Curriculum Vitae

You Chen, Ph.D.

Office Address:
Vanderbilt University Medical Center
Department of Biomedical Informatics
2525 West End Avenue, #1416
Nashville, TN 37203

Office Phone Number: (615) 343-1939
Office Fax Number: (615) 322-0502
Email: you.chen@vanderbilt.edu

Education

2004 – 2010 **Ph.D. in Computer Science**
Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China.

2010 – 2014 **Research Fellow in Biomedical Informatics**
Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN, USA.

Academic Appointments

2014-2016	Research Assistant Professor in Biomedical Informatics, Primary Appointment, Department of Biomedical Informatics, School of Medicine, Vanderbilt University Medical Center
2016 – present	Assistant Professor in Biomedical Informatics, Primary Appointment, Department of Biomedical Informatics, School of Medicine, Vanderbilt University Medical Center
2018 – present	Assistant Professor in Data Science, Affiliated Appointment, Data Science Institute, Vanderbilt University
2019 – present	Assistant Professor in Computer Science, Affiliated Appointment, Vanderbilt University

Professional Organizations

2003 – 2010	Association for Computing Machinery
2003 – 2010	Institute of Electrical and Electronics Engineers
2015 – 2017	The American Society of Human Genetics
2015 – 2017	Health:Further
2016 – 2018	Distributed Health
2016 – 2019	Healthcare Information and Management Systems Society
2019 – 2020	Institute of Electrical and Electronics Engineers
2010 – present	American Medical Informatics Association, Member

Professional Activities

Editorial Board Members

2015 – present	Informatics	(ISSN 2227-9709)
2019 – present	Smart Health	(ISSN: 2352-6483)
2020 – present	BioMedInformatics	(ISSN 2673-7426)

Guest Editing

2018 – 2019 editing a special issue – data-driven healthcare research in ***Informatics***.
2021 – 2022 editing a special issue - the promise of revolutionizing healthcare with artificial intelligence and Blockchain technology in ***Frontiers in Big Data***.

Thesis Defense Committee members

2018 Master student: Bryan Steitz in Biomedical Informatics
Thesis title: A Social Network Analysis of Cancer Provider Collaboration
2019 Ph.D. student: Cheng Ye in Computer Science
Thesis title: Information Retrieval in Clinical Chart Reviews

National Research Group Leader

2018– 2020 National Research Network - task workflow learning from EHR audit logs

Session Chair

2011 IEEE International Conference on Intelligence and Security Informatics
2018 AMIA Informatics Summit
2018 AMIA Annual Symposium
2019 AMIA Annual Symposium
2019 IEEE International Conference on Collaboration and Internet Computing
2021 AMIA Clinical Informatics

Conference Program Committees

2010 – 2015 International Symposium on Mining and Web
2010 – 2018 IEEE International Conference on Intelligence and Security Informatics
2010 – 2018 International Symposium on Foundation of Open Source Intelligence and Security Informatics
2010 – 2017 International Workshop on Trust, Security, and Privacy for Big Data
2015 – 2017 International Conference on Big Data
2017 – 2018 American Medical Informatics Association (AMIA) Scientific Program Committee
2018 – 2019 American Medical Informatics Association (AMIA) Scientific Program Committee
2019 – present First International Symposium on Mathematical and Computational Oncology

Manuscript Peer Review

2008 – present Computer and Security
2010 – present International Journal of Medical Informatics (IJMI)
2011 – present Journal of the American Medical Informatics Association (JAMIA)
2011 – present Journal of Biomedical Informatics (JBI)
2011 – present IEEE Transaction on Dependable and Secure Computing (TDSC)
2012 – present International Journal of Artificial Intelligence in Medicine (JAIM)
2012 – present Physics Letters A
2012 – present IEEE Journal on Biomedical and Health Informatics
2013 – present Computer Methods and Programs in Biomedicine

2013 – present Information Sciences
2013 – present BMC Medical Informatics & Decision Making
2013 – present Plos One
2015 – present American Journal of Managed Care
2015 – present Health Information Science and System
2015 – present Applied Clinical Informatics
2015 – present IEEE Access
2017 – present Nature Medicine
2019 – present Journal of Perinatology
2019 – present Journal of Medical Internet Research (JMIR)
2020 – present Experimental and Clinical Endocrinology & Diabetes
2020 – present Nature Communications Medicine

Grant Peer Review

2016 – 2019 Medical Research Council, Swindon, United Kingdom.
2017 – 2018 The Netherlands Organization for Scientific Research, The Netherlands.
2018 – 2019 The University of Applied Sciences and Arts of Southern Switzerland.
2018 – present Orthopaedic Research and Education Foundation (OREF). Rosemont, IL
2019 – present National Institute on Drug Abuse, NIH
2019 – present Edge Reviews, Vanderbilt University, Nashville, TN
2019 – present Summer Research Project Program, Vanderbilt University, Nashville, TN
2019 – present National CTSA Network. CTSA External Reviewers.
2020 – present Juvenile Diabetes Research Foundation (JDRF), New York, NY

