CURRICULUM VITAE

STEPHANIE S. HONG, M.S.,S.B, F.A.M.I.A.
Director, Tranlational Informatics Applications,
Biomedical Informatics and Data Science Section,
Division of General Internal Medicine, Department of Medicine
Johns Hopkins University School of Medicine

Date March 17, 2025

Contact Information

Business Address: Johns Hopkins University School Medicine,

General Internal Medicne, Department of Medicine Biomedical Informatics and Data Science Section

2024 E Monument St, Suite1-200

Baltimore, MD 21205

Email: shong59@jh.edu Foreign Languages: Korean (fluent)

ORCID Number: https://orcid.org/0000-0002-0795-1293

Education

1988 B.S. Electrical Engineering, Massachusetts Institute of Technology
2025 (expected March 17) M.S. Applied Health Science Informatics (online), Johns Hopkins University
School of Medicine,

Employment History

1987 - 1989 1989 - 1990 1990 - 1992 1992 - 1994 1994 - 2004 2004 - 2006 2014 - 2017	Software Engineer, Wang Laboratories, Inc., Lowell, MA Software Engineer, Clinical Information Advantages, Inc., Waltham, MA Senior Consultant/Business Analyst, Logica Data Architects, Inc., Waltham, MA Senior Software Engineer, Sigma Imaging Systems, New York, NY Senior Software Engineer, Oracle/ Hyperion Solutions, Stamford, CT Technical Lead/Senior Software Engineer, Cognos/Applix, Inc. Waltham, MA Application Developer/ Healthcare Systems Data Analyst, Vanderbilt Institute for Clinical and Translational Research, Vanderbilt University Medical Center, Nashville,
2019 - 2020 2020 – 2022	TN Bioinformatics Volunteer, University of Maryland School of Medicine, Baltimore, MD Clinical Data Analyst, Biomedical Informatics and Data Science Section, Division of General Internal Medicine, Department of Medicine, Johns Hopkins University
2022 - 2023	School of Medicine, Baltimore, MD Software Engineer, Biomedical Informatics and Data Science Section, Division of General Internal Medicine, Department of Medicine, Johns Hopkins University
2023 - present	School of Medicine, Baltimore, MD Sr. Software Engineer / Director, Translational Informatics Applications, Biomedical Informatics and Data Science Section, Division of General Internal Medicine,

Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD

Professional Society Membership

2020 - present Member, American Medical Informatics Association

Honors and Awards

2022	Fellow, American Medical Informatics Association
2023	Grand Prize Winner, Federation of American Societies for Experimental Biology
	DataWorks Competition
2023	Best Contribution Award for Observation Data Standards and Management
	category, Observational Health Data Sciences and Informatics (OHDSI) Global
	Symposium 2023

Administrative Service

National/International Data Service for COVID research Data Repository

2020 - present

• Member, The National COVID Cohort Collaborative (N3C) Data Enclave.

It is a platform that provides researchers access to COVID-related patient EHR data in OMOP CDM format. It is the largest centralized repository of COVID-related Patient EHR data in U.S, for COVID-19 surveillance. The N3C is the largest collections of secure and deidentified clinical data in OMOP to date. It includes 36+ billion rows of data from 23+ million Patients >9+ million COVID patients EHR data ingested from 86+ DTAs around United States. As the lead programmer for the Data Ingestion and Harmonization workstream, designed the data pipeline and wrote the terminology transformation code to ingest 86+ site's EHR data from their native formats, -PCORnet, ACT, OMOP, TriNetX, PEDsNet CDM datasets – and harmonizes the data into a common data model (CDM) based on the Observational Medical Outcomes Partnership (OMOP) 5.3.1 version. For this effort, I was elected Fellow of the American Medical Informatics Association in 2022. The same data pipeline generation code is currently being used to rollout N3C Clinical Pilot program to extended the research capabilities for other clinical research initiative, e.g. Cancer, COPD, Alzheimer and Renal cohort.

Contributed to the research guide to N3C online book chapters and podium presentation at AMIA Informatics Summit and OHDSI Global Symposium.

CRISP

Designed and implemented data pipeline to ingest CRISP dataset, state designated Health Information Exchange of Maryland, into OMOP CDM instance for COVID research. It is implemented in Microsoft Azure Databricks environment. Deployed the dataset to support multiple IRB approved public research projects.

AI-READI

Led data standardization and harmonization to map AI-READI data elements and harmonize survey data, enabling sharing that meet the FAIR (Findable, Accessible, Interoperable, Reusable) Principles. We minted new custom concept ids to generated AI-READI vocabularies to overcome challenges of mapping non-standard datasets. Applied approaches in standardizing and harmonizing newly collected survey instrument data for meaningful use and how we integrated data the AI-READI survey into The Artificial Intelligence Ready and Equitable Atlas for Diabetes Insights (AI-READi) project is one of the four Data Generation Projects funded by Bridge2AI, an NIH Common Fund Program aimed at setting the stage for widespread adoption of AI in health research. multiple data element mapping and ETL implementation projects with collaborative teams in multiple geographic locations, provided technical training to rotating students as well as provided technical guidance to contractors, co-team members and junior members of the technical team.

CLAD

2022 - present	Member, Bridge2Al Standards Core Team
2023 - present	Member, BioData Catalyst (BDC) Data Standards Core team. Both team aims to develop best practices for data collection, deposition, quality assurance, query, dissemination, and integration across Data Generation Projects and BDC projects.
2023 - present	Member, Observational Health Data Sciences and Informatics (OHDSI) Common Data Model (CDM) Working Group
2023 - present	Member, Observational Health Data Sciences and Informatics (OHDSI) Data Quality Working Group
2023 - present	Member, Observational Health Data Sciences and Informatics (OHDSI) Vocabulary Working Group
2023 - present	Member, Observational Health Data Sciences and Informatics (OHDSI) Eye Care and Vision Imaging Working Group
2023 - present	Member, Observational Health Data Sciences and Informatics (OHDSI) Databricks Implementation Working Group
2023 - present	Member, Observational Health Data Sciences and Informatics (OHDSI) Open Source Technical Advisory Board,

Selected Code Repository

- N3C DI&H code repository: https://github.com/National-COVID-Cohort-Collaborative/Data-Ingestion-and-Harmonization
- CRISP dashboard: https://adb-

7520550328862896.16.azuredatabricks.net/?o=7520550328862896#notebook/283341489397216 1/dashboard/2833414893972222

Supporting research projects: rit10-crisp-jhsph-prod-dbw-2, rit116-crisp-page-irb00355096-prod-dbw-1, rit116-crisp-page-irb00329822-prod-dbw-1

Selected Published Software

<u>Hong, S.S.</u> National COVID Cohort Collaborative (N3C) Data Ingestion and Harmonization.

https://github.com/stephanieshong/Data-Ingestion-and-Harmonization https://github.com/National-COVID-Cohort-Collaborative

<u>Hong, S.S.</u> National Clinical Cohort Collaborative Data Ingestion and Harmonization. https://github.com/National-Clinical-Cohort-Collaborative/Data-Ingestion-and-Harmonization

Teaching Service

Provided guidance to Biomedical Informatics and Data Science Program (BIDS) graduate students who rotated in Dr. Chute's lab:

Anas Belouali, Emir Syailendra, Shijia Zhang, and Haeun Lee.

Graduate Education

Johns Hopkins University School of Medicine, Teaching Biomedical Informatics and Data Science Program Teaching Assistant. MB.250.778, Fast Healthcare Interoperability Resources (FHIR) is transforming healthcare with an open-web services' standards approach to clinical integration. This course is a hands-on experience working on integrating digital health and clinical systems interoperability.

- 2024 spring term Teaching Assistant, Implementing Fast Healthcare Interoperability Resources, ME.250.778.0000.SP24
- 2025 spring term Teaching Assistant Implementing Fast Healthcare Interoperability Resources, ME.250.778.0000.SP25

Grant Support

Active Grants

4/1/2022 – 3/31/2027 (Key Contributor) Multi-PI: Drs. Aaron Lee and Cecelia Lee

AI-READi Project

Researching ethical data integration and harmonization for Type 2

Diabetes (T2DM) insights.

NIH Common Fund, Grant Number?

NIH/NCATS

7/1/2023 – 6/30/2027 (Key Contributor) Multi-PI Grant with Dr. Christopher Chute

N3C Harmonization Pipeline Development.

Maintaining and enhancing pipelines for harmonizing multiple CDM models

into OMOP for COVID-19 research.
NIH Common Fund, Grant Number?

NIH/NCATS

7/1/24? - ?

(Key Contributor) NHLBI BioData Catalyst, Data Management Core (BDC)

Grant Number:?

Participating in the following list of active grants

- National Covid Cohort Collaborative expires in September 2024 Extension in review
- Post-acute sequelae SARS-CoV-2 infection (PASC) will end in 2024 with extension into 2025
- National Clinical Cohort Collaborative (Pilot program for Renal, COPD, and Alzheimer ended and transferred to Cancer Enclave)
- BRIDGE2AI- AI-READi (https://bridge2ai.org/standards-core/, expires in 4yrs)
- NHLBI BioData Catalyst, Data Management Core (expires in 3 year)
- All of US, Center of Linkage and Acquisition of Data (CLAD) (expires in 5 years)

Completed Grants

 National Covid Cohort Collaborative expires in September 2024 with no cost extension into 2025

Work-in-Progress, first author papers:

- CMS paper Augmenting the National COVID Cohort Collaborative (N3C) Dataset with Centers for Medicare and Medicaid Services (CMS) Data, Secure and De-identified Clinical Dataset https://www.ohdsi.org/wp-content/uploads/2023/10/4-HongBriefreport2.pdf
- Pipeline paper National COVID Cohort Collaborative (N3C) & FastHealthcare Interoperability Resources (FHIR®) pipeline implementation to OMOP CDM - https://github.com/National-Clinical-Cohort-Collaborative/Data-Ingestion-and-Harmonization
- Strength and weaknesses of in silo replication of epidemiological studies: Benzodiazepines treatment and patient outcome: Survival Analysis of Benzodiazepines in Patients with a History of Cancer.
- Streamlining Research Data Standardization: Al-READI Survey Instrument Data Elements and
 MoCA Measurement Data Elements are curated and mapped utilizing a Standardized Value Set
 Mapping Table for transformation into the OMOP Common Data Model,
 https://www.ohdsi.org/wp-content/uploads/2024/10/12-hong-stephanie-briefReport_StreamliningResearchDataStandardization_2024symposiumFinal-Stephanie-S-Hong.pdf

Peer-reviewed journal articles

- 1. Haendel MA,..., **Hong SS**,..., Austin CP, for the National COVID Cohort Collaborative (N3C) Consortium (#26 of named authors). The national COVID cohort collaborative (N3C): rationale, design, infrastructure and deployment. *J Am Med Inform Assoc* 2021; 28:427-43. PMID:32805036.
- Bennett TD,..., Hong SS,..., Chute CG, for the National COVID Cohort Collaborative (N3C)
 Consortium (#17 of named authors). Clinical Characterization and Prediction of Clinical Severity of
 SARS-CoV-2 Infection Among US Adults Using Data From the US National COVID Cohort
 Collaborative. JAMA Netw Open 2021; 4:e2116901. PMID:34255046.
- Pfaff ER, ...Hong, SS,., Zai A, for the National COVID Cohort Collaborative (N3C) Consortium (#12 of named authors). Synergy between centralized and federated approaches to data quality: a report form the national COVID cohort collaborative. J Am Med Inform Assoc 2022; 29:609-18.
 PMID:34590684
- 4. Bradwell KR,... Hong SS,... Moffitt RA, for the National COVID Cohort Collaborative (N3C) Consortium (#14 of named authors). Harmonizing units and values of quantitative data elements

- in a very large nationally pooled electronic health record (EHR) dataset. *J Am Med Inform Assoc* 2022; 29:1172-82. PMID:35435957
- Bramante CT, ... Hong SS,..., for the National COVID Cohort Collaborative (N3C) Consortium (#11 of named authors). Diabetes medications and associations with Covid-19 outcomes in the N3C database: a national retrospective cohort study. PLoS ONE 2022; 17:e0271574. PMID:36395143
 - M. Khodaverdi, Aayush Visaria, SS Hong, Priyal Makwana, Sijin Wen, Wesley Kimble, Scott Chapman, George G. Sokos, Brijesh D. Patel, (#3 named author) CARDIOVASCULAR PATIENT OUTCOMES AND COMPLICATIONS AFTER SARS-COV-2 INFECTION, Journal of the American College of Cardiology, Volume 79, Issue 9, Supplement, 2022, Page 2113, ISSN 0735-1097, https://doi.org/10.1016/S0735-1097(22)03104
 - (https://www.sciencedirect.com/science/article/pii/S0735109722031047)
- Liu S, Wen A, Wang L,..., Hong SS, et al. (#18 of named authors). An open natural language processing (NLP) framework for EHR-based clinical research: a case demonstration using the National COVID Cohort Collaborative (N3C). J Am Med Inform Assoc 2023; 30:2036-2040. PMID:37555837
- 8. Klein KR, Abrahamsen TJ, Kahkoska AR, Alexander GC, Chute CG, Haendel M, **Hong SS**, Mehta H, Moffitt R, Sturmer T, Kvist K, Buse JB, N3C Consortium. (#7 of named authors). Association of Premorbid GLP-1RA and SGLT-2i Prescription Alone and in Combination with COVID-19 Severity. *Diabetes Ther*. 2024; 15:1169-1186. PMID:38536629
- 9. Lee L, French E, Coelho DH, et al. Increased Incidence of Vestibular Disorders in Patients With SARS-CoV-2. *Otol Neurotol Open* 2024;4:e051. PMID:38919767
- Walter KM, Clark M, Dard S, Hong SS, et al. National COVID Cohort Collaborative data enhancements: a path for expanding common data models. *J Am Med Inform Assoc* 2025; 32:391-397. PMID:39579089

N3C Consortium Authorship

- Mehta HB, An H,...,Alexander GC, for the National COVID Cohort Collaborative (N3C) Consortium. Use of hydroxychloroquine, remdesivir, and dexamethasone among adults hospitalized with COVID-19 in the United States: a retrospective cohort study. *Ann Intern Med* 2021; 174:1395-1403. PMID:34399060
- Wong R, et al. Glycemic control and clinic outcomes in U.S. patients with COVID-19: data from the National COVID Cohort Collaborative (N3C) database. *Diabetes Care* 2022: 45:1099-1106.
 PMID:35202458
- 3. Chow JH, et al. Association of early aspirin use with in-hospital mortality in patients with moderate COVID-19. *JAMA Netw Open.* 2022; 5:e223890. PMID:35323950
- 4. Thomas JA, ... Wilcox AB, for the National COVID Cohort Collaborative (N3C) Consortium.

 Demonstrating an approach for evaluating synthetic geospatial and temporal epidemiologic data

