and overall survival (HR=2.08, p=0.023), which remained significant in a multivariate analysis including clinical and biological variables.ConclusionsThese results enhance the predictive value of the biologically inspired CD8 radiomics score and suggests that tumor heterogeneity should be systematically considered in patients

treated with IORT. This CD8 radiomics signature may help select patients who are most likely to benefit from IORT.
- **Publisher:** England: BMJ Publishing Group
- **Language:** English
- **Identifier:** ISSN: 2051-1426; EISSN: 2051-1426; DOI: 10.1136/jitc-2020-001429; PMID: 33188037
- **Source:** BMJ Open Access Journals; Scopus; Publicly Available Content Database; BioMedCentral Open Access

Interactive-cut: Real-time feedback segmentation for translational research

- **Creator:** Egger, Jan ; Lüddemann, Tobias ; Schwarzenberg, Robert ; Freisleben, Bernd ; Nimsky, Christopher
- **Subjects:** Algorithms ; Analysis ; Anatomic Landmarks - pathology ; Aneurysms ; Applied sciences ; Artificial intelligence ; Brain - pathology ; Brain Diseases - pathology ; Computer science ; Computer science; control theory; systems ; Computer Systems ; Engineering ; Engineering, Biomedical ; Exact sciences and technology ; Feedback ; Free drawing templates ; Graph-cut ; Graphics software ; Humans ; Image Enhancement - methods ; Image Interpretation, Computer-Assisted - methods ; Interactive ; Internal Medicine ; Life Sciences & Biomedicine ; Magnetic Resonance Imaging - methods ; Medical imaging equipment ; Other ; Pattern recognition. Digital image processing. Computational geometry ; Predefined templates ; Radiology, Nuclear Medicine & Medical Imaging ; Real-time ; Reproducibility of Results ; Scale-invariant ; Science & Technology ; Segmentation ; Sensitivity and Specificity ; Technology ; Template-based ; Translational Medical Research - methods ; Translational research ; User-Computer Interface
- **Is Part Of:** Computerized medical imaging and graphics, 2014, Vol.38 (4), p.285-295
- **Description:** Abstract In this contribution, a scale-invariant image segmentation algorithm is introduced that “wraps” the algorithm's parameters for the user by its interactive behavior, avoiding the definition of “arbitrary” numbers that the user cannot really understand. Therefore, we designed a specific graph-based segmentation method that only requires a single seed-point inside the target-structure from the user and is thus particularly suitable for immediate processing and interactive, real-time adjustments by the user. In addition, color or gray value information that is needed for the approach can be automatically extracted around the user-defined seed point. Furthermore, the graph is constructed in such a way, so that a polynomial-time mincut computation can provide the segmentation result within a second on an up-to-date computer. The algorithm presented here has been evaluated with fixed seed points on 2D and 3D medical image data, such as brain tumors,

cerebral aneurysms and vertebral bodies. Direct comparison of the obtained automatic segmentation results with costlier, manual slice-by-slice segmentations performed by trained physicians, suggest a strong medical relevance of this interactive approach.

- **Publisher:** OXFORD: Elsevier Ltd
- **Language:** English
- **Identifier:** ISSN: 0895-6111; EISSN: 1879-0771; DOI: 10.1016/j.compmedimag.2014.01.006; PMID: 24613389
- **Source:** Scopus; ScienceDirect Journals (5 years ago - present); Web of Science - Science Citation Index Expanded - 2014 from WEB OF SCIENCE; Alma/SFX Local Collection

Research Evaluation Alongside Clinical Treatment in COVID-19 (REACT COVID-19): an observational and biobanking study

- **Creator:** Burke, Hannah ; Freeman, Anna ; Dushianthan, Ahilanandan ; Celinski, Michael ; Batchelor, James ; Phan, Hang ; Borca, Florina ; Kipps, Christopher ; Thomas, Gareth J ; Faust, Saul N ; Sheard, Natasha ; Williams, Sarah ; Fitzpatrick, Paul ; Landers, Dónal ; Wilkinson, Tom
- **Subjects:** Alliances ; Artificial Intelligence ; Biobanks ; Biological Specimen Banks ; Cancer ; chronic airways disease ; Clinical medicine ; Collaboration ; Coronaviruses ; COVID-19 ; COVID-19 - epidemiology ; COVID-19 - therapy ; Critical care ; Data collection ; health informatics ; Hospitals ; Humans ; Infections ; Medical research ; Natural history ; Pandemics ; Patients ; Prospective Studies ; protocols & guidelines ; respiratory infections ; respiratory medicine (see thoracic medicine) ; SARS-CoV-2 ; Severe acute respiratory syndrome coronavirus 2 ; Translational Medical Research
- **Is Part Of:** BMJ open, 2021-01-22, Vol.11 (1), p.e043012-e043012
- **Description:** IntroductionThe COVID-19 pandemic caused by SARS-CoV-2 places immense worldwide demand on healthcare services. Earlier identification of patients at risk of severe disease may allow intervention with experimental targeted treatments, mitigating the course of their disease and reducing critical care service demand.Methods and analysisThis prospective observational study of patients tested or treated for SARS-CoV-2, who are under the care of the tertiary University Hospital Southampton NHS Foundation Trust (UHSFT), captured data from admission to discharge; data collection commenced on 7 March 2020. Core demographic and clinical information, as well as results of disease-defining characteristics, was captured and recorded electronically from hospital clinical record systems at the point of testing. Manual data were collected and recorded by the clinical research team for assessments which are not part of the structured electronic healthcare record, for example, symptom onset date. Thereafter, participant records were continuously updated during hospital stay and their follow-up period.

Participants aged >16 years were given the opportunity to provide consent for excess clinical sample storage with optional further biological sampling. These anonymised samples were linked to the clinical data in the Real-time Analytics for Clinical Trials platform and were stored within a biorepository at UHSFT.Ethics and disseminationEthical approval was obtained from the HRA Specific Review Board (REC 20/HRA/2986) for waiver of informed consent for the database-only cohort; the procedures conform with the Declaration of Helsinki. The study design, protocol and patient-facing documentation for the biobanking arm of the study have been approved by North West Research Ethics Committee (REC 17/NW/0632) as an amendment to the National Institute for Health Research Southampton Clinical Research Facility-managed Southampton Research Biorepository. This study will be published as peer-reviewed articles and presented at conferences, presentations and workshops.

- **Publisher:** England: BMJ Publishing Group LTD
- **Language:** English
- **Identifier:** ISSN: 2044-6055; EISSN: 2044-6055; DOI: 10.1136/bmjopen-2020-043012; PMID: 33483446
- **Source:** BMJ Open Access Journals; Scopus; Publicly Available Content Database

[Data-driven translational prostate cancer research: From biomarker discovery to clinical decision]

- **Creator:** Lin, Yuxin ; Zhao, Xiaojun ; Miao, Zhijun ; Ling, Zhixin ; Wei, Xuedong ; Pu, Jinxian ; Hou, Jianquan ; Shen, Bairong
- **Subjects:** Artificial Intelligence ; Biomarker discovery ; Biomarkers ; Clinical application ; Humans ; Male ; Precision Medicine ; Prostate cancer ; Prostatic Neoplasms - diagnosis ; Prostatic Neoplasms - genetics ; Prostatic Neoplasms - therapy ; Review ; Systems medicine ; Translational informatics ; Translational Medical Research
- **Is Part Of:** Journal of translational medicine, 2020-03-07, Vol.18 (1), p.119-119
- **Description:** Prostate cancer (PCa) is a common malignant tumor with increasing incidence and high heterogeneity among males worldwide. In the era of big data and artificial intelligence, the paradigm of biomarker discovery is shifting from traditional experimental and small data-based identification toward big data-driven and systems-level screening. Complex interactions between genetic factors and environmental effects provide opportunities for systems modeling of PCa genesis and evolution. We hereby review the current research frontiers in informatics for PCa clinical translation. First, the heterogeneity and complexity in PCa development and clinical theranostics are introduced to raise the concern for PCa systems biology studies. Then biomarkers and risk factors ranging from molecular alternations to

clinical phenotype and lifestyle changes are explicated for PCa personalized management. Methodologies and applications for multi-dimensional data integration and computational modeling are discussed. The future perspectives and challenges for PCa systems medicine and holistic healthcare are finally provided.

- **Publisher:** England: BioMed Central
- **Language:** English
- **Identifier:** ISSN: 1479-5876; EISSN: 1479-5876; DOI: 10.1186/s12967-020-02281-4; PMID: 32143723
- **Source:** Gale Academic OneFile; Gale OneFile: Health and Medicine; Scopus; Gale OneFile: Nursing and Allied Health; Publicly Available Content Database; BioMedCentral Open Access; Academic Search Ultimate; PubMed Central; Alma/SFX Local Collection; DOAJ Directory of Open Access Journals - Not for CDI Discovery

Evaluating temporal relations in clinical text: 2012 i2b2 Challenge

- **Creator:** Sun, Weiyi ; Rumshisky, Anna ; Uzuner, Ozlem
- **Subjects:** Artificial Intelligence ; clinical language processing ; Computer Science ; Computer Science, Information Systems ; Computer Science, Interdisciplinary Applications ; Electronic Health Records ; Health Care Sciences & Services ; Humans ; Information Science & Library Science ; Life Sciences & Biomedicine ; Medical Informatics ; medical language processing ; Natural Language Processing ; Patient Discharge Summaries ; Review ; Science & Technology ; sharedtask challenges ; Technology ; temporal reasoning ; Time ; Translational Medical Research
- **Is Part Of:** Journal of the American Medical Informatics Association : JAMIA, 2013, Vol.20 (5), p.806-813
- **Description:** Background The Sixth Informatics for Integrating Biology and the Bedside (i2b2) Natural Language Processing Challenge for Clinical Records focused on the temporal relations in clinical narratives. The organizers provided the research community with a corpus of discharge summaries annotated with temporal information, to be used for the development and evaluation of temporal reasoning systems. 18 teams from around the world participated in the challenge. During the workshop, participating teams presented comprehensive reviews and analysis of their systems, and outlined future research directions suggested by the challenge contributions. Methods The challenge evaluated systems on the information extraction tasks that targeted: (1) clinically significant events, including both clinical concepts such as problems, tests, treatments, and clinical departments, and events relevant to the patient's clinical timeline, such as admissions, transfers between departments, etc; (2) temporal expressions, referring to the dates, times, durations, or

frequencies phrases in the clinical text. The values of the extracted temporal expressions had to be normalized to an ISO specification standard; and (3) temporal relations, between the clinical events and temporal expressions. Participants determined pairs of events and temporal expressions that exhibited a temporal relation, and identified the temporal relation between them. Results For event detection, statistical machine learning (ML) methods consistently showed superior performance. While ML and rule based methods seemed to detect temporal expressions equally well, the best systems overwhelmingly adopted a rule based approach for value normalization. For temporal relation classification, the systems using hybrid approaches that combined ML and heuristics based methods produced the best results.

- **Publisher:** OXFORD: OXFORD UNIV PRESS
- **Language:** English
- **Identifier:** ISSN: 1067-5027; EISSN: 1527-974X; DOI: 10.1136/amiajnl-2013-001628; PMID: 23564629
- **Source:** Scopus; Medline Complete; Web of Science - Social Sciences Citation Index - 2013 from WEB OF SCIENCE; Web of Science - Science Citation Index Expanded - 2013 from WEB OF SCIENCE

[Translational Analysis of Mouse and Human Placental Protein and mRNA Reveals Distinct Molecular Pathologies in Human Preeclampsia]

- **Creator:** Cox, Brian ; Sharma, Parveen ; Evangelou, Andreas I ; Whiteley, Kathie ; Ignatchenko, Vladimir ; Ignatchenko, Alex ; Baczyk, Dora ; Czikk, Marie ; Kingdom, John ; Rossant, Janet ; Gramolini, Anthony O ; Adamson, S. Lee ; Kislinger, Thomas
- **Subjects:** Algorithms ; Animals ; Antigens, CD - genetics ; Antigens, CD - metabolism ; Artificial Intelligence ; Bayes Theorem ; Biochemical Research Methods ; Biochemistry & Molecular Biology ; Biomarkers - metabolism ; Endoglin ; Endothelium - metabolism ; Female ; Gene Expression ; Giant Cells - metabolism ; GTP-Binding Protein alpha Subunits, G12-G13 - genetics ; GTP-Binding Protein alpha Subunits, G12-G13 - metabolism ; Humans ; Life Sciences & Biomedicine ; MAP Kinase Signaling System ; Membrane Proteins - genetics ; Membrane Proteins - metabolism ; Mice ; Mice, Inbred C57BL ; Placenta - metabolism ; Placenta - pathology ; Pre-Eclampsia - diagnosis ; Pre-Eclampsia - genetics ; Pre-Eclampsia - metabolism ; Pregnancy ; Receptors, Cell Surface - genetics ; Receptors, Cell Surface - metabolism ; Research ; RNA, Messenger - metabolism ; Science & Technology ; Software ; Translational Medical Research ; Trophoblasts - metabolism ; Vascular Endothelial Growth Factor Receptor-1 - genetics ; Vascular Endothelial Growth Factor Receptor-1 - metabolism

- **Is Part Of:** Molecular & cellular proteomics, 2011-12, Vol.10 (12), p.M111.012526-M111.012526
- **Description:** Preeclampsia (PE) adversely impacts ~5% of pregnancies. Despite extensive research, no consistent biomarkers or cures have emerged, suggesting that different molecular mechanisms may cause clinically similar disease. To address this, we undertook a proteomics study with three main goals: (1) to identify a panel of cell surface markers that distinguish the trophoblast and endothelial cells of the placenta in the mouse; (2) to translate this marker set to human via the Human Protein Atlas database; and (3) to utilize the validated human trophoblast markers to identify subgroups of human preeclampsia. To achieve these goals, plasma membrane proteins at the blood tissue interfaces were extracted from placentas using intravascular silica-bead perfusion, and then identified using shotgun proteomics. We identified 1181 plasma membrane proteins, of which 171 were enriched at the maternal blood-trophoblast interface and 192 at the fetal endothelial interface with a 70% conservation of expression in humans. Three distinct molecular subgroups of human preeclampsia were identified in existing human microarray data by using expression patterns of trophoblast-enriched proteins. Analysis of all misexpressed genes revealed divergent dysfunctions including angiogenesis (subgroup 1), MAPK signaling (subgroup 2), and hormone biosynthesis and metabolism (subgroup 3). Subgroup 2 lacked expected changes in known preeclampsia markers (sFLT1, sENG) and uniquely overexpressed GNA12. In an independent set of 40 banked placental specimens, GNA12 was overexpressed during preeclampsia when co-incident with chronic hypertension. In the current study we used a novel translational analysis to integrate mouse and human trophoblast protein expression with human microarray data. This strategy identified distinct molecular pathologies in human preeclampsia. We conclude that clinically similar preeclampsia patients exhibit divergent placental gene expression profiles thus implicating divergent molecular mechanisms in the origins of this disease.
- **Publisher:** ROCKVILLE: Elsevier Inc
- **Language:** English
- **Identifier:** ISSN: 1535-9476; EISSN: 1535-9484; DOI: 10.1074/mcp.M111.012526; PMID: 21986993
- **Source:** Scopus; HighWire Press (Free Journals); Web of Science - Science Citation Index Expanded - 2011 from WEB OF SCIENCE; Alma/SFX Local Collection; DOAJ Directory of Open Access Journals - Not for CDI Discovery

Adapting translational research methods to water, sanitation, and hygiene

- **Creator:** Setty, Karen ; Cronk, Ryan ; George, Shannan ; Anderson, Darcy ; O'flaherty, Ghanja ; Bartram, Jamie

- **Subjects:** Dissemination ; Environmental Sciences ; Environmental Sciences & Ecology ; Evidence-based practice ; Humans ; Hygiene ; Implementation science ; Knowledge translation ; Life Sciences & Biomedicine ; Participatory research ; Public, Environmental & Occupational Health ; Quality improvement ; Research design ; Review ; Sanitation ; Science & Technology ; Translational Medical Research ; Water ; Water Supply
- **Is Part Of:** International journal of environmental research and public health, 2019-10-01, Vol.16 (20), p.4049
- **Description:** Translational research applies scientific techniques to achieve practical outcomes, connecting pure research and pure practice. Many translational research types have arisen since the mid-1900s, reflecting the need to better integrate scientific advancement with policy and practice. Water, sanitation, and hygiene (WaSH) development efforts have aimed to reduce morbidity and mortality and improve service delivery; thus, associated research has a strong orientation toward applied studies that use diverse methods to support decision-making. Drawing from knowledge that emerged to support other professional fields, such as manufacturing and clinical healthcare, we characterize different types of translational research and clarify nomenclature and principles. We describe study approaches relevant to translational research questions, and offer overarching recommendations, specific examples, and resources for further study as practical advice to professionals who seek to apply translational methods to WaSH problems. To enhance collective outcomes, professionals should mindfully align projects within the translational spectrum. We further recommend overarching good practices such as documenting intervention adaptations, overtly considering contextual factors, and better distinguishing efficacy from effectiveness research by replicating studies in different contexts. By consciously improving the compatibility and linkages between WaSH science and practice, this guide can accelerate urgently needed progress toward global development goals.
- **Publisher:** BASEL: MDPI
- **Language:** English
- **Identifier:** ISSN: 1661-7827; ISSN: 1660-4601; EISSN: 1660-4601; DOI: 10.3390/ijerph16204049; PMID: 31652610
- **Source:** Scopus; Materials Science & Engineering Collection; Medline Complete; Publicly Available Content Database; Web of Science - Science Citation Index Expanded - 2019 from WEB OF SCIENCE; Alma/SFX Local Collection; Web of Science - Social Sciences Citation Index – 2019 from WEB OF SCIENCE; DOAJ Directory of Open Access Journals - Not for CDI Discovery

AIBench: A rapid application development framework for translational research in biomedicine

- **Creator:** Glez-Peña, D ; Reboiro-Jato, M ; Maia, P ; Rocha, M ; Díaz, F ; Fdez-Riverola, F
- **Subjects:** Application framework ; Artificial Intelligence ; Bioengineering - statistics & numerical data ; Biological and medical sciences ; Biomedical informatics ; Computational Biology ; Computer Graphics ; Computer Science ; Computer Science, Interdisciplinary Applications ; Computer Science, Theory & Methods ; Computer Systems ; Data Mining - statistics & numerical data ; Engineering ; Engineering, Biomedical ; Genomics - statistics & numerical data ; Internal Medicine ; Life Sciences & Biomedicine ; Medical Informatics ; Medical sciences ; Open software ; Other ; Radiotherapy. Instrumental treatment. Physiotherapy. Reeducation. Rehabilitation, orthophony, crenotherapy. Diet therapy and various other treatments (general aspects) ; Reusable component model ; Science & Technology ; Scientific software development ; Software Design ; Technology ; Technology. Biomaterials. Equipments. Material. Instrumentation ; Translational Medical Research - statistics & numerical data
- **Is Part Of:** Computer methods and programs in biomedicine, 2009, Vol.98 (2), p.191-203
- **Description:** Abstract Applied research in both biomedical discovery and translational medicine today often requires the rapid development of fully featured applications containing both advanced and specific functionalities, for real use in practice. In this context, new tools are demanded that allow for efficient generation, deployment and reutilization of such biomedical applications as well as their associated functionalities. In this context this paper presents AIBench, an open-source Java desktop application framework for scientific software development with the goal of providing support to both fundamental and applied research in the domain of translational biomedicine. AIBench incorporates a powerful plug-in engine, a flexible scripting platform and takes advantage of Java annotations, reflection and various design principles in order to make it easy to use, lightweight and non-intrusive. By following a basic input–processing–output life cycle, it is possible to fully develop multiplatform applications using only three types of concepts: operations , data-types and views . The framework automatically provides functionalities that are present in a typical scientific application including user parameter definition, logging facilities, multi-threading execution, experiment repeatability and user interface workflow management, among others. The proposed framework architecture defines a reusable component model which also allows assembling new applications by the reuse of libraries from past projects or third-party software.
- **Publisher:** CLARE: Elsevier Ireland Ltd
- **Language:** English
- **Identifier:** ISSN: 0169-2607; EISSN: 1872-7565; DOI: 10.1016/j.cmpb.2009.12.003; PMID: 20047774

- **Source:** Life Sciences Japan [FCJLS]; Scopus; ScienceDirect Physical Sciences College Edition Backfile; Elsevier:ScienceDirect:Health Sciences Subject Collection:2018; ScienceDirect Journals (Transactional Access); ScienceDirect Government Edition; ScienceDirect Polish National Consort 2007; ScienceDirect Journals (5 years ago - present); Web of Science - Science Citation Index Expanded - 2010 from WEB OF SCIENCE; Elsevier:ScienceDirect:Computer Science Subject Collection:2018; Alma/SFX Local Collection

Finding temporal gene expression patterns for translational research

- **Creator:** Tusch, Guenter ; Tolea, Olvi ; Kutsumi, Yuka ; Sam, Vincent K ; Mamidi, Lakshmi
- **Subjects:** Algorithms ; Artificial Intelligence ; bioinformatics ; clinical and epidemiological research and trials ; Data Mining - methods ; Database Management Systems ; Databases, Genetic ; education ; enhancing biological ; Gene Expression Profiling - methods ; Gene Expression Regulation - genetics ; Health technology assessment ; Medical Record Linkage - methods ; Meta-Analysis as Topic ; Natural Language Processing ; Oligonucleotide Array Sequence Analysis - methods ; Research ; Software ; Translational Medical Research - methods ; translational research
- **Is Part Of:** Studies in Health Technology and Informatics, 2013, Vol.192 (1-2), p.1173-1173
- **Description:** Translational research of time-series of gene-expression microarray datasets makes use on gene expression profiles that have been obtained at different points in time. Our web-based multi-user program helps a researcher find temporal patterns like peaks in large pre-selected microarray data sets that include data from different but related studies in publicly available databases. If all studies use the same platform, data can be combined for a meta-analysis type approach. For combination of data from different platforms we allow only Affymetrix GeneChips, for which a method for pooling of information exists. To search for time patterns, the data are transformed into an abstract layer that is independent from the particular selection of time point in the individual studies.
- **Publisher:** Netherlands
- **Language:** English
- **Identifier:** ISSN: 0926-9630; ISBN: 1614992886; ISBN: 9781614992882; EISSN: 1879-8365; DOI: 10.3233/978-1-61499-289-9-1173; PMID: 23920947
- **Source:** Scopus; Medline Complete; Alma/SFX Local Collection

Neuroimaging-Based Biomarkers in Psychiatry: Clinical Opportunities of a Paradigm Shift

- **Creator:** Fu, Cynthia H Y ; Costafreda, Sergi G
- **Subjects:** Alzheimer ; Artificial Intelligence ; Behavior ; Biomarkers ; Biomedical research ; Brain - pathology ; Brain - physiopathology ; Clinical medicine ; Depression ; Diagnosis, Differential ; Early Diagnosis ; Functional Neuroimaging - methods ; Humans ; Life Sciences & Biomedicine ; Machine learning ; Magnetic resonance imaging ; Medical imaging ; Medical research ; Mental Disorders - diagnosis ; Mental Disorders - physiopathology ; Mental Disorders - therapy ; Monitoring, Physiologic - methods ; Neural Networks (Computer) ; Neuroimaging ; Neurosciences ; Patients ; Personalized ; Prognosis ; Psychiatry ; Psychiatry - methods ; Schizophrenia ; Science & Technology ; Statistical methods ; Studies ; Translational Medical Research ; Treatment Outcome
- **Is Part Of:** The Canadian Journal of Psychiatry, 2013-09, Vol.58 (9), p.499-508
- **Description:** Neuroimaging research has substantiated the functional and structural abnormalities underlying psychiatric disorders but has, thus far, failed to have a significant impact on clinical practice. Recently, neuroimaging-based diagnoses and clinical predictions derived from machine learning analysis have shown significant potential for clinical translation. This review introduces the key concepts of this approach, including how the multivariate integration of patterns of brain abnormalities is a crucial component. We survey recent findings that have potential application for diagnosis, in particular early and differential diagnoses in Alzheimer disease and schizophrenia, and the prediction of clinical response to treatment in depression. We discuss the specific clinical opportunities and the challenges for developing biomarkers for psychiatry in the absence of a diagnostic gold standard. We propose that longitudinal outcomes, such as early diagnosis and prediction of treatment response, offer definite opportunities for progress. We propose that efforts should be directed toward clinically challenging predictions in which neuroimaging may have added value, compared with the existing standard assessment. We conclude that diagnostic and prognostic biomarkers will be developed through the joint application of expert psychiatric knowledge in addition to advanced methods of analysis.
- **Publisher:** Los Angeles, CA: SAGE Publications
- **Language:** English
- **Identifier:** ISSN: 0706-7437; EISSN: 1497-0015; DOI: 10.1177/070674371305800904; PMID: 24099497
- **Source:** Scopus; Web of Science - Social Sciences Citation Index - 2013  ; Web of Science - Science Citation Index Expanded - 2013  ; SAGE Journals Premier 2021 (PREM2021)

Translational meta-analysis tool for temporal gene expression profiles

- **Creator:** Tusch, Guenter ; Tole, Olvi
- **Subjects:** Artificial Intelligence ; Bioinformatics ; Data Mining - methods ; Database Management Systems ; Databases, Genetic ; Epidemiological research and clinical trials ; Gene Expression Profiling - methods ; Health technology assessment ; Medical Record Linkage - methods ; Meta-Analysis as Topic ; Oligonucleotide Array Sequence Analysis - methods ; Research and education ; Software ; Translational Medical Research - methods ; Translational research
- **Is Part Of:** Studies in Health Technology and Informatics, 2012, Vol.180, p.1156-1158
- **Description:** Widespread use of microarray technology that led to highly complex datasets often is addressing similar or related biological questions. In translational medicine research is often based on measurements that have been obtained at different points in time. However, the researcher looks at them as a progression over time. If a biological stimulus shows an effect on a particular gene that is reversed over time, this would show, for instance, as a peak in the gene's temporal expression profile. Our program SPOT helps researchers find these patterns in large sets of microarray data. We created the software tool using open-source platforms and the Semantic Web tool Protégé-OWL.
- **Publisher:** Netherlands
- **Language:** English
- **Identifier:** ISSN: 0926-9630; ISBN: 9781614991007; ISBN: 1614991006; EISSN: 1879-8365; DOI: 10.3233/978-1-61499-101-4-1156; PMID: 22874385
- **Source:** Scopus; Medline Complete; Alma/SFX Local Collection

A hybrid system for temporal information extraction from clinical text

- **Creator:** Tang, Buzhou ; Wu, Yonghui ; Jiang, Min ; Chen, Yukun ; Denny, Joshua C ; Xu, Hua
- **Subjects:** Artificial Intelligence ; Clinic event extraction ; Computer Science ; Computer Science, Information Systems ; Computer Science, Interdisciplinary Applications ; Electronic Health Records ; Health Care Sciences & Services ; Humans ; Information Science & Library Science ; Life Sciences & Biomedicine ; Machine learning ; Medical Informatics ; Natural Language Processing ; Patient Discharge Summaries ; Research and Applications ; Science & Technology ; Technology ; Temporal expression extraction ; Temporal information extraction ; Temporal relation extraction ; Time ; Translational Medical Research

- **Is Part Of:** Journal of the American Medical Informatics Association : JAMIA, 2013, Vol.20 (5), p.828-835
- **Description:** Objective To develop a comprehensive temporal information extraction system that can identify events, temporal expressions, and their temporal relations in clinical text. This project was part of the 2012 i2b2 clinical natural language processing (NLP) challenge on temporal information extraction. Materials and methods The 2012 i2b2 NLP challenge organizers manually annotated 310 clinic notes according to a defined annotation guideline: a training set of 190 notes and a test set of 120 notes. All participating systems were developed on the training set and evaluated on the test set. Our system consists of three modules: event extraction, temporal expression extraction, and temporal relation (also called Temporal Link, or TLink') extraction. The TLink extraction module contains three individual classifiers for TLinks: (1) between events and section times, (2) within a sentence, and (3) across different sentences. The performance of our system was evaluated using scripts provided by the i2b2 organizers. Primary measures were micro-averaged Precision, Recall, and F-measure. Results Our system was among the top ranked. It achieved F-measures of 0.8659 for temporal expression extraction (ranked fourth), 0.6278 for end-to-end TLink track (ranked first), and 0.6932 for TLink-only track (ranked first) in the challenge. We subsequently investigated different strategies for TLink extraction, and were able to marginally improve performance with an F-measure of 0.6943 for TLink-only track.
- **Publisher:** OXFORD: OXFORD UNIV PRESS
- **Language:** English
- **Identifier:** ISSN: 1067-5027; EISSN: 1527-974X; DOI: 10.1136/amiajnl-2013-001635; PMID: 23571849
- **Source:** Scopus; Medline Complete; Web of Science - Social Sciences Citation Index - 2013 from WEB OF SCIENCE; Web of Science - Science Citation Index Expanded - 2013 from WEB OF SCIENCE

[Predicting translational progress in biomedical research]

- **Creator:** Ian Hutchins, B ; Davis, Matthew T ; Meseroll, Rebecca A ; Santangelo, M
- **Contributor:** Kimmelman, Jonathan
- **Subjects:** Artificial intelligence ; Bibliometrics ; Biochemistry & Molecular Biology ; Biology ; Biomedical research ; Biomedical Research - trends ; Citation analysis ; Classification ; Clinical Trials as Topic ; Computer and Information Sciences ; Health aspects ; Humans ; Knowledge ; Learning algorithms ; Life Sciences & Biomedicine ; Life Sciences & Biomedicine - Other Topics ; Machine Learning ; Medical research ; Medical Subject Headings-MeSH ; Medical treatment ; Medicine ; Medicine and Health Sciences ; Medicine, Experimental ; Meta ; Methods ; National libraries ;

Periodicals as Topic ; Physical Sciences ; Practice Guidelines as Topic ; Predictions ; Research ; Research and Analysis Methods ; Science & Technology ; Time Factors ; Translation ; Translational Medical Research - trends
- **Is Part Of:** PLoS biology, 2019, Vol.17 (10), p.e3000416-e3000416
- **Description:** Fundamental scientific advances can take decades to translate into improvements in human health. Shortening this interval would increase the rate at which scientific discoveries lead to successful treatment of human disease. One way to accomplish this would be to identify which advances in knowledge are most likely to translate into clinical research. Toward that end, we built a machine learning system that detects whether a paper is likely to be cited by a future clinical trial or guideline. Despite the noisiness of citation dynamics, as little as 2 years of postpublication data yield accurate predictions about a paper's eventual citation by a clinical article (accuracy = 84%, F1 score = 0.56; compared to 19% accuracy by chance). We found that distinct knowledge flow trajectories are linked to papers that either succeed or fail to influence clinical research. Translational progress in biomedicine can therefore be assessed and predicted in real time based on information conveyed by the scientific community's early reaction to a paper.
- **Publisher:** SAN FRANCISCO: PUBLIC LIBRARY SCIENCE
- **Language:** English
- **Identifier:** ISSN: 1544-9173; ISSN: 1545-7885; EISSN: 1545-7885; DOI: 10.1371/journal.pbio.3000416; PMID: 31600189
- **Source:** Gale Academic OneFile; Biological Science Collection; Gale OneFile: Health and Medicine; Scopus; Gale OneFile: Agriculture; Gale In Context: Science; Gale In Context: Opposing Viewpoints; Public Library of Science (PLoS); Medline Complete; Publicly Available Content Database; Academic Search Ultimate; Web of Science - Science Citation Index Expanded - 2019  ; PubMed Central; DOAJ Directory of Open Access Journals - Not for CDI Discovery

Semantic SenseLab: Implementing the vision of the Semantic Web in neuroscience

- **Creator:** Samwald, Matthias ; Chen, Huajun ; Ruttenberg, Alan ; Lim, Ernest ; Marenco, Luis ; Miller, Perry ; Shepherd, Gordon ; Cheung, Kei-Hoi
- **Subjects:** Brain Mapping - methods ; Computer Science ; Computer Science, Artificial Intelligence ; Description logic ; Engineering ; Engineering, Biomedical ; Humans ; Information Dissemination ; Integration ; Internal Medicine ; Internet ; Life Sciences & Biomedicine ; Medical Informatics ; Nerve Net ; Neuroscience ; Neurosciences ; Ontology mapping ; Other ; Science & Technology ; Semantic Web ; Semantics ; Technology ; Translational Medical Research ; Web Ontology Language
- **Is Part Of:** Artificial intelligence in medicine, 2009, Vol.48 (1), p.21-28

- **Description:** Abstract Objective Integrative neuroscience research needs a scalable informatics framework that enables semantic integration of diverse types of neuroscience data. This paper describes the use of the Web Ontology Language (OWL) and other Semantic Web technologies for the representation and integration of molecular-level data provided by several of SenseLab suite of neuroscience databases. Methods Based on the original database structure, we semi-automatically translated the databases into OWL ontologies with manual addition of semantic enrichment. The SenseLab ontologies are extensively linked to other biomedical Semantic Web resources, including the Subcellular Anatomy Ontology, Brain Architecture Management System, the Gene Ontology, BIRNLex and UniProt. The SenseLab ontologies have also been mapped to the Basic Formal Ontology and Relation Ontology, which helps ease interoperability with many other existing and future biomedical ontologies for the Semantic Web. In addition, approaches to representing contradictory research statements are described. The SenseLab ontologies are designed for use on the Semantic Web that enables their integration into a growing collection of biomedical information resources. Conclusion We demonstrate that our approach can yield significant potential benefits and that the Semantic Web is rapidly becoming mature enough to realize its anticipated promises. The ontologies are available online at http://neuroweb.med.yale.edu/senselab/.
- **Publisher:** AMSTERDAM: Elsevier B.V
- **Language:** English
- **Identifier:** ISSN: 0933-3657; EISSN: 1873-2860; DOI: 10.1016/j.artmed.2009.11.003; PMID: 20006477
- **Source:** Life Sciences Japan [FCJLS]; ScienceDirect Pi2 Collection; Scopus; ScienceDirect Physical Sciences College Edition Backfile; Elsevier:ScienceDirect:Health Sciences Subject Collection:2018; ScienceDirect Journals (Transactional Access); ScienceDirect Government Edition; ScienceDirect Polish National Consort 2007; ScienceDirect Journals (5 years ago - present); Web of Science - Social Sciences Citation Index - 2010 from WEB OF SCIENCE; Web of Science - Science Citation Index Expanded - 2010 from WEB OF SCIENCE; Elsevier:ScienceDirect:Computer Science Subject Collection:2018; Alma/SFX Local Collection; ScienceDirect Health & Life Sciences College Edition Backfile

Ordinal response prediction using bootstrap aggregation, with application to a high-throughput methylation data set

- **Creator:** Archer, K. J ; Mas, V. R
- **Subjects:** Artificial Intelligence ; Bootstrap aggregating ; Bootstrap method ; Classification trees ; Computer Simulation ; CpG Islands - physiology ; Data Interpretation, Statistical ; Gene expression ; Genomics ; Humans ; Life Sciences &

Biomedicine ; Machine learning ; Mathematical & Computational Biology ; Mathematics ; Medical Informatics ; Medical statistics ; Medicine, Research & Experimental ; Metastasis ; Methylation ; Oligonucleotide Array Sequence Analysis - methods ; Ordinal response ; Physical Sciences ; Public, Environmental & Occupational Health ; Research & Experimental Medicine ; Science & Technology ; Statistics & Probability ; Translational Medical Research - methods ; Tumors

- **Is Part Of:** Statistics in medicine, 2009-12-20, Vol.28 (29), p.3597-3610
- **Description:** Many investigators conducting translational research are performing high−throughput genomic experiments and then developing multigenic classifiers using the resulting high−dimensional data set. In a large number of applications, the class to be predicted may be inherently ordinal. Examples of ordinal outcomes include tumor−node−metastasis (TNM) stage (I, II, III, IV); drug toxicity evaluated as none, mild, moderate, or severe; and response to treatment classified as complete response, partial response, stable disease, or progressive disease. While one can apply nominal response classification methods to ordinal response data, in doing so some information is lost that may improve the predictive performance of the classifier. This study examined the effectiveness of alternative ordinal splitting functions combined with bootstrap aggregation for classifying an ordinal response. We demonstrate that the ordinal impurity and ordered twoing methods have desirable properties for classifying ordinal response data and both perform well in comparison to other previously described methods. Developing a multigenic classifier is a common goal for microarray studies, and therefore application of the ordinal ensemble methods is demonstrated on a high−throughput methylation data set. Copyright © 2009 John Wiley & Sons, Ltd.
- **Publisher:** Chichester, UK: John Wiley & Sons, Ltd
- **Language:** English
- **Identifier:** ISSN: 0277-6715; EISSN: 1097-0258; DOI: 10.1002/sim.3707; PMID: 19697302; CODEN: SMEDDA
- **Source:** Scopus; Wiley Online Library Full Collection 2016; Wiley Online Library Journals + OA Offset 2015-2017; Wiley-Blackwell PALCI Collection; Orbis-Cascade Wiley-Blackwell Shared Titles 2011; Web of Science - Science Citation Index Expanded - 2009 from WEB OF SCIENCE; Wiley Online Library All Journals

Computational translation of genomic responses from experimental model systems to humans

- **Creator:** Brubaker, Douglas K ; Proctor, Elizabeth A ; Haigis, Kevin M ; Lauffenburger, Douglas A
- **Contributor:** Bonneau, Richard A

- **Subjects:** Acids ; Animal models ; Animals ; Artificial intelligence ; Biochemical Research Methods ; Biochemistry & Molecular Biology ; Bioengineering ; Biology ; Biology and Life Sciences ; Biomedical materials ; Cancer ; Computation ; Computer and Information Sciences ; Computer applications ; Consortia ; Data sets ; Databases ; Datasets ; Design of experiments ; Disease Models, Animal ; Engineering ; Experimental design ; Failure rates ; Gene expression ; Gene Expression Profiling - methods ; Genetic research ; Genetic translation ; Genomes ; Genomics ; Genomics - methods ; Genotype & phenotype ; Humans ; Inference ; Inflammation - genetics ; Inflammation - metabolism ; Inflammatory diseases ; Learning ; Life Sciences & Biomedicine ; Machine learning ; Mathematical & Computational Biology ; Medical research ; Medicine and Health Sciences ; Mice ; Neural networks ; Neural Networks, Computer ; Ontology ; Phenotypes ; Physical Sciences ; Research ; Research and Analysis Methods ; Rodents ; Science & Technology ; Sepsis ; Streptococcus infections ; Supervised Machine Learning ; Transcriptome - genetics ; Translation ; Translational Medical Research - methods
- **Is Part Of:** PLoS computational biology, 2019, Vol.15 (1), p.e1006286-e1006286
- **Description:** The high failure rate of therapeutics showing promise in mouse models to translate to patients is a pressing challenge in biomedical science. Though retrospective studies have examined the fidelity of mouse models to their respective human conditions, approaches for prospective translation of insights from mouse models to patients remain relatively unexplored. Here, we develop a semi-supervised learning approach for inference of disease-associated human differentially expressed genes and pathways from mouse model experiments. We examined 36 transcriptomic case studies where comparable phenotypes were available for mouse and human inflammatory diseases and assessed multiple computational approaches for inferring human biology from mouse datasets. We found that semi-supervised training of a neural network identified significantly more true human biological associations than interpreting mouse experiments directly. Evaluating the experimental design of mouse experiments where our model was most successful revealed principles of experimental design that may improve translational performance. Our study shows that when prospectively evaluating biological associations in mouse studies, semi-supervised learning approaches, combining mouse and human data for biological inference, provide the most accurate assessment of human in vivo disease processes. Finally, we proffer a delineation of four categories of model system-to-human "Translation Problems" defined by the resolution and coverage of the datasets available for molecular insight translation and suggest that the task of translating insights from model systems to human disease contexts may be better accomplished by a combination of translation-minded experimental design and computational approaches.
- **Publisher:** SAN FRANCISCO: PUBLIC LIBRARY SCIENCE
- **Language:** English

- **Identifier:** ISSN: 1553-734X; ISSN: 1553-7358; EISSN: 1553-7358; DOI: 10.1371/journal.pcbi.1006286; PMID: 30629591
- **Source:** Gale Academic OneFile; Biological Science Collection; Scopus; Gale In Context: Science; Public Library of Science (PLoS); Medline Complete; Gale OneFile: Nursing and Allied Health; Publicly Available Content Database; Academic Search Ultimate; Web of Science - Science Citation Index Expanded - 2019  ; PubMed Central; DOAJ Directory of Open Access Journals - Not for CDI Discovery

Combining rules and machine learning for extraction of temporal expressions and events from clinical narratives

- **Creator:** Kovačević, Aleksandar ; Dehghan, Azad ; Filannino, Michele ; Keane, John A ; Nenadic, Goran
- **Subjects:** Artificial Intelligence ; clinical NLP ; clinical text mining ; Computer Science ; Computer Science, Information Systems ; Computer Science, Interdisciplinary Applications ; Electronic Health Records ; event extraction ; Health Care Sciences & Services ; Humans ; Information Science & Library Science ; Information Storage and Retrieval - methods ; Life Sciences & Biomedicine ; Medical Informatics ; Natural Language Processing ; Research and Applications ; Science & Technology ; Technology ; termporal expression extraction ; termporal expression normalization ; Time ; Translational Medical Research
- **Is Part Of:** Journal of the American Medical Informatics Association : JAMIA, 2013, Vol.20 (5), p.859-866
- **Description:** Objective Identification of clinical events (eg, problems, tests, treatments) and associated temporal expressions (eg, dates and times) are key tasks in extracting and managing data from electronic health records. As part of the i2b2 2012 Natural Language Processing for Clinical Data challenge, we developed and evaluated a system to automatically extract temporal expressions and events from clinical narratives. The extracted temporal expressions were additionally normalized by assigning type, value, and modifier. Materials and methods The system combines rule-based and machine learning approaches that rely on morphological, lexical, syntactic, semantic, and domain-specific features. Rule-based components were designed to handle the recognition and normalization of temporal expressions, while conditional random fields models were trained for event and temporal recognition. Results The system achieved micro F scores of 90% for the extraction of temporal expressions and 87% for clinical event extraction. The normalization component for temporal expressions achieved accuracies of 84.73% (expression's type), 70.44% (value), and 82.75% (modifier). Discussion Compared to the initial agreement between human annotators (87-89%), the system provided comparable performance

for both event and temporal expression mining. While (lenient) identification of such mentions is achievable, finding the exact boundaries proved challenging. Conclusions The system provides a state-of-the-art method that can be used to support automated identification of mentions of clinical events and temporal expressions in narratives either to support the manual review process or as a part of a large-scale processing of electronic health databases.

- **Publisher:** OXFORD: OXFORD UNIV PRESS
- **Language:** English
- **Identifier:** ISSN: 1067-5027; EISSN: 1527-974X; DOI: 10.1136/amiajnl-2013-001625; PMID: 23605114
- **Source:** Scopus; Medline Complete; Web of Science - Social Sciences Citation Index - 2013 from WEB OF SCIENCE; Web of Science - Science Citation Index Expanded - 2013 from WEB OF SCIENCE

An end-to-end system to identify temporal relation in discharge summaries: 2012 i2b2 challenge

- **Creator:** Xu, Yan ; Wang, Yining ; Liu, Tianren ; Tsujii, Junichi ; Chang, Eric I-Chao
- **Subjects:** Artificial Intelligence ; Computer Science ; Computer Science, Information Systems ; Computer Science, Interdisciplinary Applications ; Context-Free Grammar ; Coordination ; Dependency Tree ; Electronic Health Records ; Health Care Sciences & Services ; Humans ; Information Science & Library Science ; Information Storage and Retrieval - methods ; Labeled Sequential Pattern ; Life Sciences & Biomedicine ; Medical Informatics ; Natural Language Processing ; Patient Discharge Summaries ; Research and Applications ; Science & Technology ; Syntax ; Technology ; Time ; Translational Medical Research
- **Is Part Of:** Journal of the American Medical Informatics Association : JAMIA, 2013, Vol.20 (5), p.849-858
- **Description:** Objective To create an end-to-end system to identify temporal relation in discharge summaries for the 2012 i2b2 challenge. The challenge includes event extraction, timex extraction, and temporal relation identification. Design An end-to-end temporal relation system was developed. It includes three subsystems: an event extraction system (conditional random fields (CRF) name entity extraction and their corresponding attribute classifiers), a temporal extraction system (CRF name entity extraction, their corresponding attribute classifiers, and context-free grammar based normalization system), and a temporal relation system (10 multi-support vector machine (SVM) classifiers and a Markov logic networks inference system) using labeled sequential pattern mining, syntactic structures based on parse trees, and results from a coordination classifier. Micro-averaged precision (P), recall (R),

averaged P&R (P&R), and F measure (F) were used to evaluate results. Results For event extraction, the system achieved 0.9415 (P), 0.8930 (R), 0.9166 (P&R), and 0.9166 (F). The accuracies of their type, polarity, and modality were 0.8574, 0.8585, and 0.8560, respectively. For timex extraction, the system achieved 0.8818, 0.9489, 0.9141, and 0.9141, respectively. The accuracies of their type, value, and modifier were 0.8929, 0.7170, and 0.8907, respectively. For temporal relation, the system achieved 0.6589, 0.7129, 0.6767, and 0.6849, respectively. For end-to-end temporal relation, it achieved 0.5904, 0.5944, 0.5921, and 0.5924, respectively. With the F measure used for evaluation, we were ranked first out of 14 competing teams (event extraction), first out of 14 teams (timex extraction), third out of 12 teams (temporal relation), and second out of seven teams (end-to-end temporal relation). Conclusions The system achieved encouraging results, demonstrating the feasibility of the tasks defined by the i2b2 organizers. The experiment result demonstrates that both global and local information is useful in the 2012 challenge.

- **Publisher:** LONDON: BMJ PUBLISHING GROUP
- **Language:** English
- **Identifier:** ISSN: 1067-5027; EISSN: 1527-974X; DOI: 10.1136/amiajnl-2012-001607; PMID: 23467472
- **Source:** Scopus; Medline Complete; Web of Science - Social Sciences Citation Index - 2013 from WEB OF SCIENCE; Web of Science - Science Citation Index Expanded - 2013 from WEB OF SCIENCE

[Biomedical informatics and translational medicine]

- **Creator:** Sarkar, Indra N
- **Subjects:** Analysis ; Decision Making ; Electronic Health Records ; Evidence-based medicine ; Health aspects ; Humans ; Information Storage and Retrieval ; Life Sciences & Biomedicine ; Medical informatics ; Medical Informatics - manpower ; Medical Informatics - methods ; Medical Informatics - standards ; Medicine, Research & Experimental ; Natural Language Processing ; Research ; Research & Experimental Medicine ; Review ; Science & Technology ; Translational Medical Research - manpower ; Translational Medical Research - methods ; Translational Medical Research - standards ; Usage
- **Is Part Of:** Journal of translational medicine, 2010-02-26, Vol.8 (1), p.22-22
- **Description:** Biomedical informatics involves a core set of methodologies that can provide a foundation for crossing the "translational barriers" associated with translational medicine. To this end, the fundamental aspects of biomedical informatics (e.g., bioinformatics, imaging informatics, clinical informatics, and public health informatics) may be essential in helping improve the ability to bring basic research findings to the bedside, evaluate the efficacy of interventions across

communities, and enable the assessment of the eventual impact of translational medicine innovations on health policies. Here, a brief description is provided for a selection of key biomedical informatics topics (Decision Support, Natural Language Processing, Standards, Information Retrieval, and Electronic Health Records) and their relevance to translational medicine. Based on contributions and advancements in each of these topic areas, the article proposes that biomedical teams.

- **Publisher:** LONDON: BIOMED CENTRAL LTD
- **Language:** English
- **Identifier:** ISSN: 1479-5876; EISSN: 1479-5876; DOI: 10.1186/1479-5876-8-22; PMID: 20187952
- **Source:** Gale Academic OneFile; Gale OneFile: Health and Medicine; Scopus; Gale OneFile: Nursing and Allied Health; Publicly Available Content Database; BioMedCentral Open Access; Academic Search Ultimate; Web of Science - Science Citation Index Expanded - 2010 from WEB OF SCIENCE; PubMed Central; Alma/SFX Local Collection; DOAJ Directory of Open Access Journals - Not for CDI Discovery

A flexible framework for recognizing events, temporal expressions, and temporal relations in clinical text

- **Creator:** Roberts, Kirk ; Rink, Bryan ; Harabagiu, Sanda M
- **Subjects:** Artificial Intelligence ; Clinical Informatics ; Computer Science ; Computer Science, Information Systems ; Computer Science, Interdisciplinary Applications ; Electronic Health Records ; Health Care Sciences & Services ; Humans ; Information Science & Library Science ; Information Storage and Retrieval - methods ; Life Sciences & Biomedicine ; Medical Informatics ; Medical Records Systems, Computerized ; Natural Language Processing ; Research and Applications ; Science & Technology ; Technology ; Time ; Translational Medical Research
- **Is Part Of:** Journal of the American Medical Informatics Association : JAMIA, 2013, Vol.20 (5), p.867-875
- **Description:** Objective To provide a natural language processing method for the automatic recognition of events, temporal expressions, and temporal relations in clinical records. Materials and Methods A combination of supervised, unsupervised, and rule-based methods were used. Supervised methods include conditional random fields and support vector machines. A flexible automated feature selection technique was used to select the best subset of features for each supervised task. Unsupervised methods include Brown clustering on several corpora, which result in our method being considered semisupervised. Results On the 2012 Informatics for Integrating Biology and the Bedside (i2b2) shared task data, we achieved an overall event F1-measure of 0.8045, an overall temporal expression F1-measure of 0.6154, an overall

temporal link detection F1-measure of 0.5594, and an end-to-end temporal link detection F1-measure of 0.5258. The most competitive system was our event recognition method, which ranked third out of the 14 participants in the event task. Discussion Analysis reveals the event recognition method has difficulty determining which modifiers to include/exclude in the event span. The temporal expression recognition method requires significantly more normalization rules, although many of these rules apply only to a small number of cases. Finally, the temporal relation recognition method requires more advanced medical knowledge and could be improved by separating the single discourse relation classifier into multiple, more targeted component classifiers. Conclusions Recognizing events and temporal expressions can be achieved accurately by combining supervised and unsupervised methods, even when only minimal medical knowledge is available. Temporal normalization and temporal relation recognition, however, are far more dependent on the modeling of medical knowledge.

- **Publisher:** LONDON: BMJ PUBLISHING GROUP
- **Language:** English
- **Identifier:** ISSN: 1067-5027; EISSN: 1527-974X; DOI: 10.1136/amiajnl-2013-001619; PMID: 23686936
- **Source:** Scopus; Medline Complete; Web of Science - Social Sciences Citation Index - 2013 from WEB OF SCIENCE; Web of Science - Science Citation Index Expanded - 2013 from WEB OF SCIENCE

Synthetic Lethal Networks for Precision Oncology: Promises and Pitfalls

- **Creator:** Shen, John Paul ; Ideker, Trey
- **Subjects:** Animals ; Biochemistry & Molecular Biology ; Biomarkers, Tumor ; Cancer Genomics ; Computational Biology - methods ; Disease Susceptibility ; Drug approval ; Epistasis, Genetic ; Genetic Interaction ; Genomics ; Genomics - methods ; Humans ; Life Sciences & Biomedicine ; Neoplasms - diagnosis ; Neoplasms - etiology ; Neoplasms - metabolism ; Neoplasms - therapy ; Precision Medicine ; Precision Medicine - methods ; Science & Technology ; Synthetic Lethal ; Synthetic Lethal Mutations ; Systems Biology ; Translational Medical Research
- **Is Part Of:** Journal of molecular biology, 2018-09-14, Vol.430 (18), p.2900-2912
- **Description:** Synthetic lethal interactions, in which the simultaneous loss of function of two genes produces a lethal phenotype, are being explored as a means to therapeutically exploit cancer-specific vulnerabilities and expand the scope of precision oncology. Currently, three Food and Drug Administration-approved drugs work by targeting the synthetic lethal interaction between BRCA1/2 and PARP. This review examines additional efforts to discover networks of synthetic lethal interactions and discusses both challenges and opportunities regarding the translation

of new synthetic lethal interactions into the clinic. [Display omitted] •Synthetic lethal interactions have the potential to greatly expand the scope of precision oncology.•New technologies allow for high-throughput screens to identify new synthetic lethal interactions.•Consideration of cellular context will be critical when designing synthetic lethal cancer therapies.

- **Publisher:** LONDON: Elsevier Ltd
- **Language:** English
- **Identifier:** ISSN: 0022-2836; EISSN: 1089-8638; DOI: 10.1016/j.jmb.2018.06.026; PMID: 29932943
- **Source:** Scopus; ScienceDirect Journals (5 years ago - present); Web of Science - Science Citation Index Expanded - 2018 from WEB OF SCIENCE

[Global informetric perspective studies on translational medical research]

- **Creator:** Yao, Qiang ; Lyu, Peng-Hui ; Ma, Fei-Cheng ; Yao, Lan ; Zhang, Shi-Jing
- **Subjects:** Bibliometrics ; Cluster analysis ; Developing countries ; Evidence-based medicine ; Female ; Homeopathy ; Humans ; Information Seeking Behavior ; Informetrics ; Life Sciences & Biomedicine ; Materia medica and therapeutics ; Medical colleges ; Medical Informatics ; Science & Technology ; Social network analysis ; Social Networking ; Social networks ; Systems Integration ; Therapeutics ; Translational medical research ; Translational Medical Research - standards ; Translational Medical Research - statistics & numerical data ; Translational Medical Research - trends ; Translational research ; Usage ; Web of science
- **Is Part Of:** BMC medical informatics and decision making, 2013, Vol.13 (1), p.77-77
- **Description:** Background: Translational medical research literature has increased rapidly in the last few decades and played a more and more important role during the development of medicine science. The main aim of this study is to evaluate the global performance of translational medical research during the past few decades. Methods: Bibliometric, social network analysis, and visualization technologies were used for analyzing translational medical research performance from the aspects of subject categories, journals, countries, institutes, keywords, and MeSH terms. Meanwhile, the co-author, co-words and cluster analysis methods were also used to trace popular topics in translational medical research related work. Results: Research output suggested a solid development in translational medical research, in terms of increasing scientific production and research collaboration. We identified the core journals, mainstream subject categories, leading countries, and institutions in translational medical research. There was an uneven distribution of publications at authorial, institutional, and national levels. The most commonly used keywords that appeared in the articles were "translational research", "translational medicine",

"biomarkers", "stroke", "inflammation", "cancer", and "breast cancer". Conclusions: The subject categories of "Research & Experimental Medicine", "Medical Laboratory Technology", and "General & Internal Medicine" play a key role in translational medical research both in production and in its networks. Translational medical research and CTS, etc. are core journals of translational research. G7 countries are the leading nations for translational medical research. Some developing countries, such as P.R China, also play an important role in the communication of translational research. The USA and its institutions play a dominant role in the production, collaboration, citations and high quality articles. The research trends in translational medical research involve drug design and development, pathogenesis and treatment of disease, disease model research, evidence-based research, and stem and progenitor cells.

- **Publisher:** LONDON: BMC
- **Language:** English
- **Identifier:** ISSN: 1472-6947; EISSN: 1472-6947; DOI: 10.1186/1472-6947-13-77; PMID: 23885955
- **Source:** Gale Academic OneFile; Biological Science Collection; Gale OneFile: Health and Medicine; Scopus; Medline Complete; Gale OneFile: Nursing and Allied Health; Publicly Available Content Database; BioMedCentral Open Access; Academic Search Ultimate; Web of Science - Social Sciences Citation Index - 2013  ; Web of Science - Science Citation Index Expanded - 2013  ; PubMed Central; Alma/SFX Local Collection; DOAJ Directory of Open Access Journals - Not for CDI Discovery

À la recherche du temps perdu: Extracting temporal relations from medical text in the 2012 i2b2 NLP challenge

- **Creator:** Cherry, Colin ; Zhu, Xiaodan ; Martin, Joel ; de Bruijn, Berry
- **Subjects:** 1506 ; Artificial Intelligence ; clinical text ; Computer Science ; Computer Science, Information Systems ; Computer Science, Interdisciplinary Applications ; Electronic Health Records ; Health Care Sciences & Services ; Humans ; information extraction ; Information Science & Library Science ; Information Storage and Retrieval - methods ; Life Sciences & Biomedicine ; Medical Informatics ; Natural Language Processing ; Patient Discharge Summaries ; relation extraction ; Research and Applications ; Science & Technology ; Technology ; temporal reasoning ; Time ; Translational Medical Research
- **Is Part Of:** Journal of the American Medical Informatics Association : JAMIA, 2013, Vol.20 (5), p.843-848
- **Description:** Objective An analysis of the timing of events is critical for a deeper understanding of the course of events within a patient record. The 2012 i2b2 NLP

challenge focused on the extraction of temporal relationships between concepts within textual hospital discharge summaries. Materials and methods The team from the National Research Council Canada (NRC) submitted three system runs to the second track of the challenge: typifying the time-relationship between pre-annotated entities. The NRC system was designed around four specialist modules containing statistical machine learning classifiers. Each specialist targeted distinct sets of relationships: local relationships, sectime'-type relationships, non-local overlap-type relationships, and non-local causal relationships. Results The best NRC submission achieved a precision of 0.7499, a recall of 0.6431, and an F1 score of 0.6924, resulting in a statistical tie for first place. Post hoc improvements led to a precision of 0.7537, a recall of 0.6455, and an F1 score of 0.6954, giving the highest scores reported on this task to date. Discussion and conclusions Methods for general relation extraction extended well to temporal relations, and gave top-ranked state-of-the-art results. Careful ordering of predictions within result sets proved critical to this success.

