Mary M Lucas

PhD Candidate, Information Science (Expected 2026)
mary.m.lucas@drexel.edu | (215) 581-8007 | LinkedIn | Website
Philadelphia, PA

PROFESSIONAL SUMMARY

PhD Candidate and Clinical AI Researcher with 10+ years of trauma and emergency nursing experience. Develops reliable, human-centered AI systems for high-stakes healthcare settings. Research spans algorithmic fairness and uncertainty quantification in predictive modeling, as well as prompting and evaluation strategies to improve the consistency and clinical reasoning performance of Large Language Models. Uses large-scale real-world evidence to bridge the gap between technical model performance and clinical utility.

EDUCATION

Doctor of Philosophy (Ph.D.) in Information Science | *Expected 2026* Drexel University, College of Computing & Informatics | Philadelphia, PA

- Dissertation Focus: Algorithmic Fairness, Uncertainty Quantification, and Clinical NLP and LLMs Explainability.
- Working Title: A Negotiated Learning Framework for AI Fairness and Uncertainty Quantification in Clinical Risk Prediction

Master of Science (M.S.) in Computer Information Systems | *2021* Boston University (Metropolitan College) | Boston, MA

- Concentrations: Health Informatics and Data Analytics (Dual Concentration)

Certificate of Achievement, Medical Statistics | Stanford Center for Health Education, Stanford Medicine | December 2021

Master of Philosophy (M.Phil.) in Physics | *2000* Moi University | Eldoret, Kenya

Bachelor of Science (B.Sc.) in Mathematics and Physics | *1997* Moi University | Eldoret, Kenya.

RESEARCH EXPERIENCE

Graduate Research Fellow (PhD Candidate) | College of Computing & Informatics, Drexel University | 2021 - Present

Algorithmic Fairness, Uncertainty Quantification, and Clinical NLP & LLMs

- Designed a fairness auditing framework for clinical risk prediction models across CKD and oncology datasets, integrating novel mitigation techniques and benchmarking them against existing approaches.
- Leveraged 10+ years of nursing experience to guide feature engineering and problem formulation, distinguishing physiological signals from artifacts of clinical workflow.

- Quantified algorithmic disparities in readmission prediction and treatment completion, showing how common fairness metrics often fail to capture systemic healthcare inequities.
- Developed **Ensemble Reasoning**, an iterative prompting framework that improved medical question-answering performance and consistency on USMLE datasets (+3-5%) across both closed (GPT-4) and open-source clinical LLMs (JAMIA).
- Designed and evaluated an LLM-assisted pipeline to automate extraction of population demographics from biomedical literature for equity analysis (IDCC).
- Developed a novel **Neighborhood-Adaptive Difficulty Score** combining kNN topology with conformal prediction to identify inherently hard-to-predict cases and to distinguish model weakness from case complexity.

Large-Scale Real-World Evidence (RWE) & Population Health Disparities Research

- Developed a mixed-order Markov chain simulation using VA Corporate Data Warehouse data to model prostate cancer treatment sequences and identify which empirically observed pathways are associated with elevated mortality risk for specific racial groups.
- Conducted a national registry study using the AAO IRIS® Registry to quantify racial and gender disparities in retinal vein occlusion treatment using multivariable logistic regression, highlighting inequities in access to anti-VEGF therapy with implications for equitable treatment initiatives.

Research Assistant | Health Informatics Lab, Boston University | 2019-2022

Conducted large-scale epidemiological modeling using IBM MarketScan databases, developing SQL workflows to process multi-million-row healthcare data files converted from SAS formats.

Research Fellow (Physics) | Department of Physics, Uppsala University, Sweden | 2000-2002

Conducted laboratory-based research on thin film technology as part of an international physics research collaboration.

PUBLICATIONS

Peer-reviewed Journal Articles

Mary M. Lucas, Mario Schootman, Jonathan A. Laryea, Sonia T. Orcutt, Chenghui Li, Jun Ying, Jennifer A. Rumpel, Christopher C. Yang (2024), "Bias in Prediction Models to Identify Patients With Colorectal Cancer at High Risk for Readmission After Resection", *JCO Clinical Cancer Informatics* **8**, e2300194(2024). https://doi.org/10.1200/CCI.23.00194

Mary M Lucas, Justin Yang, Jon K Pomeroy, Christopher C Yang (2024), "Reasoning with large language models for medical question answering", *Journal of the American Medical Informatics Association*, **31(9)**, 1964–1975, https://doi.org/10.1093/jamia/ocae131

Julia A. Haller, Maurizio Tomaiuolo, **Mary M. Lucas**, Christopher C. Yang, Leslie Hyman,

IRIS Registry Analytic Center Consortium (2024). “Disparities in Retinal Vein Occlusion Presentation and Initiation of Anti-VEGF Therapy: an Academy IRIS® Registry Analysis.” *Ophthalmology Retina*, **8(7)**, 657–665. https://doi.org/10.1016/j.oret.2024.01.017

Peer-reviewed Conference Proceedings

N. Nikita, T.D. Tran, C.C. Yang, J.A. Parrish, A.M.D. Hoedt, A. