

## Curriculum Vitae: updated July 2022

### I. IDENTIFYING DATA

Name Alison Callahan, PhD  
Current Position Instructor, Stanford Center for Biomedical Informatics Research  
Current Affiliation Stanford University

### II. EDUCATION HISTORY

#### Colleges and Universities Attended

2007 Bachelor of Science with Honors (Integrated Science), Carleton University, Ottawa Canada  
2009 Master of Information Studies (ALA accredited), Faculty of Information, St. George Campus, University of Toronto, Toronto Canada  
2014 Doctor of Philosophy, Biology, Department of Biology, Carleton University (laboratory of Dr. Michel Dumontier), Ottawa Canada

#### Postdoctoral Training

2014 – 2016 Postdoctoral Scholar, Stanford Center for Biomedical Informatics Research, Stanford University School of Medicine, Stanford California (laboratory of Dr. Nigam H. Shah)

### III. EMPLOYMENT

#### Academic Appointments:

4/1/2016 – 05/15/2022 Research Scientist, Stanford Center for Biomedical Informatics Research  
Stanford University School of Medicine, Stanford, CA (reappointed in 2021, 5 year term)  
05/15/2022 – present Instructor, Stanford Center for Biomedical Informatics Research, Stanford University School of Medicine, Stanford, CA

#### Other Appointments:

2009-2012 Teaching assistant, Department of Biology, Carleton University, Ottawa Canada  
2008-2009 Graduate student library assistant, Gerstein Science Information Center, University of Toronto, Toronto Canada  
Summer 2008, 2009 Researcher, Canada Institute for Scientific and Technical Information, National Research Council, Ottawa Canada  
Summer 2005, 2006, 2007 Researcher, Department of Biology, Carleton University, Ottawa Canada

### IV. HONORS AND AWARDS

2005-2007	Natural Sciences and Engineering Research Council Undergraduate Student Research Award
2007	Carleton University Senate Medal for outstanding academic achievement
2008	Social Sciences and Humanities Research Council Canada Graduate Scholarship (Masters)
2011	Natural Sciences and Engineering Research Council Canada Graduate Scholarship (Doctoral)
2014 – present	Member, American Medical Informatics Association

## V. BIBLIOGRAPHY

### Peer-reviewed original research (33 total)

1. Fraser S, **Callahan A**, Klassen D, Sherratt TN. Empirical tests of the role of disruptive coloration in reducing detectability. *Proc Biol Sci* 274(1615):1325-1331, 2007.
2. Webster RJ, **Callahan A**, Godin J.-G., Sherratt TN. Behaviourally-mediated crypsis in two nocturnal moths with contrasting appearance. *Philos Trans R Soc Lond B Biol Sci* 364: 503-510, 2008.
3. **Callahan A**, Hockema S, Eysenbach G. Contextual cocitation: Augmenting cocitation analysis and its applications. *J Assoc Inf Sci Technol* 61(6): 1130-1143, 2010.
4. **Callahan A**, Dumontier M, Shah NH. HyQue: Evaluating hypotheses using Semantic Web technologies. *J Biomed Semantics* 2(Suppl 2): S3, 2011.
5. **Callahan A**, Cruz-Toledo J, Dumontier M. Ontology-Based Querying with Bio2RDF's Linked Open Data. *J Biomed Semantics* 4(Suppl 1): S1, 2013.
6. Dumontier M, Baker CJO, Baran J, **Callahan A**, Chepelev L, Cruz-Toledo J, Del Rio NR, Duck G, Furlong LI, Keath N, Klassen D, McCusker JP, Queralt-Rosinach N, Samwald M, Villanueva-Rosales N, Wilkinson MD, Hoehndorf R. The SemanticScience Integrated Ontology (SIO) for biomedical research and knowledge discovery. *J Biomed Semantics* 5:14, 2014.
7. Harpaz R, **Callahan A**, Tamang S, Low Y, Odgers D, Finlayson S, Jung K, LePendu P, Shah NH. Text mining for adverse drug events: The promise, challenges and state of the art. *Drug Saf* 37(10): 777-790, 2014.
8. Rodríguez González A, **Callahan A**, Cruz-Toledo J, Garcia A, Egaña Aranguren M, Dumontier M, Wilkinson MD. Automatically exposing OpenLifeData via SADI semantic web services. *J Biomed Semantics* 5:46, 2014.
9. **Callahan A**, Cifuentes JJ, Dumontier M. An evidence-based approach to identify aging-related genes in *Caenorhabditis elegans*. *BMC Bioinf* 16:40, 2015.
10. **Callahan A**, Pernek I, Stiglic G, Leskovec J, Strasberg HR, Shah NH. Analyzing Information Seeking and Drug-Safety Alert Response by Health Care Professionals as New Methods for Surveillance. *J Med Internet Res* 17(8): e204, 2015.
11. Banda JM, **Callahan A\***, Winnenburg R, Strasberg HR, Cami A, Reis BY, Vilar S, Hripcsak G, Dumontier M, Shah NH. Feasibility of prioritizing drug-drug-event associations found in electronic health records. *Drug Saf* 39(1): 45-57, 2016. \*Contributed to the data interpretation, and drafting of all manuscript sections.
12. **Callahan A**, Abeyruwan SW, Al-Ali H, Sakurai K, Ferguson AR, Popovich PG, Shah NH, Visser U, Bixby JL, Lemmon VP. RegenBase: a knowledge base of spinal cord injury biology for translational research. *Database (Oxford)* 16: baw040, 2016.