- **Publisher:** OXFORD: OXFORD UNIV PRESS
- **Language:** English
- **Identifier:** ISSN: 1067-5027; EISSN: 1527-974X; DOI: 10.1136/amiajnl-2013-001624; PMID: 23523875
- **Source:** Scopus; Medline Complete; Web of Science - Social Sciences Citation Index - 2013 from WEB OF SCIENCE; Web of Science - Science Citation Index Expanded - 2013 from WEB OF SCIENCE

[Strategies for maintaining patient privacy in i2b2]

- **Creator:** Murphy, Shawn N ; Gainer, Vivian ; Mendis, Michael ; Churchill, Susanne ; Kohane, Isaac
- **Subjects:** Algorithms ; Artificial Intelligence ; automated learning ; bioinformatics ; clinical informatics ; clinical research ; clinical research informatics ; common rule ; Computer Science ; Computer Science, Information Systems ; Computer Science, Interdisciplinary Applications ; Computer Systems ; Confidentiality ; data exchange ; data models ; discovery, text and data mining methods ; genomics ; Health Care Sciences & Services ; Health Insurance Portability and Accountability Act ; HIPAA ; Humans ; image representation ; Information Science & Library Science ; Information Storage and Retrieval ; knowledge representations ; Life Sciences & Biomedicine ; linking the genotype and phenotype ; Medical Informatics ; medical records ; natural-language processing ; patient privacy ; processing and display analysis ; Research and Applications ; Science & Technology ; software architecture ; Systems Integration ; Technology ; Translational Medical Research - organization & administration ; United States ; visualization of data and knowledge

- **Is Part Of:** Journal of the American Medical Informatics Association : JAMIA, 2011-12, Vol.18 (1), p.103-108
- **Description:** Background The re-use of patient data from electronic healthcare record systems can provide tremendous benefits for clinical research, but measures to protect patient privacy while utilizing these records have many challenges. Some of these challenges arise from a misperception that the problem should be solved technically when actually the problem needs a holistic solution. Objective The authors' experience with informatics for integrating biology and the bedside (i2b2) use cases indicates that the privacy of the patient should be considered on three fronts: technical de-identification of the data, trust in the researcher and the research, and the security of the underlying technical platforms. Methods The security structure of i2b2 is implemented based on consideration of all three fronts. It has been supported with several use cases across the USA, resulting in five privacy categories of users that serve to protect the data while supporting the use cases. Results The i2b2 architecture is designed to provide consistency and faithfully implement these user privacy categories. These privacy categories help reflect the policy of both the Health Insurance Portability and Accountability Act and the provisions of the National Research Act of 1974, as embodied by current institutional review boards. Conclusion By implementing a holistic approach to patient privacy solutions, i2b2 is able to help close the gap between principle and practice.
- **Publisher:** LONDON: B M J PUBLISHING GROUP
- **Language:** English
- **Identifier:** ISSN: 1067-5027; EISSN: 1527-974X; DOI: 10.1136/amiajnl-2011-000316; PMID: 21984588
- **Source:** Scopus; Medline Complete; Web of Science - Social Sciences Citation Index - 2011 from WEB OF SCIENCE; Web of Science - Science Citation Index Expanded - 2011 from WEB OF SCIENCE

Chapter 16: Text Mining for Translational Bioinformatics

- **Creator:** Cohen, K. Bretonnel ; Hunter, Lawrence E
- **Contributor:** Lewitter, Fran ; Kann, Maricel
- **Subjects:** Algorithms ; Animals ; Artificial Intelligence ; Biochemical Research Methods ; Biochemistry & Molecular Biology ; Bioinformatics ; Biology ; Computational biology ; Computational Biology - methods ; Computer Science ; Computer Simulation ; Data mining ; Data Mining - methods ; Education ; Genetic translation ; Humans ; Life Sciences & Biomedicine ; Linguistics ; Mathematical & Computational Biology ; Phenotype ; Programming Languages ; Research ; Science & Technology ; Software ; Studies ; Translational Medical Research
- **Is Part Of:** PLoS computational biology, 2013-04, Vol.9 (4), p.e1003044-e1003044

- **Description:** Text mining for translational bioinformatics is a new field with tremendous research potential. It is a subfield of biomedical natural language processing that concerns itself directly with the problem of relating basic biomedical research to clinical practice, and vice versa. Applications of text mining fall both into the category of T1 translational research-translating basic science results into new interventions-and T2 translational research, or translational research for public health. Potential use cases include better phenotyping of research subjects, and pharmacogenomic research. A variety of methods for evaluating text mining applications exist, including corpora, structured test suites, and post hoc judging. Two basic principles of linguistic structure are relevant for building text mining applications. One is that linguistic structure consists of multiple levels. The other is that every level of linguistic structure is characterized by ambiguity. There are two basic approaches to text mining: rule-based, also known as knowledge-based; and machine-learning-based, also known as statistical. Many systems are hybrids of the two approaches. Shared tasks have had a strong effect on the direction of the field. Like all translational bioinformatics software, text mining software for translational bioinformatics can be considered health-critical and should be subject to the strictest standards of quality assurance and software testing.
- **Publisher:** SAN FRANCISCO: PUBLIC LIBRARY SCIENCE
- **Language:** English
- **Identifier:** ISSN: 1553-734X; ISSN: 1553-7358; EISSN: 1553-7358; DOI: 10.1371/journal.pcbi.1003044; PMID: 23633944
- **Source:** Gale Academic OneFile; Biological Science Collection; Scopus; Gale In Context: Science; Public Library of Science (PLoS); Medline Complete; Gale OneFile: Nursing and Allied Health; Publicly Available Content Database; Academic Search Ultimate; Web of Science - Science Citation Index Expanded - 2013  ; PubMed Central; DOAJ Directory of Open Access Journals - Not for CDI Discovery

[Virus-based nanomaterials as positron emission tomography and magnetic resonance contrast agents: from technology development to translational medicine]

- **Creator:** Shukla, Sourabh ; Steinmetz, Nicole F
- **Subjects:** Biomedical Technology - methods ; Contrast Media - chemistry ; Humans ; Life Sciences & Biomedicine ; Magnetic Resonance Imaging ; Medicine, Research & Experimental ; Nanoscience & Nanotechnology ; Nanostructures - chemistry ; Positron-Emission Tomography ; Research & Experimental Medicine ; Science & Technology ; Science & Technology - Other Topics ; Translational Medical Research ; Viruses - chemistry

- **Is Part Of:** Wiley interdisciplinary reviews. Nanomedicine and nanobiotechnology, 2015-09, Vol.7 (5), p.708-721
- **Description:** Viruses have recently emerged as ideal protein scaffolds for a new class of contrast agents that can be used in medical imaging procedures such as positron emission tomography (PET) and magnetic resonance imaging (MRI). Whereas synthetic nanoparticles are difficult to produce as homogeneous formulations due to the inherently stochastic nature of the synthesis process, virus−based nanoparticles are genetically encoded and are therefore produced as homogeneous and monodisperse preparations with a high degree of quality control. Because the virus capsids have a defined chemical structure that has evolved to carry cargoes of nucleic acids, they can be modified to carry precisely defined cargoes of contrast agents and can be decorated with spatially defined contrast reagents on the internal or external surfaces. Viral nanoparticles can also be genetically programed or conjugated with targeting ligands to deliver contrast agents to specific cells, and the natural biocompatibility of viruses means that they are cleared rapidly from the body. Nanoparticles based on bacteriophages and plant viruses are safe for use in humans and can be produced inexpensively in large quantities as self−assembling recombinant proteins. Based on these considerations, a new generation of contrast agents has been developed using bacteriophages and plant viruses as scaffolds to carry positron−emitting radioisotopes such as [18F] fluorodeoxyglucose for PET imaging and iron oxide or Gd3+ for MRI. Although challenges such as immunogenicity, loading efficiency, and regulatory compliance remain to be address, virus−based nanoparticles represent a promising new enabling technology for a new generation of highly biocompatible and biodegradable targeted imaging reagents. WIREs Nanomed Nanobiotechnol 2015, 7:708–721. doi: 10.1002/wnan.1335 This article is categorized under: Biology-Inspired Nanomaterials > Protein and Virus-Based Structures
- **Publisher:** Hoboken, USA: John Wiley & Sons, Inc
- **Language:** English
- **Identifier:** ISSN: 1939-5116; EISSN: 1939-0041; DOI: 10.1002/wnan.1335; PMID: 25683790
- **Source:** Scopus; Medline Complete; Web of Science - Science Citation Index Expanded - 2015 from WEB OF SCIENCE™; Wiley Online Library All Journals

[Computational psychopharmacology: a translational and pragmatic approach]

- **Creator:** Robbins, Trevor W ; Cardinal, Rudolf N
- **Contributor:** Robbins, Trevor ; Flagel, Shelly B ; Baunez, Christelle ; Paulus, Martin

- **Subjects:** Animals ; Biomedicine ; Computational Biology - methods ; Computational Biology - trends ; Computer modelling ; Depression ; Depression, Mental ; Depressive Disorder - diagnostic imaging ; Depressive Disorder - metabolism ; Dopamine ; Dopamine - metabolism ; Drug discovery ; Humans ; Learning - physiology ; Life Sciences & Biomedicine ; Nervous system diseases ; Neurosciences ; Neurosciences & Neurology ; Pharmacology & Pharmacy ; Pharmacology/Toxicology ; Psychiatry ; Psychopharmacology - methods ; Psychopharmacology - trends ; Reinforcement learning ; Reinforcement, Psychology ; Research ; Review ; Schizophrenia ; Schizophrenia - diagnostic imaging ; Schizophrenia - metabolism ; Science & Technology ; Serotonin ; Serotonin - metabolism ; Translational Medical Research - methods ; Translational Medical Research - trends
- **Is Part Of:** Psychopharmacology, 2019-08-01, Vol.236 (8), p.2295-2305
- **Description:** Psychopharmacology needs novel quantitative measures and theoretical approaches based on computational modelling that can be used to help translate behavioural findings from experimental animals to humans, including patients with neuropsychiatric disorders.This brief review exemplifies this approach when applied to recent published studies of the effects of manipulating central dopaminergic and serotoninergic systems in rodents and marmoset monkeys, and possible comparisons with healthy human volunteers receiving systemic agents or patients with depression and schizophrenia.Behavioural effects of central depletions of dopamine or serotonin in monkeys in probabilistic learning paradigms are characterised further by computational modelling methods and related to rodent and human data.Several examples are provided of the power of computational modelling to derive new measures and reappraise conventional explanations of regional neurotransmitter depletion and other drug effects, whilst enhancing construct validation in patient groups. Specifically, effects are shown on such parameters as ‘stimulus stickiness’ and ‘side stickiness’, which occur over and above effects on standard parameters of reinforcement learning, reminiscent of some early innovations in data analysis in psychopharmacology.Computational modelling provides a useful methodology for further detailed analysis of behavioural mechanisms that are affected by pharmacological manipulations across species and will aid the translation of experimental findings to understand the therapeutic effects of medications in neuropsychiatric disorders, as well as facilitating future drug discovery.
- **Publisher:** Berlin/Heidelberg: Springer Berlin Heidelberg
- **Language:** English
- **Identifier:** ISSN: 0033-3158; EISSN: 1432-2072; DOI: 10.1007/s00213-019-05302-3; PMID: 31273400
- **Source:** Gale Academic OneFile; Gale OneFile: Health and Medicine; Springer Online Journals Complete; Scopus; Gale OneFile: Criminal Justice; SPORTDiscus with Full Text; Medline Complete; Gale OneFile: Nursing and Allied Health; Gale OneFile: Psychology; CINAHL Complete; Academic Search Ultimate; Web of

Science - Science Citation Index Expanded - 2019 from WEB OF SCIENCE; Web of Science - Social Sciences Citation Index – 2019 from WEB OF SCIENCE

Artificial Intelligence and the Implementation Challenge

- **Creator:** Shaw, James ; Rudzicz, Frank ; Jamieson, Trevor ; Goldfarb, Avi
- **Subjects:** Artificial intelligence ; Artificial Intelligence - standards ; Ethics ; Health Care Sciences & Services ; Humans ; Implementation science ; Life Sciences & Biomedicine ; Machine learning ; Machine Learning - standards ; Medical Informatics ; Science & Technology ; Telemedicine - methods ; Viewpoint
- **Is Part Of:** Journal of medical Internet research, 2019-07-01, Vol.21 (7), p.e13659-e13659
- **Description:** Background: Applications of artificial intelligence (AI) in health care have garnered much attention in recent years, but the implementation issues posed by AI have not been substantially addressed. Objective: In this paper, we have focused on machine learning (ML) as a form of AI and have provided a framework for thinking about use cases of ML in health care. We have structured our discussion of challenges in the implementation of ML in comparison with other technologies using the framework of Nonadoption, Abandonment, and Challenges to the Scale-Up, Spread, and Sustainability of Health and Care Technologies (NASSS). Methods: After providing an overview of AI technology, we describe use cases of ML as falling into the categories of decision support and automation. We suggest these use cases apply to clinical, operational, and epidemiological tasks and that the primary function of ML in health care in the near term will be decision support. We then outline unique implementation issues posed by ML initiatives in the categories addressed by the NASSS framework, specifically including meaningful decision support, explainability, privacy, consent, algorithmic bias, security, scalability, the role of corporations, and the changing nature of health care work. Results: Ultimately, we suggest that the future of ML in health care remains positive but uncertain, as support from patients, the public, and a wide range of health care stakeholders is necessary to enable its meaningful implementation. Conclusions: If the implementation science community is to facilitate the adoption of ML in ways that stand to generate widespread benefits, the issues raised in this paper will require substantial attention in the coming years.
- **Publisher:** TORONTO: JMIR PUBLICATIONS, INC
- **Language:** English
- **Identifier:** ISSN: 1438-8871; ISSN: 1439-4456; EISSN: 1438-8871; DOI: 10.2196/13659; PMID: 31293245
- **Source:** Social Science Premium Collection; Scopus; Medline Complete; Publicly Available Content Database; Web of Science - Science Citation Index Expanded -

