

# Hongfang Liu

## List of Publications by Year in Descending Order

**Source:** <https://exaly.com/author-pdf/8411794/hongfang-liu-publications-by-year.pdf>

**Version:** 2024-04-27

This document has been generated based on the publications and citations recorded by exaly.com. For the latest version of this publication list, visit the link given above.

The third column is the impact factor (IF) of the journal, and the fourth column is the number of citations of the article.

234  
papers

3,801  
citations

31  
h-index

51  
g-index

294  
ext. papers

5,295  
ext. citations

4.4  
avg, IF

5.77  
L-index

| #   | Paper   | IF  | Citations |
|-----|---|-----|-----------|
| 234 | Identifying Information Gaps in Electronic Health Records by Using Natural Language Processing: Gynecologic Surgery History Identification.. <i>Journal of Medical Internet Research</i> , <b>2022</b> , 24, e29015   | 7.6 | 1         |
| 233 | CancerBERT: a cancer domain-specific language model for extracting breast cancer phenotypes from electronic health records.. <i>Journal of the American Medical Informatics Association: JAMIA</i> , <b>2022</b> ,  | 8.6 | 5         |
| 232 | Risk of pneumonia in asthmatic children using inhaled corticosteroids: a nested case-control study in a birth cohort.. <i>BMJ Open</i> , <b>2022</b> , 12, e051926  | 3   |           |
| 231 | A hybrid model to identify fall occurrence from electronic health records.. <i>International Journal of Medical Informatics</i> , <b>2022</b> , 162, 104736   | 5.3 | 1         |
| 230 | POKR: Building a Computable Heterogeneous Knowledge Resource for Precision Oncology. <i>Studies in Health Technology and Informatics</i> , <b>2022</b> ,  | 0.5 |           |
| 229 | The Implication of Latent Information Quality to the Reproducibility of Secondary Use of Electronic Health Records. <i>Studies in Health Technology and Informatics</i> , <b>2022</b> ,   | 0.5 |           |
| 228 | Establishing an expert consensus for the operational definitions of asthma-associated infectious and inflammatory multimorbidities for computational algorithms through a modified Delphi technique. <i>BMC Medical Informatics and Decision Making</i> , <b>2021</b> , 21, 310 | 3.6 |           |
| 227 | Early Detection of Post-Surgical Complications using Time-series Electronic Health Records. <i>AMIA Summits on Translational Science Proceedings</i> , <b>2021</b> , 2021, 152-160  | 1.1 |           |
| 226 | An Examination of the Statistical Laws of Semantic Change in Clinical Notes. <i>AMIA Summits on Translational Science Proceedings</i> , <b>2021</b> , 2021, 515-524   | 1.1 |           |
| 225 | Characterizing Long COVID: Deep Phenotype of a Complex Condition. <i>EBioMedicine</i> , <b>2021</b> , 74, 103722  | 8.8 | 25        |
| 224 | Deployment of an Interdisciplinary Predictive Analytics Task Force to Inform Hospital Operational Decision-Making During the COVID-19 Pandemic. <i>Mayo Clinic Proceedings</i> , <b>2021</b> , 96, 690-698  | 6.4 | 5         |
| 223 | Natural Language Processing and Machine Learning for Identifying Incident Stroke From Electronic Health Records: Algorithm Development and Validation. <i>Journal of Medical Internet Research</i> , <b>2021</b> , 23, e22951   | 7.6 | 2         |
| 222 | Agreement between neuroimages and reports for natural language processing-based detection of silent brain infarcts and white matter disease. <i>BMC Neurology</i> , <b>2021</b> , 21, 189   | 3.1 | 0         |
| 221 | Probing Patient Messages Enhanced by Natural Language Processing: A Top-Down Message Corpus Analysis. <i>Health Data Science</i> , <b>2021</b> , 2021, 1-10   |     | 1         |
| 220 | Longitudinal cohorts for harnessing the electronic health record for disease prediction in a US population. <i>BMJ Open</i> , <b>2021</b> , 11, e044353   | 3   | 2         |
| 219 | Are synthetic clinical notes useful for real natural language processing tasks: A case study on clinical entity recognition. <i>Journal of the American Medical Informatics Association: JAMIA</i> , <b>2021</b> , 28, 2193-2201  | 8.6 | 3         |
| 218 | Using ensemble of ensemble machine learning methods to predict outcomes of cardiac resynchronization. <i>Journal of Cardiovascular Electrophysiology</i> , <b>2021</b> , 32, 2504-2514  | 2.7 | 3         |