13. Dumontier M, Gray AJG, Marshall MS, Alexiev M, Ansell P, Bader G, Baran J, Bolleman JT, **Callahan A**, Cruz-Toledo J, Gaudet P, Gombocz E, Gonzalez-Beltran AN, Groth P, Haendel M, Ito M, Jupp S, Juty M, Katayama T, Kobayashi N, Krishnaswami K, Laibe C, Le Novère N, Lin S, Malone J, Miller M, Mungall CJ, Rietveld L, Wimalaratne SM, Yamaguchi A. The health care and life sciences community profile for dataset descriptions. *PeerJ* 4:e2331, 2016. *\*Contributed to review and editing of manuscript drafts.*
14. **Callahan A**, Anderson KD, Beattie MS, Bixby JL, Ferguson AR, Fouad K, Jakeman LB, Nielson JL, Popovich PG, Schwab JM, Lemmon VP. Developing a data sharing community for spinal cord injury research. *Exp Neurol* 295: 135-143, 2017.
15. **Callahan A**, Winnenburt R, Shah NH. U-Index, a dataset and an impact metric for informatics tools and databases. *Sci Data* 5: 180043, 2018.
16. Schuler A, **Callahan A\***, Jung K, Shah NH. Performing an informatics consult: methods and challenges. *J Am Coll Radiol* 15 (3): 563-568, 2018. *\*Contributed extensively to conception and design, drafting and critical revision of the manuscript.*
17. Wendling T, Jung K, **Callahan A\***, Schuler A, Shah NH, Gallego B. Comparing methods for estimation of heterogeneous treatment effects using observational data from health care databases. *Stat Med* 37 (23): 3309-3324, 2018. *\*Contributed to data acquisition and drafting of the Methods section of the manuscript.*
18. Nikfarjam A, Ransohoff JD, **Callahan A\***, Jones E, Loew B, Kwong BY, Sarin KY, Shah NH. Early detection of adverse drug reactions in social health networks: A natural language processing pipeline for signal detection. *JMIR Public Health Surveill* 5(2):e11264, 2019. *\*Contributed to data acquisition and data interpretation, and critical revision of the manuscript.*
19. Gombar S\*, **Callahan A\***, Califf R, Harrington R, Shah NH. It is time to learn from patients like mine. *npj Digit Med* 2, 16, 2019. *\*Co-first authors.*
20. Nikfarjam A, Ransohoff JD, **Callahan A\***, Polony V, Shah NH. Profiling off-label prescriptions in cancer treatment using social health networks. *JAMIA Open* 2(3): 301-305, 2019. *\*Contributed to data acquisition and data interpretation, and critical revision of the manuscript.*
21. **Callahan A\***, Fries JA\*, Ré C, Huddleston III JI, Giori NJ, Delp SL, Shah NH. Medical device surveillance with electronic health records. *npj Digit Med* 2:94, 2019. *\*Co-first authors.*
22. Fouad K, Bixby JL, **Callahan A\***, Grethe JS, Jakeman LB, Lemmon VP, Magnuson DSK, Martone ME, Nielson JL, Schwab JM, Taylor-Burds C, Tetzlaff W, Torres-Espin A, Ferguson AR and the FAIR-SCI Ahead Workshop Participants. FAIR SCI Ahead: The Evolution of the Open Data Commons for Pre-Clinical Spinal Cord Injury Research. *J Neurotrauma* 37(6): 831-838, 2020. *\*Contributed to data acquisition and drafting of the manuscript.*
23. **Callahan A**, Shah NH, Chen JH. Research and reporting considerations for observational studies using electronic health record data. *Ann Intern Med* 172 (11\_Supplement): S79-S84, 2020.
24. Kashyap S, Gombar S, Yadlowsky S, **Callahan A\***, Fries J, Pinsky BA, Shah NH. Measure what matters: counts of hospitalized patients are a better metric for health system capacity planning for a reopening. *J Am Med Assoc* 327(7):1126-1131, 2020. *\*Contributed to data acquisition, analysis and interpretation, and to critical revision of the manuscript for important intellectual content.*
25. **Callahan A\***, Steinberg E\*, Fries JA, Gombar S, Patel B, Corbin CK, Shah NH. Estimating the efficacy of symptom-based screening for COVID-19. *npj Digit Med* 3(1):1-3, 2020. *\*Co-first authors.*