2019 from WEB OF SCIENCE; PubMed Central; Web of Science - Social Sciences Citation Index – 2019 from WEB OF SCIENCE; DOAJ Directory of Open Access Journals - Not for CDI Discovery

Contributions from the 2018 Literature on Bioinformatics and Translational Informatics

- **Creator:** Smaïl-Tabbone, Malika ; Rance, Bastien
- **Subjects:** Section 8: Bioinformatics and Translational Informatics
- **Is Part Of:** Yearbook of medical informatics, 2019-08, Vol.28 (1), p.190-193
- **Description:** Summary Objectives : To summarize recent research and select the best papers published in 2018 in the field of Bioinformatics and Translational Informatics (BTI) for the corresponding section of the International Medical Informatics Association (IMIA) Yearbook. Methods : A literature review was performed for retrieving from PubMed papers indexed with keywords and free terms related to BTI. Independent review allowed the two section editors to select a list of 14 candidate best papers which were subsequently peer-reviewed. A final consensus meeting gathering the whole IMIA Yearbook editorial committee was organized to finally decide on the selection of the best papers. Results : Among the 636 retrieved papers published in 2018 in the various subareas of BTI, the review process selected four best papers. The first paper presents a computational method to identify molecular markers for targeted treatment of acute myeloid leukemia using multi-omics data (genome-wide gene expression profiles) and in vitro sensitivity to 160 chemotherapy drugs. The second paper describes a deep neural network approach to predict the survival of patients suffering from glioma on the basis of digitalised pathology images and genomics biomarkers. The authors of the third paper adopt a pan-cancer approach to take benefit of multi-omics data for drug repurposing. The fourth paper presents a graph-based semi-supervised method to accurate phenotype classification applied to ovarian cancer. Conclusions : Thanks to the normalization of open data and open science practices, research in BTI continues to develop and mature. Noteworthy achievements are sophisticated applications of leading edge machine-learning methods dedicated to personalized medicine.
- **Publisher:** Stuttgart: Georg Thieme Verlag KG
- **Language:** English
- **Identifier:** ISSN: 0943-4747; EISSN: 2364-0502; DOI: 10.1055/s-0039-1677945
- **Source:** Thieme Open Access Journals

The digital revolution in phenotyping

- **Creator:** Oellrich, Anika ; Collier, Nigel ; Groza, Tudor ; Rebholz-Schuhmann, Dietrich ; Shah, Nigam ; Bodenreider, Olivier ; Boland, Mary Regina ; Georgiev, Ivo ; Liu, Hongfang ; Livingston, Kevin ; Luna, Augustin ; Mallon, Ann-Marie ; Manda, Prashanti ; Robinson, Peter N ; Rustici, Gabriella ; Simon, Michelle ; Wang, Liqin ; Winnenburg, Rainer ; Dumontier, Michel
- **Subjects:** Acquisition ; Biochemical Research Methods ; Biochemistry & Molecular Biology ; Humans ; Information Storage and Retrieval ; Interoperability ; Knowledge discovery ; Life Sciences & Biomedicine ; Mathematical & Computational Biology ; Papers ; Phenomics ; Phenotype ; Phenotypes ; Research Design ; Science & Technology ; Semantic representation ; Translational Medical Research
- **Is Part Of:** Briefings in bioinformatics, 2016-09, Vol.17 (5), p.819-830
- **Description:** Phenotypes have gained increased notoriety in the clinical and biological domain owing to their application in numerous areas such as the discovery of disease genes and drug targets, phylogenetics and pharmacogenomics. Phenotypes, defined as observable characteristics of organisms, can be seen as one of the bridges that lead to a translation of experimental findings into clinical applications and thereby support 'bench to bedside' efforts. However, to build this translational bridge, a common and universal understanding of phenotypes is required that goes beyond domain-specific definitions. To achieve this ambitious goal, a digital revolution is ongoing that enables the encoding of data in computer-readable formats and the data storage in specialized repositories, ready for integration, enabling translational research. While phenome research is an ongoing endeavor, the true potential hidden in the currently available data still needs to be unlocked, offering exciting opportunities for the forthcoming years. Here, we provide insights into the state-of-the-art in digital phenotyping, by means of representing, acquiring and analyzing phenotype data. In addition, we provide visions of this field for future research work that could enable better applications of phenotype data.
- **Publisher:** OXFORD: OXFORD UNIV PRESS
- **Language:** English
- **Identifier:** ISSN: 1467-5463; EISSN: 1477-4054; DOI: 10.1093/bib/bbv083; PMID: 26420780
- **Source:** Scopus; Business Source Premier; Medline Complete; Academic Search Ultimate; Web of Science - Social Sciences Citation Index - 2016 from WEB OF SCIENCE™; Web of Science - Science Citation Index Expanded - 2016 from WEB OF SCIENCE™

Ontologies, Knowledge Representation, and Machine Learning for Translational Research: Recent Contributions

- **Creator:** Robinson, Peter N ; Haendel, Melissa A
- **Subjects:** Section 6: Knowledge Representation and Management

- **Is Part Of:** Yearbook of medical informatics, 2020-08, Vol.29 (1), p.159-162
- **Description:** Summary Objectives : To select, present, and summarize the most relevant papers published in 2018 and 2019 in the field of Ontologies and Knowledge Representation, with a particular focus on the intersection between Ontologies and Machine Learning. Methods : A comprehensive review of the medical informatics literature was performed to select the most interesting papers published in 2018 and 2019 and that document the utility of ontologies for computational analysis, including machine learning. Results : Fifteen articles were selected for inclusion in this survey paper. The chosen articles belong to three major themes: (i) the identification of phenotypic abnormalities in electronic health record (EHR) data using the Human Phenotype Ontology ; (ii) word and node embedding algorithms to supplement natural language processing (NLP) of EHRs and other medical texts; and (iii) hybrid ontology and NLP-based approaches to extracting structured and unstructured components of EHRs. Conclusion : Unprecedented amounts of clinically relevant data are now available for clinical and research use. Machine learning is increasingly being applied to these data sources for predictive analytics, precision medicine, and differential diagnosis. Ontologies have become an essential component of software pipelines designed to extract, code, and analyze clinical information by machine learning algorithms. The intersection of machine learning and semantics is proving to be an innovative space in clinical research.
- **Publisher:** Stuttgart: Georg Thieme Verlag KG
- **Language:** English
- **Identifier:** ISSN: 0943-4747; EISSN: 2364-0502; DOI: 10.1055/s-0040-1701991
- **Source:** Thieme Open Access Journals

Crowdsourcing for translational research: Analysis of biomarker expression using cancer microarrays

- **Creator:** Lawson, Jonathan ; Robinson-Vyas, Rupesh J ; McQuillan, Janette P ; Paterson, Andy ; Christie, Sarah ; Kidza-Griffiths, Matthew ; McDuffus, Leigh-Anne ; Moutasim, Karwan A ; Shaw, Emily C ; Kiltie, Anne E ; Howat, William J ; Hanby, Andrew M ; Thomas, Gareth J ; Smittenaar, Peter
- **Subjects:** biomarker ; Biomarkers, Tumor - metabolism ; cancer ; crowdsourcing ; Crowdsourcing - methods ; Data Interpretation, Statistical ; Humans ; Image Processing, Computer-Assisted - methods ; Immunohistochemistry ; Life Sciences & Biomedicine ; Molecular Diagnostics ; Neoplasms - metabolism ; Oncology ; pathology ; Patient Selection ; Science & Technology ; Tissue Array Analysis ; tissue microarray ; Translational Medical Research - methods
- **Is Part Of:** British journal of cancer, 2017-01-17, Vol.116 (2), p.237-245

- **Description:** Background: Academic pathology suffers from an acute and growing lack of workforce resource. This especially impacts on translational elements of clinical trials, which can require detailed analysis of thousands of tissue samples. We tested whether crowdsourcing -enlisting help from the public -is a sufficiently accurate method to score such samples. Methods: We developed a novel online interface to train and test lay participants on cancer detection and immunohistochemistry scoring in tissue microarrays. Lay participants initially performed cancer detection on lung cancer images stained for CD8, and we measured how extending a basic tutorial by annotated example images and feedback-based training affected cancer detection accuracy. We then applied this tutorial to additional cancer types and immunohistochemistry markers -bladder/ki67, lung/EGFR, and oesophageal/CD8 -to establish accuracy compared with experts. Using this optimised tutorial, we then tested lay participants' accuracy on immunohistochemistry scoring of lung/EGFR and bladder/p53 samples. Results: We observed that for cancer detection, annotated example images and feedback-based training both improved accuracy compared with a basic tutorial only. Using this optimised tutorial, we demonstrate highly accurate (40.90 area under curve) detection of cancer in samples stained with nuclear, cytoplasmic and membrane cell markers. We also observed high Spearman correlations between lay participants and experts for immunohistochemistry scoring (0.91 (0.78, 0.96) and 0.97 (0.91, 0.99) for lung/EGFR and bladder/p53 samples, respectively). Conclusions: These results establish crowdsourcing as a promising method to screen large data sets for biomarkers in cancer pathology research across a range of cancers and immunohistochemical stains.
- **Publisher:** LONDON: NATURE PUBLISHING GROUP
- **Language:** English
- **Identifier:** ISSN: 0007-0920; EISSN: 1532-1827; DOI: 10.1038/bjc.2016.404; PMID: 27959886
- **Source:** Nature Open Access; Biological Science Collection; Scopus; Web of Science - Science Citation Index Expanded - 2017 from WEB OF SCIENCE; Nature Journals Online

[From Code to Bedside: Implementing Artificial Intelligence Using Quality Improvement Methods]

- **Creator:** Smith, Margaret ; Sattler, Amelia ; Hong, Grace ; Lin, Steven
- **Subjects:** Analysis ; Artificial intelligence ; design thinking ; Health care reform ; implementation science ; Medical colleges ; Methods ; Quality control ; quality improvement
- **Is Part Of:** Journal of general internal medicine : JGIM, 2021-04-01, Vol.36 (4), p.1061-1066

- **Description:** Despite increasing interest in how artificial intelligence (AI) can augment and improve healthcare delivery, the development of new AI models continues to outpace adoption in existing healthcare processes. Integration is difficult because current approaches separate the development of AI models from the complex healthcare environments in which they are intended to function, resulting in models developed without a clear and compelling use case and not tested or scalable in a clinical setting. We propose that current approaches and traditional research methods do not support successful AI implementation in healthcare and outline a repeatable mixed-methods approach, along with several examples, that facilitates uptake of AI technologies into human-driven healthcare processes. Unlike traditional research, these methods do not seek to control for variation, but rather understand it to learn how a technology will function in practice coupled with user-centered design techniques. This approach, leveraging design thinking and quality improvement methods, aims to increase the adoption of AI in healthcare and prompt further study to understand which methods are most successful for AI implementations.
- **Publisher:** United States: Springer
- **Language:** English
- **Identifier:** ISSN: 0884-8734; EISSN: 1525-1497; DOI: 10.1007/s11606-020-06394-w; PMID: 33469745
- **Source:** Springer Online Journals Complete; Scopus

Realistic simulation of virtual multi-scale, multi-modal patient trajectories using Bayesian networks and sparse auto-encoders