|     |   |      |    |
|-----|---|------|----|
| 217 | An aberration detection-based approach for sentinel syndromic surveillance of COVID-19 and other novel influenza-like illnesses. <i>Journal of Biomedical Informatics</i> , <b>2021</b> , 113, 103660   | 10.2 | 6  |
| 216 | Modeling cancer clinical trials using HL7 FHIR to support downstream applications: A case study with colorectal cancer data. <i>International Journal of Medical Informatics</i> , <b>2021</b> , 145, 104308  | 5.3  | 2  |
| 215 | Digital Pathology-based Study of Cell- and Tissue-level Morphologic Features in Serous Borderline Ovarian Tumor and High-grade Serous Ovarian Cancer. <i>Journal of Pathology Informatics</i> , <b>2021</b> , 12, 24  | 4.4  | 1  |
| 214 | Family History Extraction From Synthetic Clinical Narratives Using Natural Language Processing: Overview and Evaluation of a Challenge Data Set and Solutions for the 2019 National NLP Clinical Challenges (n2c2)/Open Health Natural Language Processing (OHNLP) Competition. <i>JMIR Medical Informatics</i> , <b>2021</b> , 9, e24008 | 3.6  | 4  |
| 213 | Recommendations for the safe, effective use of adaptive CDS in the US healthcare system: an AMIA position paper. <i>Journal of the American Medical Informatics Association: JAMIA</i> , <b>2021</b> , 28, 677-684  | 8.6  | 19 |
| 212 | Risk, Mechanisms and Implications of Asthma-Associated Infectious and Inflammatory Multimorbidities (AIMs) among Individuals With Asthma: a Systematic Review and a Case Study. <i>Allergy, Asthma and Immunology Research</i> , <b>2021</b> , 13, 697-718  | 5.3  | 0  |
| 211 | Quantifying the Importance of COVID-19 Vaccination to Our Future Outlook. <i>Mayo Clinic Proceedings</i> , <b>2021</b> , 96, 1890-1895  | 6.4  | 3  |
| 210 | Association of Silent Cerebrovascular Disease Identified Using Natural Language Processing and Future Ischemic Stroke. <i>Neurology</i> , <b>2021</b> , 97, e1313-e1321   | 6.5  | 1  |
| 209 | Artificial intelligence-assisted clinical decision support for childhood asthma management: A randomized clinical trial. <i>PLoS ONE</i> , <b>2021</b> , 16, e0255261   | 3.7  | 1  |
| 208 | Generating real-world evidence from unstructured clinical notes to examine clinical utility of genetic tests: use case in BRCAness. <i>BMC Medical Informatics and Decision Making</i> , <b>2021</b> , 21, 3  | 3.6  | 1  |
| 207 | Ascertainment of delirium status using natural language processing from electronic health records. <i>Journals of Gerontology - Series A Biological Sciences and Medical Sciences</i> , <b>2020</b> ,   | 6.4  | 3  |
| 206 | Developing an FHIR-Based Computational Pipeline for Automatic Population of Case Report Forms for Colorectal Cancer Clinical Trials Using Electronic Health Records. <i>JCO Clinical Cancer Informatics</i> , <b>2020</b> , 4, 201-209  | 5.2  | 10 |
| 205 | Constructing co-occurrence network embeddings to assist association extraction for COVID-19 and other coronavirus infectious diseases. <i>Journal of the American Medical Informatics Association: JAMIA</i> , <b>2020</b> , 27, 1259-1267  | 8.6  | 6  |
| 204 | COVID-19 TestNorm: A tool to normalize COVID-19 testing names to LOINC codes. <i>Journal of the American Medical Informatics Association: JAMIA</i> , <b>2020</b> , 27, 1437-1442   | 8.6  | 8  |
| 203 | Expert artificial intelligence-based natural language processing characterises childhood asthma. <i>BMJ Open Respiratory Research</i> , <b>2020</b> , 7,  | 5.6  | 10 |
| 202 | Impact of Diverse Data Sources on Computational Phenotyping. <i>Frontiers in Genetics</i> , <b>2020</b> , 11, 556   | 4.5  | 1  |
| 201 | Investigating the impact of disease and health record duration on the eMERGE algorithm for rheumatoid arthritis. <i>Journal of the American Medical Informatics Association: JAMIA</i> , <b>2020</b> , 27, 601-605  | 8.6  | 2  |
| 200 | Characterizing Chronic Pain Episodes in Clinical Text at Two Health Care Systems: Comprehensive Annotation and Corpus Analysis. <i>JMIR Medical Informatics</i> , <b>2020</b> , 8, e18659   | 3.6  |    |

|     |   |      |    |
|-----|---|------|----|
| 199 | Automating the Transformation of Free-Text Clinical Problems into SNOMED CT Expressions. <i>AMIA Summits on Translational Science Proceedings</i> , <b>2020</b> , 2020, 497-506   | 1.1  | 1  |
| 198 | Natural Language Processing for the Evaluation of Methodological Standards and Best Practices of EHR-based Clinical Research. <i>AMIA Summits on Translational Science Proceedings</i> , <b>2020</b> , 2020, 171-180                                    | 1.1  | 1  |
| 197 | Data-driven Sublanguage Analysis for Cancer Genomics Knowledge Modeling: Applications in Mining Oncological Genetics Information from Patients' Genetic Reports. <i>AMIA Summits on Translational Science Proceedings</i> , <b>2020</b> , 2020, 720-729 | 1.1  | 1  |
| 196 | Detecting and Filtering Immune-Related Adverse Events Signal Based on Text Mining and Observational Health Data Sciences and Informatics Common Data Model: Framework Development Study. <i>JMIR Medical Informatics</i> , <b>2020</b> , 8, e17353      | 3.6  | 4  |
| 195 | Implementation of a Cohort Retrieval System for Clinical Data Repositories Using the Observational Medical Outcomes Partnership Common Data Model: Proof-of-Concept System Validation. <i>JMIR Medical Informatics</i> , <b>2020</b> , 8, e17376        | 3.6  | 5  |
| 194 | The 2019 n2c2/OHNLNLP Track on Clinical Semantic Textual Similarity: Overview. <i>JMIR Medical Informatics</i> , <b>2020</b> , 8, e23375  | 3.6  | 11 |
| 193 | An Aberration Detection-Based Approach for Sentinel Syndromic Surveillance of COVID-19 and Other Novel Influenza-Like Illnesses <b>2020</b> ,   |      | 4  |
| 192 | Artificial intelligence approaches using natural language processing to advance EHR-based clinical research. <i>Journal of Allergy and Clinical Immunology</i> , <b>2020</b> , 145, 463-469   | 11.5 | 40 |
| 191 | Unsupervised machine learning for the discovery of latent disease clusters and patient subgroups using electronic health records. <i>Journal of Biomedical Informatics</i> , <b>2020</b> , 102, 103364  | 10.2 | 24 |
| 190 | A corpus-driven standardization framework for encoding clinical problems with HL7 FHIR. <i>Journal of Biomedical Informatics</i> , <b>2020</b> , 110, 103541  | 10.2 | 6  |
| 189 | Correlations Between COVID-19 Cases and Google Trends Data in the United States: A State-by-State Analysis. <i>Mayo Clinic Proceedings</i> , <b>2020</b> , 95, 2370-2381  | 6.4  | 24 |
| 188 | Subgrouping Rare Disease Patients Leveraging the Human Phenotype Ontology Embeddings <b>2020</b> ,  |      | 1  |
| 187 | Evaluation of patient-level retrieval from electronic health record data for a cohort discovery task. <i>JAMIA Open</i> , <b>2020</b> , 3, 395-404  | 2.9  | 2  |
| 186 | Time event ontology (TEO): to support semantic representation and reasoning of complex temporal relations of clinical events. <i>Journal of the American Medical Informatics Association: JAMIA</i> , <b>2020</b> , 27, 1046-1056                       | 8.6  | 3  |
| 185 | Clinical concept extraction: A methodology review. <i>Journal of Biomedical Informatics</i> , <b>2020</b> , 109, 103526   | 10.2 | 30 |
| 184 | Enrich Rare Disease Phenotypic Characterizations via a Graph Convolutional Network Based Recommendation System <b>2020</b> ,  |      | 1  |
| 183 | MedSTS: a resource for clinical semantic textual similarity. <i>Language Resources and Evaluation</i> , <b>2020</b> , 54, 57-72   | 1.8  | 33 |
| 182 | Cohort Profile: The Right Drug, Right Dose, Right Time: Using Genomic Data to Individualize Treatment Protocol (RIGHT Protocol). <i>International Journal of Epidemiology</i> , <b>2020</b> , 49, 23-24k  | 7.8  | 19 |

|     |  |      |    |
|-----|--|------|----|
| 181 | Assessment of the impact of EHR heterogeneity for clinical research through a case study of silent brain infarction. <i>BMC Medical Informatics and Decision Making</i> , <b>2020</b> , 20, 60                                       | 3.6  | 9  |
| 180 | Test collections for electronic health record-based clinical information retrieval. <i>JAMIA Open</i> , <b>2019</b> , 2, 360-368   | 2.9  | 7  |
| 179 | Automatic extraction and assessment of lifestyle exposures for Alzheimer's disease using natural language processing. <i>International Journal of Medical Informatics</i> , <b>2019</b> , 130, 103943                                | 5.3  | 9  |
| 178 | A privacy-preserving distributed filtering framework for NLP artifacts. <i>BMC Medical Informatics and Decision Making</i> , <b>2019</b> , 19, 183   | 3.6  | 9  |
| 177 | Automated extraction of sudden cardiac death risk factors in hypertrophic cardiomyopathy patients by natural language processing. <i>International Journal of Medical Informatics</i> , <b>2019</b> , 128, 32-38                     | 5.3  | 15 |
| 176 | Deep learning and alternative learning strategies for retrospective real-world clinical data. <i>Npj Digital Medicine</i> , <b>2019</b> , 2, 43  | 15.7 | 91 |
| 175 | Early Identification of Childhood Asthma: The Role of Informatics in an Era of Electronic Health Records. <i>Frontiers in Pediatrics</i> , <b>2019</b> , 7, 113  | 3.4  | 3  |
| 174 | Natural language processing of radiology reports for identification of skeletal site-specific fractures. <i>BMC Medical Informatics and Decision Making</i> , <b>2019</b> , 19, 73   | 3.6  | 13 |
| 173 | Rare disease knowledge enrichment through a data-driven approach. <i>BMC Medical Informatics and Decision Making</i> , <b>2019</b> , 19, 32  | 3.6  | 16 |
| 172 | Saliency of Medical Concepts of Inside Clinical Texts and Outside Medical Records for Referred Cardiovascular Patients. <i>Journal of Healthcare Informatics Research</i> , <b>2019</b> , 3, 200-219                                 | 4    | 3  |
| 171 | HPO2Vec+: Leveraging heterogeneous knowledge resources to enrich node embeddings for the Human Phenotype Ontology. <i>Journal of Biomedical Informatics</i> , <b>2019</b> , 96, 103246   | 10.2 | 13 |
| 170 | Developing a FHIR-based EHR phenotyping framework: A case study for identification of patients with obesity and multiple comorbidities from discharge summaries. <i>Journal of Biomedical Informatics</i> , <b>2019</b> , 99, 103310 | 10.2 | 18 |
| 169 | Ensembles of natural language processing systems for portable phenotyping solutions. <i>Journal of Biomedical Informatics</i> , <b>2019</b> , 100, 103318  | 10.2 | 9  |
| 168 | Developing a scalable FHIR-based clinical data normalization pipeline for standardizing and integrating unstructured and structured electronic health record data. <i>JAMIA Open</i> , <b>2019</b> , 2, 570-579                      | 2.9  | 14 |
| 167 | Discovering associations between problem list and practice setting. <i>BMC Medical Informatics and Decision Making</i> , <b>2019</b> , 19, 69  | 3.6  | 3  |
| 166 | Evaluating the Impact of Dictionary Updates on Automatic Annotations Based on Clinical NLP Systems. <i>AMIA Summits on Translational Science Proceedings</i> , <b>2019</b> , 2019, 714-721   | 1.1  | 3  |
| 165 | Coverage Evaluation of CTCAE for Capturing the Immune-related Adverse Events Leveraging Text Mining Technologies. <i>AMIA Summits on Translational Science Proceedings</i> , <b>2019</b> , 2019, 771-778                             | 1.1  | 8  |
| 164 | Developing Customizable Cancer Information Extraction Modules for Pathology Reports Using CLAMP. <i>Studies in Health Technology and Informatics</i> , <b>2019</b> , 264, 1041-1045  | 0.5  | 2  |