Honors

2003 First Place of National Mathematical Contest in Modeling.
2008 Outstanding student of the Chinese Academy of Sciences.
2009 Bewinner Communications Award for outstanding Ph.D. candidate students.
2009 Outstanding student of the Chinese Academy of Sciences.
2011 Paper "Leveraging Social Networks to Detect Anomalous Insider Actions in Collaborative Environments" was honored as the best paper of the IEEE Intelligence and Security Informatics Conference.
2012 Paper "Detecting Anomalous Insiders in Collaborative Information Systems" was named one of the top 100 articles of 2012 in computer science by the journal ACM Computing Reviews.
2013 Recipient of a training grant from the National Library of Medicine (NLM/NIH).
2015 Paper "Building bridges Across Electronic Health Record Systems through Inferred Phenotypic Topics" published in Journal of Biomedical Informatics has been selected as the Editors' Choice paper for Volume 55
2018 Paper "DMMS: a decentralized blockchain ledger for the management of medication histories" received high school student scholarship from AMIA annual symposium.
2018 Paper "Interaction patterns of trauma providers are associated with length of stay. Journal of the American Medical Informatics Association" published in the Journal of American Medical Informatics Associations was selected as the Editors' Choice paper.
2018 Paper " Predicting Neonatal Encephalopathy From Maternal Data in Electronic Medical Records" won second place in the student paper competition in AMIA Informatics Summit.
2019 System "Deep Imputation of Temporal Data" won second place in the ICHI Data Analytics Challenge on Missing Data Imputation.

2019 System “XGBoost Imputation for Time Series Data” won third place in the ICHI Data Analytics Challenge on Missing Data Imputation.
2019 Undergraduate student's project to help health care organizations design management strategies to improve efficiencies of electronic health record systems received Vanderbilt University Data Science Institute summer research program award.
2019 Graduate student’s project on Investigating the workflow from electronic health records to verify if medical trainees exhibit clinical reasoning received Globalink Research Award.
2020 an Editorial award for demonstrating editorial excellence as an Associate Editor for Smart Health.
2021 Inducted into the Fellows of AMIA (FAMIA).

Media News

2016 Study tracks the makeup of VUMC collaborative care teams. VUMC Reporter. http://news.vumc.org/2016/10/27/study-tracks-makeup-of-vumc-collaborative-care-teams/
2018 Centered on the patient: Using medical records to improve care. Research Features. https://researchfeatures.com/2018/05/30/using-medical-records-improve-care/
2019 Study to explore care coordination’s impact on patient outcomes. VUMC Reporter. http://news.vumc.org/2019/09/25/study-to-explore-care-coordinations-impact-on-patient-outcomes/
2019 VUMC awarded $1.5M to analyze EHR data for care coordination patterns. Becker's Health IT and CIO Report. https://www.beckershospitalreview.com/ehrs/vumc-awarded-1-5m-to-analyze-ehr-data-for-care-coordination-patterns.html
2019 Data Science Institute welcomes the first cohort of undergraduate summer research fellows. https://engineering.vanderbilt.edu/news/2019/data-science-institute-welcomes-first-cohort-of-undergraduate-summer-research-fellows/
2021 COVID-19 met with intensive teamwork. VUMC Reporter. https://news.vumc.org/2021/04/01/covid-intensive-teamwork/
2021 Can machine learning provide advance notice of an extreme preterm birth? Communications of the ACM. https://cacm.acm.org/news/254598-predicting-preterm-birth/fulltext

Teaching Activities:

Graduate School Courses

2011 Guest Lecture on Anomaly Detection, course (BMIF-380 / CS-396), Vanderbilt University,
2013 Guest Lecture on Medical Information Management, course (BMIF-380/ CS-396)
2013 Guest Lecture on EMR System Usage Auditing, course (BMIF-380/ CS-396)
2015 Guest Lecture on Insider Threats Detection, course (BMIF-380/ CS-396)
2015 Guest Lecture on Collaborative Patterns Learning, course (BMIF-380/ CS-396)
2016 Guest Lecture on Learning of Healthcare Systems I, course (BMIF-380/ CS-396)
2016 Guest Lecture on Learning of Healthcare Systems II, course (BMIF-380/ CS-396)
2018 Guest Lecture on Insider Threat Detection, course (BMIF-380/ CS-396)
2018 Guest Lecture on Blockchain Technology, course (BMIF-380/ CS-396)
2018 Guest Lecture on Temporal Patterns Learning, course (BMIF-380/ CS-396)
2019 Guest Lecture on Temporal Patterns Learning, course (BMIF-380/ CS-396)
2019 Developed Data Science Case Study for Graduate Students major in Biomedical Informatics
2021 CS5891/CS3891 Network Analysis in Healthcare

Scientific Presentations

2011 Leveraging Social networks to Detect Anomalous insider actions in Collaborative Environments, IEEE International Conference on Intelligence and Security Informatics, July 11, 2011, Beijing, China

2011 Uncovering Anomalous Usage of Medical Records via Social Network Analysis, Team for Research in Ubiquitous Secure Technology, November, 2nd, 2011, Washington DC, U.S.

2012 Auditing Medical Records Accesses via Healthcare Interaction Networks, American Medical Informatics Association (AMIA) 2012 Annual Symposium, November 06, 2012, Chicago, IL

2012 Protecting Patients through Dynamic Network Analysis of Hospital Department Relationships, Team for Research in Ubiquitous Secure Technology, November 15, 2012, Washington DC, U.S.