- utility: results from analyzing >1.8 million SARS-CoV-2 tests in the United States National COVID Cohort Collaborative (N3C). *J Am Med Inform Assoc* 2022; 29:1130-65. PMID:35357487.
- 5. Pfaff ER, et al. Identifying who has long COVID in the USA: a machine learning approach using N3C data. *Lancet Digit Health* 2022; 4:e532-541. PMID:35589549
- Wong R, et al. Effect of SARS-CoV-2 Infection and Infection Severity on Longer-Term Glycemic Control and Weight in People With Type 2 Diabetes. *Diabetes Care* 2022; 45:2709-2717. PMID:36098660
- 7. Alur-Gupta S, Boland MR, Dokras A, N3C Consortium. Risk of severe acute respiratory syndrome coronavirus 2 infection among women with polycystic ovary syndrome. *Fertil Steril.* 2023; 119:847-857. PMID:36693555
- 8. Verhagen NB, Koerber NK, et al. Vaccination Against SARS-CoV-2 Decreases Risk of Adverse Events in Patients who Develop COVID-19 Following Cancer Surgery*Ann Surg Oncol* 2023; 30:1305-1308. PMID:36479662
- 9. Lyu T, et al. Risk for stillbirth among pregnant individuals with SARS-CoV-2 infection varied by gestational age *Am J Obstet Gynecol* 2023; 229:288.e1-e13. PMID:36858096
- 10. Vinson JA, et al. Hormone replacement therapy and COVID-19 outcomes in solid organ transplant recipients compared with the general population. *Am J Transplant*. 2023; 23:1035-1047. PMID:37105315
- 11. Widere JC, Davis CL, Loomba JJ, et al. Early Empiric Antibiotic Use in Patients Hospitalized With COVID-19: A Retrospective Cohort Study. *Crit Care Med* 2023; 51:1168-1176. PMID:37125800.
- 12. Lee E, Bates B, Kuhrt N, et al. National Trends in Anticoagulation Therapy for COVID-19 Hospitalized Adults in the United States: Analyses of the National COVID Cohort Collaborative. *J Infect Dis.* 2023; 228:895-906. PMID:37265224
- 13. Pincavitch JD, et al. Thirty-Day Mortality and Complication Rates in Total Joint Arthroplasty After a Recent COVID-19 Diagnosis: A Retrospective Cohort in the National COVID Cohort Collaborative (N3C). *J Bone Joint Surg Am.* 2023;105:1362-1372. PMID:37352338
- 14. Xiao X, et al. Potential drug-drug interactions among U.S. adults treated with nirmatrelvir/ritonavir: A cross-sectional study of the National Covid Cohort Collaborative (N3C). *Pharmacotherapy* 2023:43:1251-1261. PMID:37539477
- 15. Jones SE, et al. Who is pregnant? Defining real-world data-based pregnancy episodes in the National COVID Cohort Collaborative (N3C). *JAMIA Open* 2023; 6:00ad067. PMID: 37600074
- Patel B, Chapman SA, Neumann JT, et al. Outcomes of patients with active cancers and preexisting cardiovascular diseases infected with SARS-CoV-2. *Cardiooncology* 2023; 9:36. PMID:37803479
- 17. Grimseley EA, Torikashvili JV, et al. Nonelective coronary artery bypass graft outcomes are adversely impacted by Coronavirus disease 2019 infection, but not altered processes of care: A

- National COVID Cohort Collaborative and National Surgery Quality Improvement Program analysis. *JTCVS Open.* 2023; 16:342-352. PMID:38204718
- 18. Sharathkuma A, Wendt L, Ortman C, et al. COVID-19 outcomes in persons with hemophilia: results from a US-based national COVID-19 surveillance registry. *J Thromb Haemost* 2024; 22:61-75. PMID:37182697
- 19. Li CXR, Shaik T, et al. Characterizing the impact of SARS-CoV-2 reinfection on surgical outcomes. *Br J Surg.* 2024; 111:znad436. PMID:38215236
- 20. Kunz M, Rott KW, et al. The Intersections of COVID-19, HIV, and Race/Ethnicity: Machine Learning Methods to Identify and Model Risk Factors for Severe COVID-19 in a Large U.S. National Dataset. *AIDS Behav.* 2024; 28:5-21. PMID:38326668
- 21. SenthilKumar G, Verhage NB, et al. Risk of Early Postoperative Cardiovascular and Cerebrovascular Complication in Patients with Preoperative COVID-19 Undergoing Cancer Surgery. *J Am Coll Surg.* 238:1085-1097. PMID:38348959
- 22. Vinson Aj, Schissel M, et al. The prevalence of postacute sequelae of coronavirus disease 2019 in solid organ transplant recipients: Evaluation of risk in the National COVID Cohort Collaborative. *Am J Transplant* 2024; 24:1675-1689. PMID:38857785
- 23. Verhagen NB, Geissler T, et al. From Alpha to Omicron and Beyond: Associations Between SARS-CoV-2 Variants and Surgical Outcomes. *J Surg Res* 2024; 301:71-79. PMID:38917576
- 24. Mastoloni EM, et al. Impact of Nutritional Status on COVID-19-Induced Olfactory Dysfunction. *Laryngoscope* 2024; 134:4338-4343. PMID:39077963
- 25. Butzin-Dozier Z, Ji Y, Li H, et al. Predicting Long COVID in the National COVID Cohort Collaborative Using Super Learner: Cohort Study. *JMIR Public Health Surveill*. 2024; 10:e53322. PMID:39146534
- 26. Zhang X, et al. Chronic Lung Disease as a Risk Factor for Long COVID in Patients Diagnosed With Coronavirus Disease 2019: A Retrospective Cohort Study. *Open Forum Infect Dis.* 2024;11:ofae424. PMID:39183811
- 27. Islam JY, Hurwiz E, Li D, et al. Associations of County-Level Social Determinants of Health with COVID-19 Related Hospitalization Among People with HIV: A Retrospective Analysis of the U.S. National COVID Cohort Collaborative (N3C). *AIDS Behav.* 2024; 28:136-148. PMID:39292319
- 28. Hung P, Yu J, Harrison SE, et al. Racial and Ethnic and Rural Variations in the Use of Hybrid Prenatal Care in the US. *JAMA Netw Open* 2024; 7:e2449243. PMID:39641928
- 29. Berquist T, Loomba J, Pfatt E, et al. Crowd-sourced machine learning prediction of long COVID using data from the National COVID Cohort Collaborative. *EBioMedicine* 2024; 108:105333. PMID: 39321500
- 30. Patanwala AE, Xiao X, Hills TE, et al. Comparative Effectiveness of Baricitinib Versus Tocilizumab in Hospitalized Patients With COVID-19: A Retrospective Cohort Study of the National Covid Collaborative. *Crit Care Med.* 2025; 53:e29-e41. PMID:39365115
- 31. Wang WK, Jeong H, Hershkovich L, et al. Tree-based classification model for Long-COVID infection prediction with age stratification using data from the National COVID Cohort Collaborative. *JAMIA Open* 2024; 7:00ae111. PMID: 39524607
- 32. Vinson AJ, Anzalone AJ, Schissel M, et al. Association of COVID-19 With Risk of Posttransplant Diabetes Mellitus. *Transplantation* 2024; doi:10.1097/TP.000000000005227. PMID:39531312
- 33. Soff S, Yoo YJ, Bramante C, et al. Association of glycemic control with Long COVID in patients with type 2 diabetes: findings from the National COVID Cohort Collaborative (N3C). BMJ Open Diabetes Res Care. 2025; 13:e004536. PMID:39904520
- 34. Martin B, DeWitt PE, Russell S, et al. The Recent Increase in Invasive Bacterial Infections: A Report From the National COVID Cohort Collaborative. *Pediatr Infect Dis J* 2025; 44:217-227. PMID:39951364