|     |  |      |     |
|-----|--|------|-----|
| 163 | Natural Language Processing for the Identification of Silent Brain Infarcts From Neuroimaging Reports. <i>JMIR Medical Informatics</i> , <b>2019</b> , 7, e12109   | 3.6  | 18  |
| 162 | Detecting Lifestyle Risk Factors for Chronic Kidney Disease With Comorbidities: Association Rule Mining Analysis of Web-Based Survey Data. <i>Journal of Medical Internet Research</i> , <b>2019</b> , 21, e14204                                  | 7.6  | 5   |
| 161 | Artificial intelligence to organize patient portal messages: a journey from an ensemble deep learning text classification to rule-based named entity recognition <b>2019</b> ,   |      | 3   |
| 160 | Desiderata for delivering NLP to accelerate healthcare AI advancement and a Mayo Clinic NLP-as-a-service implementation. <i>Npj Digital Medicine</i> , <b>2019</b> , 2, 130  | 15.7 | 26  |
| 159 | Integrating word embedding neural networks with PubMed abstracts to extract keyword proximity of chronic diseases <b>2019</b> ,  |      | 3   |
| 158 | Natural language processing for populating lung cancer clinical research data. <i>BMC Medical Informatics and Decision Making</i> , <b>2019</b> , 19, 239  | 3.6  | 6   |
| 157 | Detection of Surgical Site Infection Utilizing Automated Feature Generation in Clinical Notes. <i>Journal of Healthcare Informatics Research</i> , <b>2019</b> , 3, 267-282  | 4    | 6   |
| 156 | A clinical text classification paradigm using weak supervision and deep representation. <i>BMC Medical Informatics and Decision Making</i> , <b>2019</b> , 19, 1   | 3.6  | 130 |
| 155 | Relationship Between Very Cold Outside Weather and Surgical Outcome: Integrating Shallow and Deep Artificial Neural Nets. <i>Studies in Health Technology and Informatics</i> , <b>2019</b> , 264, 1783-1784                                       | 0.5  | 1   |
| 154 | Clinical information extraction applications: A literature review. <i>Journal of Biomedical Informatics</i> , <b>2018</b> , 77, 34-49  | 10.2 | 282 |
| 153 | Ascertainment of asthma prognosis using natural language processing from electronic medical records. <i>Journal of Allergy and Clinical Immunology</i> , <b>2018</b> , 141, 2292-2294.e3   | 11.5 | 15  |
| 152 | Clinical documentation variations and NLP system portability: a case study in asthma birth cohorts across institutions. <i>Journal of the American Medical Informatics Association: JAMIA</i> , <b>2018</b> , 25, 353-359                          | 8.6  | 35  |
| 151 | CLAMP - a toolkit for efficiently building customized clinical natural language processing pipelines. <i>Journal of the American Medical Informatics Association: JAMIA</i> , <b>2018</b> , 25, 331-336  | 8.6  | 137 |
| 150 | Impact of Patient Reminders on Papanicolaou Test Completion for High-Risk Patients Identified by a Clinical Decision Support System. <i>Journal of Women's Health</i> , <b>2018</b> , 27, 569-574  | 3    | 3   |
| 149 | Use of echocardiography in outpatients with chest pain and normal resting electrocardiograms referred to Mayo Clinic Rochester. <i>American Heart Journal</i> , <b>2018</b> , 196, 49-55   | 4.9  | 7   |
| 148 | Natural language processing of clinical notes for identification of critical limb ischemia. <i>International Journal of Medical Informatics</i> , <b>2018</b> , 111, 83-89   | 5.3  | 42  |
| 147 | Automated chart review utilizing natural language processing algorithm for asthma predictive index. <i>BMC Pulmonary Medicine</i> , <b>2018</b> , 18, 34   | 3.5  | 30  |
| 146 | Innovative Informatics Approaches for Peripheral Artery Disease: Current State and Provider Survey of Strategies for Improving Guideline-Based Care. <i>Mayo Clinic Proceedings Innovations, Quality &amp; Outcomes</i> , <b>2018</b> , 2, 129-136 | 3.1  | 8   |

|     |   |      |    |
|-----|---|------|----|
| 145 | Ten-Year Trends in Antiemetic Prescribing in Patients Receiving Highly Emetogenic Chemotherapy. <i>Journal of the National Comprehensive Cancer Network: JNCCN</i> , <b>2018</b> , 16, 294-299  | 7.3  | 9  |
| 144 | Natural Language Processing for Asthma Ascertainment in Different Practice Settings. <i>Journal of Allergy and Clinical Immunology: in Practice</i> , <b>2018</b> , 6, 126-131  | 5.4  | 29 |
| 143 | Early genetic aberrations in patients with sporadic colorectal cancer. <i>Molecular Carcinogenesis</i> , <b>2018</b> , 57, 114-124  | 5    | 16 |
| 142 | Using data-driven sublanguage pattern mining to induce knowledge models: application in medical image reports knowledge representation. <i>BMC Medical Informatics and Decision Making</i> , <b>2018</b> , 18, 61                                       | 3.6  | 7  |
| 141 | Postoperative bleeding risk prediction for patients undergoing colorectal surgery. <i>Surgery</i> , <b>2018</b> , 164, 1209-1216  | 3.6  | 16 |
| 140 | Constructing Node Embeddings for Human Phenotype Ontology to Assist Phenotypic Similarity Measurement <b>2018</b> ,   |      | 4  |
| 139 | Detecting Pharmacovigilance Signals Combining Electronic Medical Records With Spontaneous Reports: A Case Study of Conventional Disease-Modifying Antirheumatic Drugs for Rheumatoid Arthritis. <i>Frontiers in Pharmacology</i> , <b>2018</b> , 9, 875 | 5.6  | 13 |
| 138 | Modeling asynchronous event sequences with RNNs. <i>Journal of Biomedical Informatics</i> , <b>2018</b> , 83, 167-177   | 10.2 | 19 |
| 137 | Systematic identification of latent disease-gene associations from PubMed articles. <i>PLoS ONE</i> , <b>2018</b> , 13, e0191568  | 3.7  | 8  |
| 136 | Integrating Structured and Unstructured EHR Data Using an FHIR-based Type System: A Case Study with Medication Data. <i>AMIA Summits on Translational Science Proceedings</i> , <b>2018</b> , 2017, 74-83   | 1.1  | 10 |
| 135 | The Sublanguage of Clinical Problem Lists: A Corpus Analysis <b>2018</b> , 2018, 1451-1460  | 0.7  |    |
| 134 | Incorporating Knowledge-Driven Insights into a Collaborative Filtering Model to Facilitate the Differential Diagnosis of Rare Diseases <b>2018</b> , 2018, 1505-1514  | 0.7  | 2  |
| 133 | Standardizing Heterogeneous Annotation Corpora Using HL7 FHIR for Facilitating their Reuse and Integration in Clinical NLP <b>2018</b> , 2018, 574-583  | 0.7  | 6  |
| 132 | Utilization of Electronic Medical Records and Biomedical Literature to Support the Diagnosis of Rare Diseases Using Data Fusion and Collaborative Filtering Approaches. <i>JMIR Medical Informatics</i> , <b>2018</b> , 6, e11301                       | 3.6  | 16 |
| 131 | Assessing Unmet Information Needs of Breast Cancer Survivors: Exploratory Study of Online Health Forums Using Text Classification and Retrieval. <i>JMIR Cancer</i> , <b>2018</b> , 4, e10  | 3.2  | 14 |
| 130 | Privacy-Preserving Predictive Modeling: Harmonization of Contextual Embeddings From Different Sources. <i>JMIR Medical Informatics</i> , <b>2018</b> , 6, e33   | 3.6  | 8  |
| 129 | Burden of hospitalization in clinically diagnosed peripheral artery disease: A community-based study. <i>Vascular Medicine</i> , <b>2018</b> , 23, 23-31  | 3.3  | 8  |
| 128 | Electronic Health Record Phenotypes for Precision Medicine: Perspectives and Caveats From Treatment of Breast Cancer at a Single Institution. <i>Clinical and Translational Science</i> , <b>2018</b> , 11, 85-92                                       | 4.9  | 12 |