2014 Creating Interpretable Collaborative Patterns to Detect Insider Threats, ICERM, Brown University. Invited talk. October 22-24, 2014, Providence, Rhode Island

2015 Inferring Clinical Workflow Efficiency via Electronic Medical Record Utilization. American Medical Informatics Association (AMIA) 2015 Annual Symposium, November 16, 2015, SF, CA

2017 Discovering Care Coordination Practice Patterns in the EHR: Interpretation and Impact on Patient Outcomes. International Symposium on Data-Driven Healthcare. Plenary talk. August 29-30, 2017, Singapore

2019 Provider Networks in the Neonatal Intensive Care Unit Affect Length of Stay. IEEE CIC 2019. December 12-14, 2019. Los Angeles, USA

2019 Learning tasks from EHR audit logs, AMIA 2019. Panel presentation. November 16-18, Washington DC, USA

2019 Learning EHR interaction workflows from audit logs. Invited talk. September 17, 2019. Ohio State University. Columbus, Ohio, USA.

2019 Learning to identify severe maternal morbidity from electronic health records, MedInfo 2019. August 23-30, 2019, Lyon, France

2019 Leveraging electronic health records to learn the progression path for severe maternal morbidity, MedInfo 2019. August 23-30, 2019, Lyon, France

2019 A Deep Learning Approach to Predict Neonatal Encephalopathy from Electronic Health Records, International Conference on Healthcare Informatics. June 10-13, 2019. Xi'an, China.

2019 Obstetric Patients with Repetitious Hospital Location Transfers Have Prolonged Stays, International Conference on Healthcare Informatics. June 10-13, 2019. Xi'an, China

2019 Leveraging EHR Audit Logs to Discover Effective Care Coordination Practice Patterns. National Research Network. Hosted by Dr. Julia Adler-Milstein. May 10, 2019.

2019 Leveraging Electronic Health Records to Discover Effective Care Coordination Practice Patterns, 60 minutes of oral presentation. HIMSS 2019, March 10, 2019, Orlando, FL.

2021 Leveraging Audit Log Data to Discover Effective Interactions of Healthcare Workers with Electronic Health Record Systems. Harvard University Biomedical Informatics Journal Club. Hosted by Dr. Isaac Kohane, February 16 2021

2021 The Collaboration Structures of Critical Care Teams and Patient Outcomes: a Retrospective Network Analysis. The University of Alabama at Birmingham, Powertalk, hosted by March 26, 2021. Hosted by Dr. James J. Cimino, February 16 2021

2021 Mining tasks and task characteristics from electronic health record audit logs with unsupervised machine learning. National Research Network Webinar on April 23, 2021. Dr. Julia Adler-Milstein from UCSF hosted the talk.

2021 Collaborative Care Team Structures and Patient Outcomes: Network Analysis. Wake Forest University. Hosted by Dr. Metin Gurcan. July 7, 2021.

2021 Collaboration Structures and Patient Outcomes: Network Analysis Study. Improving Health Outcomes Across the Lifespan Through HOMES Symposium. Ohio State University Department of Internal Medicine Center for Health Outcomes in Medicine Scholarship and Service (HOMES). Hosted by Dr. Robert Cronin, September 10, 2021

Students

2010 – 2015 Wen Zhang – Electric Engineering and Computer Science, Vanderbilt University, Graduate Student, mentoring committee member

2015 – 2017 Zejian Zhan – Electric Engineering and Computer Science, Vanderbilt University, Student Assistant, mentor

2016 summer Matthew Draper – School of Engineering, Stony Brook University, undergraduate student, mentor

2016-2017 Sixie Yu - Electric Engineering and Computer Science, Vanderbilt University, Student Assistant, mentor

2017 summer Annie Yin – Computer Science, Duke University, undergraduate student, mentor, Annie's work was presented in AMIA Informatics Summit 2018.

2017 summer Arthur Xin – Mathematics and Computer Science, Case Western Reserve University, undergraduate student, mentor

2017 summer Thomas Li – Computer Science, Duke University, undergraduate student, mentor. Thomas's work was presented in AMIA Informatics Summit 2018.

2018 summer Charlotte Zuber - Computer Science, Rutgers University, undergraduate student, mentor, Charlotte's work was presented as a podium in MedInfo 2019.

2018 summer Cindy Kim – Mathematics, Vanderbilt University, undergraduate student, mentor, Cindy's work was presented in AMIA Clinical Informatics, 2019.

2018 summer Patrick Li – Computer Science, University of Pennsylvania, undergraduate student, mentor, Patrick's work received AMIA scholarship in AMIA 2018.

2019 summer Xinmeng Zhang – Computer Science, Vanderbilt University, mentor. Xinmeng received an award from the data science institute summer research program, Vanderbilt University, in 2019. She was selected for Honorable Mention in Outstanding Undergraduate Researcher Awards for 2020 by Computing Research Association.

2019-2020 Barret Jones – Biomedical Informatics, Vanderbilt University Medical Center, Ph.D. student, research rotation mentor

2020-2021 Bob Chen - Chemical and Physical Biology Program, Vanderbilt University Medical Center, Ph.D. student, research rotation mentor

2020 summer Peter Ju – Computer Science, UIUC, mentor

2021 summer Hannah Mannering – Computer Science, Loyola University, undergraduate student, Mentor, Hannah's work was presented in AMIA Informatics Summit, 2021.

2021 summer Patrick Li – Computer Science, University of Pennsylvania, undergraduate student, mentor, Patrick's work was presented in AMIA Informatics Summit, 2021.