26. Reps JM, Williams RD, You SC, Falconer T, Minty E, **Callahan A\***, Ryan PB, Park RW, Lim H.-S., Rijnbeek, P. Feasibility and evaluation of a large-scale external validation approach for patient-level prediction in an international data network: validation of models predicting stroke in female patients newly diagnosed with atrial fibrillation. *BMC Med Res Method* 20:102, 2020. \*Contributed to data acquisition.
27. Giori NJ, Radin R, **Callahan A\***, Fries JA, Halijaj E, Ré C, Delp SL, Shah NH, Harris AHS. Assessment of extractability and accuracy of electronic health record data for joint implant registries. *JAMA Netw Open* 4(3): e211728, 2021. \*Contributed to conception and design, and drafting of the manuscript.
28. Ibrahim B, de Freitas Mendonca MI, Gombar S, **Callahan A\***, Jung K, Capasso R. Association of systemic diseases with surgical treatment for obstructive sleep apnea compared with continuous positive airway pressure. *JAMA Otolaryngol Head Neck Surg* 147(4):329-335, 2021. \*Contributed to data acquisition, analysis and interpretation, and to critical revision of the manuscript for important intellectual content.
29. Fries JA, Steinberg E, Khattar S, Fleming SL, Posada JD, **Callahan A\***, Shah NH. Trove: Ontology-driven weak supervision for clinical entity classification. *Nat Commun* 12(1): 1-11, 2021. \*Contributed to conception and design, data acquisition and analysis, and to critical revision of the manuscript for important intellectual content.
30. **Callahan A\***, Polony V\*, Posada JD, Banda JM, Gombar S, Shah NH. ACE: The Advanced Cohort Engine for searching longitudinal patient records. *J Am Med Inform Assoc* 28 (7):1468-1479, 2021. \*Co-first authors.
31. Caswell-Jin JL\*, **Callahan A\***, Purington N, Han SS, Itakura H, John EM, Blayney DW, Sledge Jr GW, Shah NH, Kurian AW. Treatment and monitoring variability in US metastatic breast cancer care. *JCO Clin Cancer Inform* 5:600-614, 2021. \*Co-first authors.
32. **Callahan A\***, Gombar S\*, Cahan EM, Jung K, Steinberg E, Polony V, Morse K, Tibshirani R, Hastie T, Harrington R, Shah NH. Using aggregate patient data at the bedside via an on-demand consultation service. *NEJM Catalyst Innovations in Care Delivery* 2(10):10.1056/CAT.21.0224, 2021. \*Co-first authors.
33. Dash D, Gokhale A, Patel BS, **Callahan A\***, Posada J, Krishnan G, Collins W, Li R, Schulman K, Ren L, Shah NH. Building a learning health system: Creating an analytical workflow for evidence generation to inform institutional clinical care guidelines. *Applied Clinical Informatics* 13(1):315-321, 2021. \*Contributed to conception and design, data acquisition and analysis, and to critical revision of the manuscript for important intellectual content.

#### Peer-reviewed publications (other - 6 total)

1. **Callahan A**, Cruz-Toledo J, Ansell P, Klassen D, Tumarello G, Dumontier M. Improved Dataset Coverage and Interoperability with Bio2RDF Release 2. *Proceedings of the 5th International Conference on Semantic Web Applications and Tools for Life Sciences (SWAT4LS)*, Paris, France, November 28-30, 2012. (conference proceedings)
2. Dumontier M, **Callahan A**, Cruz-Toledo J, Ansell P, Emonet V, Belleau F, Droit A. Bio2RDF Release 3: A Larger Connected Network of Linked Data for the Life Sciences. *Proceedings of the International Semantic Web Conference 2014 Posters & Demonstrations Track* 1272: 401-404, 2014. (conference proceedings)

3. Odgers DJ, Harpaz R, **Callahan A**, Stiglic G, Shah NH. Analyzing search behavior of healthcare professionals for drug safety surveillance. *Proceedings of the Pacific Symposium on Biocomputing* 2015: 306-317, 2015. (conference proceedings)
4. Schuler A, Liu V, Wan J, **Callahan A\***, Udell M, Stark DE, Shah NH. Discovering patient phenotypes using generalized low rank models. *Proceedings of the Pacific Symposium on Biocomputing* 2016: 144-155, 2016. (conference proceedings) \*Contributed to data acquisition, and drafting of Methods and Discussion sections of manuscript.
5. **Callahan A**, Danzi MC, Zunino G, Cooper DJ, Shah NH, Visser U, Bixby JL, Lemmon VP. Elucidating effects of nerve injury on gene expression using RegenBase, a knowledge base of spinal cord injury biology. *Proceedings of the Bio-Ontologies SIG at Intelligent Systems for Molecular Biology, Orlando, Florida, July 8-9, 2016*. (conference proceedings)
6. **Callahan A**, Shah NH. 2018. A second opinion from observational data on second-line diabetes drugs. *JAMA Netw Open* (8), e186119, 2018. (invited commentary)

#### **Non-Peer-reviewed Articles (1 total)**

1. Lu JH, **Callahan A\***, Patel B, Morse K, Dash D, Shah NH. Low adherence to existing model reporting guidelines by commonly used clinical prediction models. medRxiv 10.1101/2021.07.21.21260282, 2021. \*Contributed to data collection and interpretation, and to drafting and critical revision of the manuscript.

#### **Book Chapters (4 total)**

1. Webster RJ, **Callahan A**, Godin J.-G., Sherratt TN. Camouflage behavior and body orientation on backgrounds containing directional patterns. In Stevens M, Merilaita S (Eds.), Animal Camouflage: Mechanisms and Function Cambridge UK: Cambridge University Press, 2011. pp. 101-117.
2. **Callahan A**, Dumontier M. Evaluating scientific hypotheses using the SPARQL Inferencing Notation. In Simperl E, Cimiano P, Polleres A, Corcho O, Presutti V (Eds.), The Semantic Web: Research and Applications, Lecture Notes in Computer Science, 7295, 2012. pp. 647-658.
3. **Callahan A**, Cruz-Toledo J, Ansell P and Dumontier M. Bio2RDF Release 2: Improved Coverage, Interoperability and Provenance of Life Science Linked Data. In Cimiano P, Corcho O, Presutti V, Hollink L, Rudolph S (Eds.), The Semantic Web: Semantics and Big Data, Lecture Notes in Computer Science 7882, 2013. pp 200-212.
4. **Callahan A**, Shah NH. Machine learning in healthcare. In Sheikh A, Cresswell KM, Wright A, Bates DW (Eds.), Key Advances in Clinical Informatics: Transforming Health Care through Health Information Technology. Elsevier, 2017. pp. 279-292.

#### **Abstracts (9 total)**

1. **Callahan A**, Sakurai K, Abeyruwan SW, Ferguson AR, Popovich PG, Visser U, Bixby JL, Lemmon VP. Minimum information about a spinal cord injury experiment (MIASCI): Concepts and integration with the RegenBase ontology. Abstracts from the 33rd Annual National Neurotrauma Symposium. *J Neurotrauma* 31(12): A-1-A-126, 2014.
2. Lemmon, VP, **Callahan A**, Abeyruwan SW, Ferguson AR, Popovich PG, Visser U, Bixby JL. Facilitating reproducibility and data integration for SCI research with MIASCI and RegenBase. Abstracts from the 33rd Annual National Neurotrauma Symposium. *J Neurotrauma* 32(12): A-1-A-152, 2015.

3. Gombar S, Moneghetti KJ, **Callahan A\***, Jung K, Wheeler MT. Beta-Blockers Inferior to Calcium Channel Blockers in Hypertrophic Cardiomyopathy. *Circulation* 136 (Suppl 1): A21307-A21307, 2017. \*Contributed to data acquisition, analysis, and interpretation, and to drafting of Methods section of manuscript.
4. Rhee J, Gombar S, Beutner K, **Callahan A\***, Jung K, Wheeler M. Cardiovascular safety of mexiletine as a therapy for myotonia in patients with myotonic dystrophy. *J Am Coll Cardiol* 73 (9 Supplement 1), 474, 2019. \*Contributed to data acquisition, analysis and interpretation, and to drafting of Methods section of manuscript.
5. Karimi Y, Gombar S, Dean L, Hansen J, **Callahan A\***, Jung K, Divi V, Blayney DW. Real-world efficacy of bone modifying agents (BMAs) in patients with breast cancer (BC) treated in an academic health system: Use of the “green button”. *J Clin Oncol* 37 (Supplement 15): e18054, 2019. \*Contributed to data acquisition, analysis and interpretation, and to drafting of Methods section of manuscript.
6. Meng L, Gombar S, **Callahan A\***, Jung K, Holubar M, Chang A, Mui E, Deresinski S. Step-down from IV to oral therapy in patients with bacteremia due to Enterobacteriaceae: fluoroquinolones (FQ) vs.  $\beta$ -lactams (BL) or trimethoprim-sulfamethoxazole (TMP-SMX). *Open Forum Infect Dis* 6 (Supplement 2): S124, 2019. \*Contributed to data acquisition, analysis and interpretation, and to drafting of Methods section of manuscript.
7. Caswell-Jin JL\*, **Callahan A\***, Purington N, Han SS, Itakura H, Sledge GW, Shah NH, Kurian AW. Linking insurance claims across time to characterize treatment, monitoring, and end-of-life care in metastatic breast cancer. *J Clin Oncol* 38 (15\_Suppl): 7063, 2020. \*Co-first authors.
8. **Callahan A**, Murugappan G, Main EK, Leonard SA. Constructing a cohort of nulliparous, term, singleton, vertex births from electronic health records. *Society for Maternal and Fetal Medicine 42<sup>nd</sup> Annual Pregnancy Meeting*, January 31-February 5 2022, Orlando FL.
9. **Callahan A**, Leonard S, Druzin M, Lathi RB, Murugappan G. Constructing a pregnancy loss cohort from electronic health records. *American College of Obstetricians & Gynecologists Annual Meeting*, May 6-8 2022, San Diego CA.

## VI. GRANT FUNDING

### Current:

9/2017 – 8/2022

Funder: NLM (R01)

Title: From enrichments to insights

Role: Key personnel (Principal Investigator: Nigam H. Shah)

### Submitted:

2022 – 2027

Funder: NLM (R01)

Title: Bridging the AI chasm with workflow-aware models

Role: Co-Investigator (PI: Nigam H. Shah)

## VII. PATENTS

1. U.S. Patent 11210296, “Systems and methods for cohort analysis using compressed data objects enabling fast memory lookups,” Dec 28, 2021. Inventors: Nigam H. Shah (Menlo Park, CA), Vladimir Polony (Stanford, CA), Juan M. Banda (Stanford, CA), Alison V. Callahan (Stanford, CA).

## VIII. EDITORIAL SERVICE

## **Ad Hoc Reviewer**

2013 – 2018	<i>Bio-Ontologies Special Interest Group Meeting at ISMB</i>
2015 – present	<i>American Medical Informatics Association Annual Symposium</i>
2018 – present	<i>Machine Learning for Health (ML4H) at NeurIPS</i>
2019 – present	<i>Nature Communications</i>
2021 – present	<i>Journal of the American Medical Informatics Association</i>
2021 – present	<i>Nature Scientific Data</i>

## **IX. SERVICE AS GRANT REVIEWER**

2017 – 2022	Biomedical Informatics Library and Data Science (BILDS) study section, National Library of Medicine, National Institutes of Health – standing member
-------------	---

## **X. INVITED PRESENTATIONS**

1. 2015 “Text mining in clinical data analysis.” Data Boot Camp, Veterans Affairs, Menlo Park, CA
2. 2016 “RegenBase: A knowledge base of spinal cord injury biology for translational research.” NINDS FAIR Share Workshop, NIH, Bethesda, MD
3. 2016 “Extracting patient-reported pain from clinical notes.” Stanford Data Science Initiative Retreat, Stanford, CA
4. 2017 “Governance for data sharing.” FAIR SCI Ahead Satellite Meeting at the Society for Neuroscience Annual Meeting, Washington, DC
5. 2017 “The Informatics Consult: Enabling precision in the practice of medicine.” Artificial Intelligence in Medicine (AIMed), Laguna Niguel, CA
6. 2018 “The Informatics Consult: Filling the inferential gap in medicine.” Children’s Hospital of Orange County Medical Intelligence and Innovation Institute, Orange, CA
7. 2019 “The Green Button: Deploying an informatics consult service at an academic medical center.” HIMSS Physician Pre-Symposium, Orlando, FL
8. 2021 “Supporting COVID-19 patient management with data.” Stanford STARR Informatics Summit, Stanford, CA