|     |   |      |     |
|-----|---|------|-----|
| 127 | BioCreative/OHNL Challenge 2018 <b>2018</b> ,   |      | 18  |
| 126 | Preliminary exploration of survival analysis using the OHDSI common data model: a case study of intrahepatic cholangiocarcinoma. <i>BMC Medical Informatics and Decision Making</i> , <b>2018</b> , 18, 116   | 3.6  | 4   |
| 125 | Leveraging Association Rule Mining to Detect Pathophysiological Mechanisms of Chronic Kidney Disease Complicated by Metabolic Syndrome <b>2018</b> ,  |      | 2   |
| 124 | Association of Ankle-Brachial Indices With Limb Revascularization or Amputation in Patients With Peripheral Artery Disease. <i>JAMA Network Open</i> , <b>2018</b> , 1, e185547   | 10.4 | 6   |
| 123 | Extracting chemical-protein relations using attention-based neural networks. <i>Database: the Journal of Biological Databases and Curation</i> , <b>2018</b> , 2018,  | 5    | 14  |
| 122 | A comparison of word embeddings for the biomedical natural language processing. <i>Journal of Biomedical Informatics</i> , <b>2018</b> , 87, 12-20  | 10.2 | 134 |
| 121 | Deep Phenotyping on Electronic Health Records Facilitates Genetic Diagnosis by Clinical Exomes. <i>American Journal of Human Genetics</i> , <b>2018</b> , 103, 58-73  | 11   | 55  |
| 120 | Mining peripheral arterial disease cases from narrative clinical notes using natural language processing. <i>Journal of Vascular Surgery</i> , <b>2017</b> , 65, 1753-1761  | 3.5  | 46  |
| 119 | Peripheral Neutrophil to Lymphocyte Ratio Improves Prognostication in Colon Cancer. <i>Clinical Colorectal Cancer</i> , <b>2017</b> , 16, 115-123.e3  | 3.8  | 27  |
| 118 | Application of a Natural Language Processing Algorithm to Asthma Ascertainment. An Automated Chart Review. <i>American Journal of Respiratory and Critical Care Medicine</i> , <b>2017</b> , 196, 430-437   | 10.2 | 42  |
| 117 | Ontology-based systematical representation and drug class effect analysis of package insert-reported adverse events associated with cardiovascular drugs used in China. <i>Scientific Reports</i> , <b>2017</b> , 7, 13819                              | 4.9  | 9   |
| 116 | On Mapping Textual Queries to a Common Data Model <b>2017</b> ,   |      | 2   |
| 115 | Dependency and AMR Embeddings for Drug-Drug Interaction Extraction from Biomedical Literature <b>2017</b> ,   |      | 10  |
| 114 | Need of informatics in designing interoperable clinical registries. <i>International Journal of Medical Informatics</i> , <b>2017</b> , 108, 78-84  | 5.3  | 12  |
| 113 | Intra-institutional EHR collections for patient-level information retrieval. <i>Journal of the Association for Information Science and Technology</i> , <b>2017</b> , 68, 2636-2648   | 2.7  | 5   |
| 112 | Proton Pump Inhibitors and the Risk for Fracture at Specific Sites: Data Mining of the FDA Adverse Event Reporting System. <i>Scientific Reports</i> , <b>2017</b> , 7, 5527  | 4.9  | 21  |
| 111 | BELMiner: adapting a rule-based relation extraction system to extract biological expression language statements from bio-medical literature evidence sentences. <i>Database: the Journal of Biological Databases and Curation</i> , <b>2017</b> , 2017, | 5    | 21  |
| 110 | Detection of clinically important colorectal surgical site infection using Bayesian network. <i>Journal of Surgical Research</i> , <b>2017</b> , 209, 168-173   | 2.5  | 29  |



|     |  |      |    |
|-----|--|------|----|
| 109 | Medical concept intersection between outside medical records and consultant notes: A case study in transferred cardiovascular patients <b>2017</b> ,   |      | 1  |
| 108 | Leveraging word embeddings and medical entity extraction for biomedical dataset retrieval using unstructured texts. <i>Database: the Journal of Biological Databases and Curation</i> , <b>2017</b> , 2017,  | 5    | 10 |
| 107 | Correlating Lab Test Results in Clinical Notes with Structured Lab Data: A Case Study in HbA1c and Glucose. <i>AMIA Summits on Translational Science Proceedings</i> , <b>2017</b> , 2017, 221-228   | 1.1  | 3  |
| 106 | Discovering adverse drug events combining spontaneous reports with electronic medical records: a case study of conventional DMARDs and biologics for rheumatoid arthritis. <i>AMIA Summits on Translational Science Proceedings</i> , <b>2017</b> , 2017, 95-103 | 1.1  | 4  |
| 105 | Systematic Analysis of Free-Text Family History in Electronic Health Record. <i>AMIA Summits on Translational Science Proceedings</i> , <b>2017</b> , 2017, 104-113  | 1.1  | 12 |
| 104 | Analysis of Clinical Variations in Asthma Care Documented in Electronic Health Records Between Staff and Resident Physicians. <i>Studies in Health Technology and Informatics</i> , <b>2017</b> , 245, 1170-1174   | 0.5  | 5  |
| 103 | Mining Hierarchies and Similarity Clusters from Value Set Repositories <b>2017</b> , 2017, 1372-1381   | 0.7  | 2  |
| 102 | Deep Learning Solutions for Classifying Patients on Opioid Use <b>2017</b> , 2017, 525-534   | 0.7  | 11 |
| 101 | Leveraging Collaborative Filtering to Accelerate Rare Disease Diagnosis <b>2017</b> , 2017, 1554-1563  | 0.7  | 11 |
| 100 | Distinction between medical and non-medical usages of short forms in clinical narratives <b>2017</b> , 2017, 1302-1311   | 0.7  | 1  |
| 99  | Phenotypic Analysis of Clinical Narratives Using Human Phenotype Ontology. <i>Studies in Health Technology and Informatics</i> , <b>2017</b> , 245, 581-585  | 0.5  | 12 |
| 98  | Recommending Education Materials for Diabetic Questions Using Information Retrieval Approaches. <i>Journal of Medical Internet Research</i> , <b>2017</b> , 19, e342   | 7.6  | 7  |
| 97  | Generalized ensemble model for document ranking in information retrieval. <i>Computer Science and Information Systems</i> , <b>2017</b> , 14, 123-151  | 0.8  | 2  |
| 96  | Using Human Phenotype Ontology for Phenotypic Analysis of Clinical Notes. <i>Studies in Health Technology and Informatics</i> , <b>2017</b> , 245, 1285  | 0.5  | 5  |
| 95  | Colorectal Cancer with Residual Polyp of Origin: A Model of Malignant Transformation. <i>Translational Oncology</i> , <b>2016</b> , 9, 280-6   | 4.9  | 7  |
| 94  | A computational framework for converting textual clinical diagnostic criteria into the quality data model. <i>Journal of Biomedical Informatics</i> , <b>2016</b> , 63, 11-21  | 10.2 | 6  |
| 93  | Toward a Learning Health-care System - Knowledge Delivery at the Point of Care Empowered by Big Data and NLP. <i>Biomedical Informatics Insights</i> , <b>2016</b> , 8, 13-22  | 4.9  | 41 |
| 92  | BELTracker: evidence sentence retrieval for BEL statements. <i>Database: the Journal of Biological Databases and Curation</i> , <b>2016</b> , 2016,  | 5    | 10 |