2016 – present Cheng Gao – Biomedical Informatics, Vanderbilt University, post-doctor, mentor

2019 – present Eugene Jeong – Biomedical Informatics, Vanderbilt University Medical Center, Ph.D. student, mentor

2020 – present Yubo Feng – Computer Science, Vanderbilt University, Ph.D. student, mentor

2020 – present Xinmeng Zhang – Computer Science, Vanderbilt University, Ph.D. Student, mentor

Ongoing Extramural and Intramural Research Support:

1R01LM012854-01A1 (PI: Chen) 08/1/2019 – 7/31/2023 5.40 calendar months
National Library of Medicine, National Institutes of Health **Total Award: $1,479,000**
Discovering Care Coordination Practice Patterns in the EMR: Interpretation and Impact on Patient Outcomes.
The overarching goal of this project is to learn care coordination patterns through the data stored in electronic medical record systems, assess their influence on care effectiveness (in the form of LOS and readmission), and evaluate the extent to which they are ready for adoption by healthcare organizations through comprehensive surveys and interviews with knowledgeable healthcare experts.
Role: PI

5 R01 AG058639-03 (PI: Ely) 04/15/2019 – 03/31/2024 0.60 calendar months
NIH/ NIA **Total Award: $17,051,430**
BRAIN-ICU-2 Study: Bringing to Light the Risk Factors and Incidence of Neuropsychological Dysfunction (Dementia) in ICU Survivors, 2nd Study
This BRAIN-ICU-2 Study [Bringing to light the Risk factors And Incidence of Neuropsychological dysfunction (dementia) in ICU Survivors, 2nd Study] will define the relationship between ICU delirium and dementia, will explain this unfortunate and life-changing brain damage in ICU survivors, and will pave the way for preventive programs, strategic rehabilitation, and targeted future interventions.
Role: Co-investigator

5R01 GM120484-03 (PI: Mayur) 01/01/2017 – 12/31/2021 0.60 calendar months
National Institute of General Medical Sciences **Total Award: $2,960,280**
The INSIGHT-ICU Study: Illuminating Neuropsychological dysfunction and Systemic Inflammatory mechanisms Gleaned after Hospitalization in Trauma-ICU Study
Goal: Following a multi-hundred-subject prospective Trauma ICU cohort for up to 12 months, to define the cognitive effects of injury and ICU survivorship that affects a large segment of the working population, while revealing pathophysiologic mechanisms underlying long-term cognitive impairment.
Role: Co-investigator

5 R01 HG006844-08 (PI: Malin) 09/21/2012 – 07/31/2020 2.40 calendar months

National Library of Medicine, National Institutes of Health **Total Award: $999,570**
A Risk Management Framework for Identifiability in Genomics Research
Goal: Developing game-theoretic models to reason about the extent to which genomic and clinical data should be designated as readily identifiable.
Role: Co-investigator

W81XWH-16-R-0033, DOD (PI: Mayur) 10/01/2016 – 09/22/2021 0.60 calendar months
Department of Defense (DOD) **Total Award: $582,754**
Linking Investigations in Trauma and Emergency Services (LITES) Network
The purpose of LITES is to create a standing research network of US trauma systems and centers with the capability to conduct prospective, multicenter, injury care and outcomes research of relevance to the Department of Defense (DOD).
Role: Co-investigator

Pending Extramural and Intramural Research Support:
1 R01 HS028628-01 (PIs: Chen, Gong, Patel) 12/01/2021 – 11/30/2026 2.40 calendar months

AHRQ **Total Award: $2,000,000**
Network Analysis: Discovering Novel Risk Factors from Electronic Health Records to Improve Patient Safety in ICUs
The timely and accurate detection and reporting of patient safety events (PSE) should improve the efficiency of decision-making, reduce clinician's workload, and ultimately enhance patient safety. The project leverages network analysis and machine learning to investigate PSE, PSE contexts, and dynamic teams in intensive care units from electronic health records. Products of this project will include a robust set of standardized PSE common data elements, a 5-axis radar graph characterizing dynamic teams, and PSE risk calculators based on machine learning.
Role: PI

2124320 (PIs: Malin, Chen, Jiang, Yan) 12/01/2021 – 11/30/2025 0.50 calendar months
NSF **Total Award: $1,200,000**
Collaborative Research: SCH: SePaLiT: Searching for Patients Like This in a Privacy-Preserving Federated Environment
The overarching goal of this project is to develop a distributed framework that can be embedded into the medical data query systems that have gained traction and wide deployment to conduct privacy-preserving searches for patients like this (SePaLiT) queries across participating organizations to enable and facilitate evidence-based medical research and, ultimately, care.
Role: PI

1 R01 HD104371-01A1 (PI: Velez Edwards) 07/01/2021 – 06/30/2026 1.80 calendar months
NIH **Total Award: $648,541**
Precision Medicine Approaches to Identifying Women at Risk for PCOS
Polycystic ovarian syndrome (PCOS) is a complex disease with a heterogenous clinical presentation that poses challenges in diagnosis. We leverage multiple large-scale electronic health record resources to develop and validate race-specific polygenic and clinical risk scores for PCOS to be used for identifying women at highest risk for PCOS. In order to better understand the relationship between health history and PCOS genetic risk we will also conduct phenome-wide association studies to identify novel conditions with shared genetic architecture with PCOS.