35. Lee E, Bates B, Kuhrt N, Andersen KM, Visaria A, Patel R, Setoguchi S; N3C Consortium. National Trends in Anticoagulation Therapy for COVID-19 Hospitalized Adults in the United States: Analyses of the National COVID Cohort Collaborative. J Infect Dis. 2023 Oct 3;228(7):895-906. doi: 10.1093/infdis/jiad194. PMID: 37265224; PMCID: PMC10547450.

Major Invited Poster Presentation/Panel

National

- 1. Augmenting the National COVID Cohort Collaborative (N3C) Dataset with Medicare and Medicaid (CMS) Data, Secure and De-identified Clinical Dataset, N3C DIH workstream presentation, Stephanie Hong November 14, 2023, N3C Workstream presentation
- 3. Panel discussion titled "Challenges and Opportunities in Building a Data Sharing Ecosystem for Health Research & Surveillance," presented at the Potomac Officers Club's 2023 Healthcare Summit, Falls Church, Virginia, The panel was moderated by John Beresny, Director of the Health Solutions Division at Noblis, and included other esteemed panelists such as Dr. Christopher Chute, Bloomberg Distinguished Professor of Health Informatics at Johns Hopkins University and Medicine, and Dr. Kimberly Marschhauser, Associate Director at the Patient-Centered Outcomes Research Institute (PCORI). Panel discussed the complexities and prospects associated with establishing a comprehensive data-sharing ecosystem to enhance health research and surveillance efforts. https://potomacofficersclub.com/events/poc-2023-healthcare-summit/

OHDSI Global Symposium:

- Augmenting the National COVID Cohort Collaborative (N3C) Dataset with Medicare and Medicaid (CMS) Data, Secure and Deidentified Clinical Dataset, Stephanie Hong, et. al, N3c Consortium, https://www.ohdsi.org/2023showcase-4/
- Streamlining Research Data Standardization: AI-READI Survey Instrument Data Elements and MoCA Measurement Data Elements are curated and mapped utilizing a Standardized Value Set Mapping Table for transformation into the OMOP Common Data Model, Stephanie S. Hong, et al, AI-READI consortium, https://www.ohdsi.org/2024showcase-12/