|    |   |      |    |
|----|---|------|----|
| 91 | Clinical element models in the SHARPN consortium. <i>Journal of the American Medical Informatics Association: JAMIA</i> , <b>2016</b> , 23, 248-56  | 8.6  | 14 |
| 90 | Exploring Gaps of Family History Documentation in EHR for Precision Medicine -A Case Study of Familial Hypercholesterolemia Ascertainment. <i>AMIA Summits on Translational Science Proceedings</i> , <b>2016</b> , 2016, 160-6   | 1.1  | 4  |
| 89 | Leveraging Terminology Services for Extract-Transform-Load Processes: A User-Centered Approach <b>2016</b> , 2016, 1010-1019  | 0.7  | 4  |
| 88 | A Topic-modeling Based Framework for Drug-drug Interaction Classification from Biomedical Text <b>2016</b> , 2016, 789-798  | 0.7  | 5  |
| 87 | A Text-Mining Framework for Supporting Systematic Reviews <b>2016</b> , 1, 1-9  |      | 4  |
| 86 | MayoNLP at SemEval-2016 Task 1: Semantic Textual Similarity based on Lexical Semantic Net and Deep Learning Semantic Model <b>2016</b> ,  |      | 11 |
| 85 | Using Social Media Data to Identify Potential Candidates for Drug Repurposing: A Feasibility Study. <i>JMIR Research Protocols</i> , <b>2016</b> , 5, e121  | 2    | 17 |
| 84 | Predicate Oriented Pattern Analysis for Biomedical Knowledge Discovery. <i>Intelligent Information Management</i> , <b>2016</b> , 8, 66-85  | 0.8  | 11 |
| 83 | Valx: A System for Extracting and Structuring Numeric Lab Test Comparison Statements from Text. <i>Methods of Information in Medicine</i> , <b>2016</b> , 55, 266-75  | 1.5  | 24 |
| 82 | Comparison of Three Information Sources for Smoking Information in Electronic Health Records. <i>Cancer Informatics</i> , <b>2016</b> , 15, 237-242   | 2.4  | 14 |
| 81 | Prioritizing Adverse Drug Reaction and Drug Repositioning Candidates Generated by Literature-Based Discovery <b>2016</b> ,  |      | 8  |
| 80 | Systematic Analysis of Adverse Event Reports for Sex Differences in Adverse Drug Events. <i>Scientific Reports</i> , <b>2016</b> , 6, 24955   | 4.9  | 50 |
| 79 | Automated Chart Review for Asthma Ascertainment: An Innovative Approach for Asthma Care and Research in the Era of Electronic Medical Record. <i>Journal of Allergy and Clinical Immunology</i> , <b>2016</b> , 137, AB196  | 11.5 | 2  |
| 78 | Identifying Peripheral Arterial Disease Cases Using Natural Language Processing of Clinical Notes. <i>IEEE-EMBS International Conference on Biomedical and Health Informatics</i> , <b>2016</b> , 2016, 126-131   | 1.9  | 12 |
| 77 | A Part-Of-Speech term weighting scheme for biomedical information retrieval. <i>Journal of Biomedical Informatics</i> , <b>2016</b> , 63, 379-389   | 10.2 | 23 |
| 76 | Time Lapse to Colorectal Cancer: Telomere Dynamics Define the Malignant Potential of Polyps. <i>Clinical and Translational Gastroenterology</i> , <b>2016</b> , 7, e188   | 4.2  | 7  |
| 75 | Rapid identification of familial hypercholesterolemia from electronic health records: The SEARCH study. <i>Journal of Clinical Lipidology</i> , <b>2016</b> , 10, 1230-9  | 4.9  | 62 |
| 74 | A Robust e-Epidemiology Tool in Phenotyping Heart Failure with Differentiation for Preserved and Reduced Ejection Fraction: the Electronic Medical Records and Genomics (eMERGE) Network. <i>Journal of Cardiovascular Translational Research</i> , <b>2015</b> , 8, 475-83 | 3.3  | 25 |

|    |   |      |     |
|----|---|------|-----|
| 73 | DEEPEN: A negation detection system for clinical text incorporating dependency relation into NegEx. <i>Journal of Biomedical Informatics</i> , <b>2015</b> , 54, 213-9  | 10.2 | 49  |
| 72 | Mining severe drug-drug interaction adverse events using Semantic Web technologies: a case study. <i>BioData Mining</i> , <b>2015</b> , 8, 12   | 4.3  | 9   |
| 71 | A bibliometric analysis on tobacco regulation investigators. <i>BioData Mining</i> , <b>2015</b> , 8, 11  | 4.3  | 5   |
| 70 | Text mining facilitates database curation - extraction of mutation-disease associations from Bio-medical literature. <i>BMC Bioinformatics</i> , <b>2015</b> , 16, 185  | 3.6  | 16  |
| 69 | Ease of adoption of clinical natural language processing software: An evaluation of five systems. <i>Journal of Biomedical Informatics</i> , <b>2015</b> , 58 Suppl, S189-S196  | 10.2 | 18  |
| 68 | Validating drug repurposing signals using electronic health records: a case study of metformin associated with reduced cancer mortality. <i>Journal of the American Medical Informatics Association: JAMIA</i> , <b>2015</b> , 22, 179-91 | 8.6  | 139 |
| 67 | A frequency-filtering strategy of obtaining PHI-free sentences from clinical data repository <b>2015</b> ,  |      | 5   |
| 66 | Use of Echocardiography in Olmsted County Outpatients With Chest Pain and Normal Resting Electrocardiograms Seen at Mayo Clinic Rochester. <i>Mayo Clinic Proceedings</i> , <b>2015</b> , 90, 1492-8                                      | 6.4  | 4   |
| 65 | Next generation informatics for big data in precision medicine era. <i>BioData Mining</i> , <b>2015</b> , 8, 34   | 4.3  | 6   |
| 64 | Retrieval of Semantically Similar Healthcare Questions in Healthcare Forums <b>2015</b> ,   |      | 2   |
| 63 | A new method for prioritizing drug repositioning candidates extracted by literature-based discovery <b>2015</b> ,   |      | 18  |
| 62 | <b>2015</b> ,   |      | 6   |
| 61 | Granular Quality Reporting for Cervical Cytology Testing. <i>AMIA Summits on Translational Science Proceedings</i> , <b>2015</b> , 2015, 178-82   | 1.1  | 2   |
| 60 | A Bibliometric Analysis on Cancer Population Science with Topic Modeling. <i>AMIA Summits on Translational Science Proceedings</i> , <b>2015</b> , 2015, 102-6  | 1.1  | 1   |
| 59 | Drug Normalization for Cancer Therapeutic and Druggable Genome Target Discovery. <i>AMIA Summits on Translational Science Proceedings</i> , <b>2015</b> , 2015, 72-6  | 1.1  | 1   |
| 58 | Operationalizing Semantic Medline for meeting the information needs at point of care. <i>AMIA Summits on Translational Science Proceedings</i> , <b>2015</b> , 2015, 152-6  | 1.1  | 1   |
| 57 | Assessing the Need of Discourse-Level Analysis in Identifying Evidence of Drug-Disease Relations in Scientific Literature. <i>Studies in Health Technology and Informatics</i> , <b>2015</b> , 216, 539-43                                | 0.5  |     |
| 56 | Identification of Patients with Family History of Pancreatic Cancer--Investigation of an NLP System Portability. <i>Studies in Health Technology and Informatics</i> , <b>2015</b> , 216, 604-8   | 0.5  | 17  |

|    |  |      |    |
|----|--|------|----|
| 55 | Network-based analysis reveals distinct association patterns in a semantic MEDLINE-based drug-disease-gene network. <i>Journal of Biomedical Semantics</i> , <b>2014</b> , 5, 33   | 2.2  | 20 |
| 54 | Standardizing adverse drug event reporting data. <i>Journal of Biomedical Semantics</i> , <b>2014</b> , 5, 36  | 2.2  | 17 |
| 53 | Using large clinical corpora for query expansion in text-based cohort identification. <i>Journal of Biomedical Informatics</i> , <b>2014</b> , 49, 275-81  | 10.2 | 24 |
| 52 | Automated recommendation for cervical cancer screening and surveillance. <i>Cancer Informatics</i> , <b>2014</b> , 13, 1-6   | 2.4  | 3  |
| 51 | Overview of the gene ontology task at BioCreative IV. <i>Database: the Journal of Biological Databases and Curation</i> , <b>2014</b> , 2014,  | 5    | 42 |
| 50 | Patient-level temporal aggregation for text-based asthma status ascertainment. <i>Journal of the American Medical Informatics Association: JAMIA</i> , <b>2014</b> , 21, 876-84  | 8.6  | 17 |
| 49 | MedXN: an open source medication extraction and normalization tool for clinical text. <i>Journal of the American Medical Informatics Association: JAMIA</i> , <b>2014</b> , 21, 858-65                                       | 8.6  | 63 |
| 48 | Integrating information retrieval with distant supervision for gene ontology annotation. <i>Database: the Journal of Biological Databases and Curation</i> , <b>2014</b> , 2014,   | 5    | 2  |
| 47 | Genome-wide analysis of loss of heterozygosity in breast infiltrating ductal carcinoma distant normal tissue highlights arm specific enrichment and expansion across tumor stages. <i>PLoS ONE</i> , <b>2014</b> , 9, e95783 | 3.7  | 3  |
| 46 | Towards pathway curation through literature mining--a case study using PharmGKB. <i>Pacific Symposium on Biocomputing Pacific Symposium on Biocomputing</i> , <b>2014</b> , 352-63   | 1.3  | 5  |
| 45 | Facilitating post-surgical complication detection through sublanguage analysis. <i>AMIA Summits on Translational Science Proceedings</i> , <b>2014</b> , 2014, 77-82   | 1.1  | 4  |
| 44 | Towards Transforming Expert-based Content to Evidence-based Content. <i>AMIA Summits on Translational Science Proceedings</i> , <b>2014</b> , 2014, 83-90  | 1.1  | 4  |
| 43 | Discovering associations among diagnosis groups using topic modeling. <i>AMIA Summits on Translational Science Proceedings</i> , <b>2014</b> , 2014, 43-9  | 1.1  | 5  |
| 42 | Analysis of medication and indication occurrences in clinical notes <b>2014</b> , 2014, 1046-55  | 0.7  | 2  |
| 41 | An automated approach for ranking journals to help in clinician decision support <b>2014</b> , 2014, 757-66  | 0.7  |    |
| 40 | A common type system for clinical natural language processing. <i>Journal of Biomedical Semantics</i> , <b>2013</b> , 4, 1   | 2.2  | 36 |
| 39 | Automated chart review for asthma cohort identification using natural language processing: an exploratory study. <i>Annals of Allergy, Asthma and Immunology</i> , <b>2013</b> , 111, 364-9                                  | 3.2  | 49 |
| 38 | Analysis of cross-institutional medication description patterns in clinical narratives. <i>Biomedical Informatics Insights</i> , <b>2013</b> , 6, 7-16   | 4.9  | 8  |