1 R01 HD107679-01 (PI: Velez Edwards) 11/01/2021 – 10/31/2025 1.20 calendar months
NIH/NICHHD **Total Award: $3,285,916**
Using the Clinical Phenome and Polygenic Risk Scores to Understand Uterine Fibroids
We leverage multiple large-scale electronic health record resources to develop and validate race-specific polygenic and clinical risk scores for fibroids and apply these scores to develop fibroid prediction models risk, size, and recurrence after treatment and conduct phenome-wide association studies to identify novel conditions with shared genetic architecture with fibroids.

1 R01 NLM (PIs: Li, Chen, Su) 04/01/2022 – 03/31/2027 1.80 calendar months
NIH/NLM **Total Award: $2,060,731**

Machine learning drives translational drug interaction and pharmacogenetics research
The proposed project will provide novel ML methodologies, and robust evidence to study a number of drug/gene/ADE associations. These PG findings will provide a valuable resource for the wider scientific community for potential prospective studies, and contribute significantly towards precision medicine and improving clinical care.
Role: PI

1 R01 NICHD (PIs: Chen, Gong) 04/01/2022 – 03/31/2027 2.40 calendar months
NIH/NICHD **Total Award: $$4,020,098.70**
Discovering and interpreting deviations in neonatal safe care practices
The project includes a direct observational study of clinician and clinical team behaviors associated with health IT system use in providing and coordinating complex care in the NICU, and a data-mining study using EHR logs to capture NREs through the innovative analysis of EHR interaction data. The product of the observational study will be an enhanced NRE classification framework that incorporates behaviors of clinicians interacting health IT systems. The product of the data mining study will be an informatics framework for electronically detecting NREs from EHR interaction data.
Role: PI

1 R01 HD100437-04 (PI: Chen) 12/01/2022 – 11/30/2026 3.60 calendar months
NIH/NICHHD **Total Award: $2,661,369.22**
Learning from Electronic Health Records to Prevent Severe Maternal Morbidity
While guidelines for managing risk of severe maternal morbidity (SMM) exist, it remains challenging to predict and prevent SMM. The project will develop an informatics framework to predict and prevent SMM, beyond current guidelines. This project will enable healthcare organizations and maternal quality improvement collaborative to enhance their processes for controlling, predicting and ultimately preventing, SMM.
Role: PI

1 U01CA253491-01 (PI: Li) 07/01/2022 – 06/30/2025 1.20 calendar months
NIH/NCI **Total Award: $1,390,619**
canRxAE: a natural language processing platform in defining cancer drug-induced adverse events
The overarching goal of this project is to apply natural language processing and machine learning algorithms to longitudinal electronic health records to extract, define and discover cancer drug-induced adverse events.
Role: Site-PI

Completed Extramural and Intramural Research Support:
NLM K99/R00LM011933 (PI: Chen) 05/01/2015 – 04/30/2019 9.0 calendar months
National Library of Medicine, National Institutes of Health **Total Award: $ 823,926**
Learning Patterns of Collaboration to Optimize the Management of Care Providers
The research goal of this R00 is to design social network analysis models to learn community, dependency and workflow patterns for care providers in the healthcare systems.
Role: PI

IIS-1418504 (PI: Denny, Malin) 09/01/2014 – 08/31/2019 (NCE) 6.60 calendar months

National Science Foundation **Total Award: $692,972**
SCH:INT: Collaborative Research: High-throughput Phenotyping on Electronic Health Records using Multi-Tensor Factorization
This project will propose a general computational framework based on multi-tensor factorization for transforming electronic health records data into meaningful phenotypes with expert guidance.
Role: Co-investigator

2 R01 LM010207-05A1 (PI: Malin) 09/01/2014 – 08/31/2019 1.20 calendar months
National Library of Medicine, National Institutes of Health **Total Award: $1,396,355**
Automated Detection of Anomalous Access to Electronic Health Records
Goal: Engineer data mining methods to extract workflows and organizational models associated with the utilization of patients' medical records to discover anomalies concerning expected use.
Role: Co-investigator

R01 LM010207-0110 (PI: Malin) 09/01/2009 – 08/31/2014 12.0 calendar months
National Library of Medicine, National Institutes of Health **Total Award: $953,226**
Automated Detection of Anomalous Access to Electronic Health Records
Goal: Engineer data mining methods to extract patterns of use of patients' medical records and to discover anomalous activities.
Role: Key personal

PUBLICATIONS

Google Scholar H-index**: 21;** i10-index: **32;** Total citations: **over 1,650**
https://scholar.google.com/citations?user=c-pOkPEAAAAJ&hl=en

1. **Chen Y**, Li Y, Cheng X, Guo L. Survey and taxonomy of feature selection algorithms in intrusion detection system. Information Security and Cryptology. LNCS4318. 2006; 153-167.

2. Li Y, Fang B, **Chen Y**, Guo L. A lightweight intrusion detection model based on feature selection and maximum entropy model. IEEE International Conference on Communication and Technology. 2006; 1-4.