- **Creator:** Sood, Meemansa ; Sahay, Akrishta ; Karki, Reagon ; Emon, Mohammad Asif ; Vrooman, Henri ; Hofmann-Apitius, Martin ; Fröhlich, Holger
- **Subjects:** Alzheimer Disease - diagnostic imaging ; Alzheimer Disease - genetics ; Bayes Theorem ; Bayesian analysis ; Brain - diagnostic imaging ; Cohort Studies ; Computer Simulation ; Databases, Factual - statistics & numerical data ; Deep Learning ; Disease Progression ; Drug development ; Humans ; Longitudinal Studies ; Mathematical models ; Models, Statistical ; Parkinson Disease - diagnosis ; Patients ; Polymorphism, Single Nucleotide ; Translation ; Translational Medical Research - methods ; Translational Medical Research - statistics & numerical data ; User-Computer Interface
- **Is Part Of:** Scientific reports, 2020-12-01, Vol.10 (1), p.10971-10971
- **Description:** Translational research of many disease areas requires a longitudinal understanding of disease development and progression across all biologically relevant scales. Several corresponding studies are now available. However, to compile a comprehensive picture of a specific disease, multiple studies need to be analyzed and compared. A large number of clinical studies is nowadays conducted in

the context of drug development in pharmaceutical research. However, legal and ethical constraints typically do not allow for sharing sensitive patient data. In consequence there exist data "silos", which slow down the overall scientific progress in translational research. In this paper, we suggest the idea of a virtual cohort (VC) to address this limitation. Our key idea is to describe a longitudinal patient cohort with the help of a generative statistical model, namely a modular Bayesian Network, in which individual modules are represented as sparse autoencoder networks. We show that with the help of such a model we can simulate subjects that are highly similar to real ones. Our approach allows for incorporating arbitrary multi-scale, multi-modal data without making specific distribution assumptions. Moreover, we demonstrate the possibility to simulate interventions (e.g. via a treatment) in the VC. Overall, our proposed approach opens the possibility to build sufficiently realistic VCs for multiple disease areas in the future.
- **Publisher:** England: Nature Publishing Group
- **Language:** English
- **Identifier:** ISSN: 2045-2322; EISSN: 2045-2322; DOI: 10.1038/s41598-020-67398-4; PMID: 32620927
- **Source:** Nature Open Access; Biological Science Collection; Scopus; Publicly Available Content Database; Academic Search Ultimate

Inter-species prediction of protein phosphorylation in the sbv IMPROVER species translation challenge

- **Creator:** Biehl, Michael ; Sadowski, Peter ; Bhanot, Gyan ; Bilal, Erhan ; Dayarian, Adel ; Meyer, Pablo ; Norel, Raquel ; Rhrissorrakrai, Kahn ; Zeller, Michael D ; Hormoz, Sahand
- **Subjects:** Algorithms ; Animals ; Bronchi - cytology ; Bronchi - metabolism ; Cells, Cultured ; Databases, Factual ; Epithelial Cells - cytology ; Epithelial Cells - metabolism ; Gene Expression Profiling ; Gene Expression Regulation ; Humans ; Improver Challenge Special Issue; Species Translation Challenge ; Oligonucleotide Array Sequence Analysis ; Phosphoproteins - metabolism ; Phosphorylation ; Rats ; Software ; Species Specificity ; Systems Biology - methods ; Translational Medical Research
- **Is Part Of:** Bioinformatics, 2015-02-15, Vol.31 (4), p.453-461
- **Description:** Animal models are widely used in biomedical research for reasons ranging from practical to ethical. An important issue is whether rodent models are predictive of human biology. This has been addressed recently in the framework of a series of challenges designed by the systems biology verification for Industrial Methodology for Process Verification in Research (sbv IMPROVER) initiative. In particular, one of the sub-challenges was devoted to the prediction of protein

phosphorylation responses in human bronchial epithelial cells, exposed to a number of different chemical stimuli, given the responses in rat bronchial epithelial cells. Participating teams were asked to make inter-species predictions on the basis of available training examples, comprising transcriptomics and phosphoproteomics data. Here, the two best performing teams present their data-driven approaches and computational methods. In addition, post hoc analyses of the datasets and challenge results were performed by the participants and challenge organizers. The challenge outcome indicates that successful prediction of protein phosphorylation status in human based on rat phosphorylation levels is feasible. However, within the limitations of the computational tools used, the inclusion of gene expression data does not improve the prediction quality. The post hoc analysis of time-specific measurements sheds light on the signaling pathways in both species. A detailed description of the dataset, challenge design and outcome is available at www.sbvimprover.com. The code used by team IGB is provided under http://github.com/uci-igb/improver2013. Implementations of the algorithms applied by team AMG are available at http://bhanot.biomaps.rutgers.edu/wiki/AMG-sc2-code.zip. meikelbiehl@gmail.com.
- **Publisher:** England: Oxford University Press
- **Language:** English
- **Identifier:** ISSN: 1367-4803; EISSN: 1460-2059; EISSN: 1367-4811; DOI: 10.1093/bioinformatics/btu407; PMID: 24994890
- **Source:** Scopus; Medline Complete; Alma/SFX Local Collection

[Need to Knowledge (NtK) Model: An evidence-based framework for generating technological innovations with socio-economic impacts]

- **Creator:** Flagg, Jennifer L ; Lane, Joseph P ; Lockett, Michelle M
- **Subjects:** Analysis ; Biomedical Engineering ; Biomedical Research - organization & administration ; Commercial transaction ; Debate ; Engineering development ; Evidence-based ; Evidence-Based Medicine ; Health Care Sciences & Services ; Health Care Sector - organization & administration ; Health Policy & Services ; Humans ; Industrial production ; Innovation ; Inventions - economics ; Inventions - utilization ; Knowledge translation ; Life Sciences & Biomedicine ; Models, Theoretical ; Research ; Research grants ; Science & Technology ; Scientific research ; Socioeconomic Factors ; States of knowledge ; Technological innovations ; Technology transfer ; Technology-based ; Translational Medical Research
- **Is Part Of:** Implementation science : IS, 2013-02-15, Vol.8 (1), p.21-21
- **Description:** Background: Traditional government policies suggest that upstream investment in scientific research is necessary and sufficient to generate technological innovations. The expected downstream beneficial socio-economic impacts are

presumed to occur through non-government market mechanisms. However, there is little quantitative evidence for such a direct and formulaic relationship between public investment at the input end and marketplace benefits at the impact end. Instead, the literature demonstrates that the technological innovation process involves a complex interaction between multiple sectors, methods, and stakeholders. Discussion: The authors theorize that accomplishing the full process of technological innovation in a deliberate and systematic manner requires an operational-level model encompassing three underlying methods, each designed to generate knowledge outputs in different states: scientific research generates conceptual discoveries; engineering development generates prototype inventions; and industrial production generates commercial innovations. Given the critical roles of engineering and business, the entire innovation process should continuously consider the practical requirements and constraints of the commercial marketplace. The Need to Knowledge (NtK) Model encompasses the activities required to successfully generate innovations, along with associated strategies for effectively communicating knowledge outputs in all three states to the various stakeholders involved. It is intentionally grounded in evidence drawn from academic analysis to facilitate objective and quantitative scrutiny, and industry best practices to enable practical application. Summary: The Need to Knowledge (NtK) Model offers a practical, market-oriented approach that avoids the gaps, constraints and inefficiencies inherent in undirected activities and disconnected sectors. The NtK Model is a means to realizing increased returns on public investments in those science and technology programs expressly intended to generate beneficial socio-economic impacts.

- **Publisher:** LONDON: BMC
- **Language:** English
- **Identifier:** ISSN: 1748-5908; EISSN: 1748-5908; DOI: 10.1186/1748-5908-8-21; PMID: 23414369
- **Source:** Gale Academic OneFile; Gale OneFile: Health and Medicine; Scopus; Gale OneFile: Nursing and Allied Health; Publicly Available Content Database; BioMedCentral Open Access; Academic Search Ultimate; Web of Science - Social Sciences Citation Index - 2013 from WEB OF SCIENCE; Web of Science - Science Citation Index Expanded - 2013 from WEB OF SCIENCE; PubMed Central; DOAJ Directory of Open Access Journals - Not for CDI Discovery

Gut microbiome, big data and machine learning to promote precision medicine for cancer

- **Creator:** Cammarota, Giovanni ; Ianiro, Gianluca ; Ahern, Anna ; Carbone, Carmine ; Temko, Andriy ; Claesson, Marcus J ; Gasbarrini, Antonio ; Tortora, Giampaolo
- **Is Part Of:** Nature reviews. Gastroenterology & hepatology, 2020-07-09, Vol.17 (10), p.635-648

- **Language:** English
- **Identifier:** ISSN: 1759-5045; EISSN: 1759-5053; DOI: 10.1038/s41575-020-0327-3
- **Source:** Gale Academic OneFile; Biological Science Collection; Gale OneFile: Health and Medicine; Gale Health and Wellness; Gale OneFile: Nursing and Allied Health

Classifying publications from the clinical and translational science award program along the translational research spectrum: A machine learning approach

- **Creator:** Surkis, Alisa ; Hogle, Janice A ; DiazGranados, Deborah ; Hunt, Joe D ; Mazmanian, Paul E ; Connors, Emily ; Westaby, Kate ; Whipple, Elizabeth C ; Adamus, Trisha ; Mueller, Meridith ; Aphinyanaphongs, Yindalon
- **Subjects:** Algorithms ; Analysis ; Area Under Curve ; Documentation ; Knowledge translation ; Life Sciences & Biomedicine ; Machine Learning ; Medicine, Research & Experimental ; Publications - classification ; Research ; Research & Experimental Medicine ; Science & Technology ; Text classification ; Translational Medical Research ; Translational research ; Usage
- **Is Part Of:** Journal of translational medicine, 2016-08-05, Vol.14 (1), p.235-235
- **Description:** Background: Translational research is a key area of focus of the National Institutes of Health (NIH), as demonstrated by the substantial investment in the Clinical and Translational Science Award (CTSA) program. The goal of the CTSA program is to accelerate the translation of discoveries from the bench to the bedside and into communities. Different classification systems have been used to capture the spectrum of basic to clinical to population health research, with substantial differences in the number of categories and their definitions. Evaluation of the effectiveness of the CTSA program and of translational research in general is hampered by the lack of rigor in these definitions and their application. This study adds rigor to the classification process by creating a checklist to evaluate publications across the translational spectrum and operationalizes these classifications by building machine learning-based text classifiers to categorize these publications. Methods: Based on collaboratively developed definitions, we created a detailed checklist for categories along the translational spectrum from T0 to T4. We applied the checklist to CTSA-linked publications to construct a set of coded publications for use in training machine learning-based text classifiers to classify publications within these categories. The training sets combined T1/T2 and T3/T4 categories due to low frequency of these publication types compared to the frequency of T0 publications. We then compared classifier performance across different algorithms and feature sets and applied the classifiers to all publications in PubMed indexed to CTSA grants. To validate the algorithm, we manually classified the articles with the top 100 scores from each classifier. Results: The definitions and checklist facilitated classification

and resulted in good inter-rater reliability for coding publications for the training set. Very good performance was achieved for the classifiers as represented by the area under the receiver operating curves (AUC), with an AUC of 0.94 for the T0 classifier, 0.84 for T1/T2, and 0.92 for T3/T4. Conclusions: The combination of definitions agreed upon by five CTSA hubs, a checklist that facilitates more uniform definition interpretation, and algorithms that perform well in classifying publications along the translational spectrum provide a basis for establishing and applying uniform definitions of translational research categories. The classification algorithms allow publication analyses that would not be feasible with manual classification, such as assessing the distribution and trends of publications across the CTSA network and comparing the categories of publications and their citations to assess knowledge transfer across the translational research spectrum.

- **Publisher:** LONDON: BMC
- **Language:** English
- **Identifier:** ISSN: 1479-5876; EISSN: 1479-5876; DOI: 10.1186/s12967-016-0992-8; PMID: 27492440
- **Source:** Gale Academic OneFile; Gale OneFile: Health and Medicine; Scopus; Gale OneFile: Nursing and Allied Health; Publicly Available Content Database; BioMedCentral Open Access; Academic Search Ultimate; Web of Science - Social Sciences Citation Index - 2016 from WEB OF SCIENCE; Web of Science - Science Citation Index Expanded - 2016 from WEB OF SCIENCE; PubMed Central; Alma/SFX Local Collection; DOAJ Directory of Open Access Journals - Not for CDI Discovery