|    |  |     |    |
|----|--|-----|----|
| 37 | <b>2013,</b>   |     | 5  |
| 36 | Identifying Abdominal Aortic Aneurysm Cases and Controls using Natural Language Processing of Radiology Reports. <i>AMIA Summits on Translational Science Proceedings</i> , <b>2013</b> , 2013, 249-53 | 1.1 | 15 |
| 35 | An information extraction framework for cohort identification using electronic health records. <i>AMIA Summits on Translational Science Proceedings</i> , <b>2013</b> , 2013, 149-53                   | 1.1 | 66 |
| 34 | Building a knowledge base of severe adverse drug events based on AERS reporting data using semantic web technologies. <i>Studies in Health Technology and Informatics</i> , <b>2013</b> , 192, 496-500 | 0.5 | 9  |
| 33 | Automatically Extracting Sentences from Medline Citations to Support Clinicians' Information Needs <b>2012,</b>  |     | 2  |
| 32 | Clinical Decision Support for Colonoscopy Surveillance Using Natural Language Processing <b>2012,</b>  |     | 4  |
| 31 | Ontology-Based Temporal Relation Modeling with MapReduce Latent Dirichlet Allocations for Big EHR Data <b>2012,</b>  |     | 1  |
| 30 | A hybrid approach to sentiment sentence classification in suicide notes. <i>Biomedical Informatics Insights</i> , <b>2012</b> , 5, 43-50   | 4.9 | 16 |
| 29 | Unified Medical Language System term occurrences in clinical notes: a large-scale corpus analysis. <i>Journal of the American Medical Informatics Association: JAMIA</i> , <b>2012</b> , 19, e149-56   | 8.6 | 50 |
| 28 | Using SNOMED-CT to encode summary level data - a corpus analysis. <i>AMIA Summits on Translational Science Proceedings</i> , <b>2012</b> , 2012, 30-7  | 1.1 | 9  |
| 27 | Feasibility of pooling annotated corpora for clinical concept extraction. <i>AMIA Summits on Translational Science Proceedings</i> , <b>2012</b> , 2012, 38  | 1.1 | 8  |
| 26 | Towards a semantic lexicon for clinical natural language processing <b>2012</b> , 2012, 568-76   | 0.7 | 10 |
| 25 | A study of transportability of an existing smoking status detection module across institutions <b>2012</b> , 2012, 577-86  | 0.7 | 33 |
| 24 | Classification Systems for Bacterial Protein-Protein Interaction Document Retrieval <b>2012</b> , 33-43  |     |    |
| 23 | Using machine learning for concept extraction on clinical documents from multiple data sources. <i>Journal of the American Medical Informatics Association: JAMIA</i> , <b>2011</b> , 18, 580-7        | 8.6 | 91 |
| 22 | Microarray probes and probe sets. <i>Frontiers in Bioscience - Elite</i> , <b>2010</b> , 2, 325-38   | 1.6 | 25 |
| 21 | Testing the theory of relative defect proneness for closed-source software. <i>Empirical Software Engineering</i> , <b>2010</b> , 15, 577-598  | 3.3 | 16 |
| 20 | Classification Systems for Bacterial Protein-Protein Interaction Document Retrieval. <i>International Journal of Computational Models and Algorithms in Medicine</i> , <b>2010</b> , 1, 34-44          |     |    |

|    |  |     |    |
|----|--|-----|----|
| 19 | BioTagger-GM: a gene/protein name recognition system. <i>Journal of the American Medical Informatics Association: JAMIA</i> , <b>2009</b> , 16, 247-55   | 8.6 | 47 |
| 18 | Software Engineering Education for Bioinformatics <b>2009</b> ,  |     | 8  |
| 17 | Semi-supervised Learning of Text Classification on Bacterial Protein-Protein Interaction Documents <b>2009</b> ,   |     | 6  |
| 16 | SpliceCenter: a suite of web-based bioinformatic applications for evaluating the impact of alternative splicing on RT-PCR, RNAi, microarray, and peptide-based studies. <i>BMC Bioinformatics</i> , <b>2008</b> , 9, 313 | 3.6 | 36 |
| 15 | iProLINK: A Framework for Linking Text Mining with Ontology and Systems Biology <b>2008</b> ,  |     | 1  |
| 14 | Document Classification for Mining Host Pathogen Protein-Protein Interactions <b>2008</b> ,  |     | 1  |
| 13 | Teaching software engineering to end-users <b>2008</b> ,   |     | 4  |
| 12 | Annotating breast cancer microarray samples using ontologies <b>2008</b> , 414-8   | 0.7 |    |
| 11 | Mapping Gene/Protein Names in Free Text to Biomedical Databases <b>2007</b> ,  |     | 2  |
| 10 | AffyProbeMiner: a web resource for computing or retrieving accurately redefined Affymetrix probe sets. <i>Bioinformatics</i> , <b>2007</b> , 23, 2385-90   | 7.2 | 62 |
| 9  | BioThesaurus: a web-based thesaurus of protein and gene names. <i>Bioinformatics</i> , <b>2006</b> , 22, 103-5   | 7.2 | 88 |
| 8  | DynGO: a tool for visualizing and mining of Gene Ontology and its associations. <i>BMC Bioinformatics</i> , <b>2005</b> , 6, 201   | 3.6 | 28 |
| 7  | A multi-aspect comparison study of supervised word sense disambiguation. <i>Journal of the American Medical Informatics Association: JAMIA</i> , <b>2004</b> , 11, 320-31  | 8.6 | 57 |
| 6  | iProLINK: an integrated protein resource for literature mining. <i>Computational Biology and Chemistry</i> , <b>2004</b> , 28, 409-16  | 3.6 | 32 |
| 5  | CliniViewer: a tool for viewing electronic medical records based on natural language processing and XML. <i>Studies in Health Technology and Informatics</i> , <b>2004</b> , 107, 639-43                                 | 0.5 | 9  |
| 4  | A Query Taxonomy Describes Performance of Patient-Level Retrieval from Electronic Health Record Data   |     | 1  |
| 3  | Implementation of a Cohort Retrieval System for Clinical Data Repositories Using the Observational Medical Outcomes Partnership Common Data Model: Proof-of-Concept System Validation (Preprint)                         |     | 1  |
| 2  | The 2019 n2c2/OHNL Track on Clinical Semantic Textual Similarity: Overview (Preprint)  |     | 2  |

|   |  |   |
|---|--|---|
| 1 | Evaluation of Patient-Level Retrieval from Electronic Health Record Data for a Cohort Discovery Task | 1 |
|---|--|---|