3. **Chen Y**, Li W, Cheng X. Toward building lightweight intrusion detection system through modified RMHC and SVM. 15th IEEE International Conference on Networks. 2007; 83-88

4. Li Y, Fang B, Guo L, **Chen Y**. Network anomaly detection based on TCM-KNN algorithm. ACM Symposium on Information, Computer and Communications Security. 2007; 13-19

5. **Chen Y**, Dai L, Li Y, Cheng X. Building lightweight intrusion detection system based on principal component analysis and C4.5 algorithm. 9th IEEE International Conference on Advanced Communication Technology. 2007; 2109-2112

6. Li Y, Fang B, Guo L, **Chen Y**. TCM-KNN algorithm for supervised network intrusion detection. Pacific Asia Workshop on Intelligence and Security Informatics (PAISI). 2007; p141-151

7. Dai L, **Chen Y**, Yun X. Optimizing IP flow classification using feature selection. The Eighth International Conference on Parallel and Distributed Computing, Applications and Technologies. 2007; 39-45.

8. **Chen Y**, Cheng X, Huang Y. A Wavelet-based model to recognize high-quality topics on web forums. IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, WI. 2008; 343-351.

9. **Chen Y**, Dai L, Cheng X. GATS-C4.5: An algorithm for optimizing features in flow classification. IEEE Consumer Communications and Networking Conference. 2008; 466-470.

10. **Chen Y**, Cheng X, Yang S. Bursty topics extraction for web forums. 11th ACM International Conference on Web Information and Data Management (WIDM). 2009; 55-58.

11. Yang S, Cheng X, **Chen Y**, Fang G, Zhang J, Xu H. Detect events on noisy textual datasets. Proceedings of the 12th Asia-Pacific Web Conference. 2010; 372-374

12. Liu X, Bai S, **Chen Y**, Cheng X. Improving the routing performance of KAD through social network analysis. IEEE symposium on Computers and Communications. 2010; 721-727

13. **Chen Y**, Nyemba S, Zhang W, and Malin B. Leveraging social networks to detect anomalous insider actions in collaborative environments. Proceedings of the 9th IEEE Intelligence and Security Informatics (ISI). 2011: 119-124. PMID: 25621314, PMCID: 4303584. (**Best paper award**)

14. **Chen Y**, Malin B. Detection of anomalous insiders in collaborative environments via relational analysis of access logs. Proceedings of the 1st ACM Conference on Data and Application Security and Privacy. 2011: 63-74. PMID: 25485309, PMCID: 4257138.

15. Li X, Xue Y, **Chen Y**, and Malin B. Context-aware anomaly detection for electronic medical record systems. 2nd USENIX Workshop on Health Security and Privacy. 2011.

16. **Chen Y**, Nyemba S, Malin B. Detecting anomalous insiders in collaborative information systems. IEEE Transactions on Dependable and Secure Computing. 2012; 9(3): 332-344. PMID: 24489520, PMCID: 3905623. (**One of the top articles of 2012 in computer science by the journal ACM Computer Review**)

17. **Chen Y**, Nyemba S, Zhang W, Malin B. Specializing network analysis to detect anomalous insider actions. Security Informatics. 2012; 1(1): 5. PMID: 23399988, PMCID: 3566705.

18. **Chen Y**, Nyemba S, Malin B. Auditing medical record accesses via healthcare interaction networks. Proceedings of the American Medical Informatics Association Annual Symposium. 2012; 93-102. PMID: 23304277, PMCID: 3540438.

19. Zhang W, **Chen Y**, Gunter C, Liebovitz D and Malin B. Evolving role definition through permission invocation patterns. Proceedings of the ACM Symposium on Access Control Models and Technologies (SACMAT). 2012:37-48

20. **Chen Y**, Lorenzi N, Nyemba S, Schildcrout J, Malin B. We work with them? health workers interpretation of organizational relations mined from electronic health records. International Journal of Medical Informatics. 2014; 83(7): 495-506. PMID: 24845147, PMCID: 4159755

21. Zhang W, **Chen Y**, Cybulski T, Fabbri D, Gunter C, Lawlor P, Liebovitz D, Malin B. Decide now or later ?: Quantifying the tradeoff between prospective and retrospective access decisions. Proceedings of the ACM SIGSAC Conference on Computer and Communications Security (CCS). 2014; 1182-1192.

22. **Chen Y**, Ghosh J, Bejan C, Gunter C, Kho A, Liebovitz D, Sun J, Denny J, Malin B. Building bridges across electronic health record systems through inferred phenotypic topics. Journal of Biomedical informatics. 2015; 55:482-93. PMID: 25841328, PMCID: 4464930 [**Editors' Choice paper**]

23. **Chen Y**, Xie W, Gunter C, Liebovitz D, Mehrotra S, Zhang H, Malin B. Inferring clinical workflow efficiency via electronic medical record utilization. Proceedings of the American Medical Informatics Annual Fall Symposium. 2015: 416-425.

24. Wang Y, Chen R, Ghosh J, Denny J, Kho A, **Chen Y**, Malin B, Sun J. Rubik: knowledge guided tensor factorization and completion for health data analytics. Proceedings of the 21st ACM SIGKDD Conference on Knowledge Discovery and Data Mining. 2015: 1265-1274.

25. Yin ZJ, **Chen Y**, Fabbri D, Sun JM and Malin B. #PrayForDad: Learning the semantics behind why social media users disclose health information. The 10th International AAAI Conference on Web and Social Media. 2016: 456-465.

26. Yan C, **Chen Y**, Li B, Liebovitz D, Malin B. Learning clinical workflows to identify subgroups of heart failure patients. Proceedings of the American Medical Informatics Annual Fall Symposium. 2016:1248-1257.

27. Gao, C., Kho, A. N., Ivory, C., Osmundson, S., Malin, B. A., & **Chen, Y**. Predicting length of stay for obstetric patients via electronic medical records. Studies in health technology and informatics. 2017; (245):1019-1023.

28. Zheng T, Xie W, Xu l, Zhang Y, Yang G, **Chen Y**. A Machine learning-based framework to identify type 2 diabetes through electronic health records. International Journal of Medical Informatics. 2017; 97:120-127.

29. **Chen Y**, Lorenzi N, Sandberg W, Wolgast K, Malin B. Identifying collaborative care teams through electronic medical record utilization patterns. Journal of the American Medical Informatics Association. 2017;24(e1):e111-e120.

30. Li T, Gao C, Yan C, Osmundson S, Malin B, **Chen Y**. Predicting neonatal encephalopathy from maternal data in electronic medical records. AMIA 2018 Informatics Summit. 359-368. (**The final list in the student paper competition**)

31. Yin W, Gao C, Xu Y, Li B, Ruderfer D, **Chen Y**. Learning opportunities for drug repositioning via GWAS and PheWAS findings. AMIA 2018 Informatics Summit. 237-246.

32. **Chen Y**, Kho A, Liebovitz D, Ivory C, Osmundson S, Bian J, Malin B. Learning bundled care opportunities from electronic medical records. Journal of Biomedical Informatics. 2018; 77:1-10

33. **Chen Y**, Patel M, McNaughton C, Malin B. Interaction patterns of trauma providers are associated with length of stay. Journal of the American Medical Informatics Association. 2018;25(7):790-799 (**Editors' Choice paper**)

34. Yan C, Yin Z, Xiang S, **Chen Y**, Vorobeychik Y, Fabbri D, Kho A, Liebovitz D, Malin B. Mining the best observational window to model social phenomena. In 2018 IEEE 4th International Conference on Collaboration and Internet Computing. 2018;46-55.

35. Gao C, Osmundson S, Edwards DR, Jackson GP, Malin BA, **Chen Y**. Deep learning predicts extreme preterm birth from electronic health records. Journal of biomedical informatics. 2019;100:103334. doi:10.1016/j.jbi.2019.103334

36. Gao C, Yan C, Malin BA, **Chen Y.** Corpus size influences clinical concept embeddings. AMIA Annual Symposium. 2019. 1583.

37. Yan C, Gao C, Zhang X, **Chen Y**, Malin B. Deep imputation of temporal data. The 7th IEEE International Conference on Healthcare Informatics. 2019; Second place in data challenge on missing data Imputation. DOI: 10.1109/ICHI.2019.8904776

38. Zhang X, Yan C, Gao C, Malin B, **Chen Y**. XGBoost imputation for time series data. The 7th IEEE International Conference on Healthcare Informatics. 2019; Third place in data challenge on missing data Imputation. DOI: 10.1109/ICHI.2019.8904666

39. Gao C, Yan C, Osmundson S, Malin BA, **Chen Y**. A deep learning approach to predict neonatal encephalopathy from electronic health records. The 7th IEEE International Conference on Healthcare Informatics. 2019; DOI: 10.1109/ICHI.2019.8904667

40. Gao C, Yan C, Osmundson S, Malin BA, **Chen Y**. Obstetric patients with repetitious hospital location transfers have prolonged stays. The 7th IEEE International Conference on Healthcare Informatics. 2019; DOI: 10.1109/ICHI.2019.8904557

41. Gao C, Osmundson S, Yan X, Edwards DV, Malin BA, **Chen Y**. Leveraging electronic health records to learn progression path for severe maternal morbidity. Stud Health Technol Inform. 2019 Aug 21;264:148-152.

42. Gao C, Osmundson S, Yan X, Edwards DV, Malin BA, **Chen Y**. Learning to identify severe maternal morbidity from electronic health records. Stud Health Technol Inform. 2019 Aug 21;264:143-147.

43. Zuber C, **Chen Y**. A temporal pattern discovery algorithm for predicting neonatal mortality in NICU postoperative patients. MedInfo 2019. In press

44. Kim C, Lehmann C, Schildcrout J, Hatch D, France D, **Chen Y**. Provider networks in the neonatal intensive care unit associate with length of stay. In 2019 IEEE 5th International Conference on Collaboration and Internet Computing. 2019. DOI: 10.1109/CIC48465.2019.00024

45. Li P, Nelson SD, Malin BA, **Chen Y**. DMMS: a decentralized blockchain ledger for the management of medication histories. Blockchain in Healthcare Today. 2019 Jan 4:2(38). https://doi.org/10.30953/bhty.v2.38

46. **Chen Y**, Lehmann C, Hatch D, Schremp E, Malin B, France D. Modeling care team structures in the neonatal intensive care unit through network analysis of EHR audit logs. Methods of Information in Medicine. 2019;58(4-05):109-123. doi:10.1055/s-0040-1702237

47. **Chen Y**, Yan C, Patel MB. Network analysis subtleties in ICU structures and outcomes. Am J Respir Crit Care Med. 2020 Dec 1;202(11):1606-1607. doi: 10.1164/rccm.202008-3114LE. PMID: 32931298; PMCID: PMC7706157.

48. Ju P, Gao C, Patel M, **Chen Y**. Predicting trauma inpatient disposition through graph convolutional networks. AMIA Informatics Summit. 2020. In press.

49. Zhang X, Yan C, Gao C, Malin B, **Chen Y**. Predicting missing values in medical data via XGBoost regression. Journal of Healthcare Informatics Research. 2020 Aug 3:1-2.

50. Sinsky A, Rule A, Cohen G, Arndt B, Shanafelt T, Sharp C, Baxter S, Seale M, Yan S, **Chen Y**, Adler J, Hribar M. Metrics for assessing physician activity using EHR log data. Journal of the American Medical Informatics Association. 2020;27(4):639-643. doi:10.1093/jamia/ocz223

51. **Chen Y**, Feng Y, Yan C, Zhang X, Gao C. Modeling covid-19 growing trends to reveal the differences in the effectiveness of non-pharmaceutical interventions among countries in the world. medRxiv. 2020 Jan 1.

52. **Chen Y**. Learning healthcare worker networks from electronic health record utilization. In Teamwork in Healthcare 2020 Sep 25. IntechOpen.

53. Dai L, Zhang L, Li L, **Chen Y**. Detecting anomalous insiders using network analysis. In International Conference in Communications, Signal Processing, and Systems. 2020:17-29. Springer, Singapore.

54. Jones B, Zhang XM, Malin B, **Chen Y**. Learning tasks of pediatric providers from electronic health record audit logs. AMIA Annu Symp Proc. 2021 Jan 25;2020:612-618. eCollection.

55. Chen B, Alrifai W, Gao C, Jones B, Novak L, Lorenzi N, France D, Malin B, **Chen Y**. Mining tasks and task characteristics from electronic health record audit logs with unsupervised machine learning. Journal of the American Medical Informatics Association. 2021 Jun 12;28(6):1168-1177. (**Featured article**)

56. Yan C, Zhang X, Gao C, Wilfong E, Casey J, France D, Gong Y, Patel M, Malin B, **Chen Y**. Collaboration structures in COVID-19 critical care: retrospective network analysis study. JMIR human factors. 2021 Mar 8;8(1):e25724.

57. Zhu Y, Chiang CW, Wang L, Brock G, Milks MW, Cao W, Su Y, **Chen Y**, Zhang PY, Zeng DL, Donneyong M, and Li L. A multi-state transition model for statin-induced myopathy and statin discontinuation. CPT: pharmacometrics and systems pharmacology. 2021, in press.

58. Li P, Chen B, Rhodes E, Slagle J, Alrifai MW, France D, **Chen Y**. Measuring collaboration through concurrent electronic health record usage. JMIR Medical Informatics. 2021 Sep 3;9(9):e28998.

59. Jeong E, Person AK, Stollings JL, Su Y, Li L, **Chen Y**. Detecting drug-drug interactions in COVID-19 patients. MedInfo 2021. In press.

60. Yan C, Gao C, Zhang Z, Chen W, Malin B, Ely W, Patel M, **Chen Y**. Predicting brain function status changes in critically ill patients. Journal of the American Medical Informatics Association. Epub ahead of print. PMID: 34402496

61. Jeong E, Osmundson S, Gao C, Edwards DRV, Malin B, **Chen Y**. Learning the impact of acute and chronic diseases on forecasting neonatal encephalopathy. Computer Methods and Programs in Biomedicine. 2021 Sep 8:106397.

62. **Chen Y,** Alrifai MW, Gong Y, Evan R, Slagle J, Malin B, France D. Perioperative care structures and non-routine events: network analysis. MedInfo 2021. In press.

63. **Chen Y**, Gong Y. Teamwork and patient safety in intensive care units: challenges and opportunities. MedInfo 2021. In press.

64. Gao C, Osmundson S, Malin B, **Chen Y**. Telehealth use in prenatal care. MedInfo 2021. In press.

65. Gao C, Malin B, **Chen Y**. Telehealth uptake into primary care during the COVID-19 pandemic. MedInfo 2021. In press.

66. Huang H, Li H, Zhang XK, Shi GZ, Xu M, Ru XJ, G, **Chen Y**, Patel M, Wesley E, Lin S, Zhang GB. Association of postoperative delirium with cognitive outcomes: a meta-analysis. Journal of Clinical Anesthesia. 2021 Dec 1;75:110496.

67. Gao C, Osmundson S, Malin B, **Chen Y**. Prenatal telehealth during the pandemic: sociodemographic and clinical associations. Telehealth and Medicine Today. In press.

68. Zhang X, Yan C, Malin B, Patel M, **Chen Y**. Predicting next day discharges through access logs of electronic health records. Journal of the American Medical Informatics Association. In press.

69. Mannering H, Yan C, Gong Y, Alrifai MW, France D, **Chen Y**. Assessing neonatal intensive care unit structures and outcomes before and during the COVID-19 pandemic. JMIR. Second round review.

70. Wei Z, **Chen Y,** Raghu P. Upender adipokines link sleep disorders to metabolic dysregulation: insights from network analysis. Metabolism. Submitted.

71. Albaugh VL, Yu D, **Chen Y**, Williams B, Spann MD, Flynn R, English W. Sex- and procedure-dependent effects underlie long-term weight loss outcomes following metabolic and bariatric surgery. Annals of Surgery Open. Submitted.

72. Zhang G, **Chen Y**, Patel M, Wesly E. Higher grade glioma increases the risk of postoperative delirium: deficient neuroplasticity as a potential mechanism. Critical Care Medicine. Submitted.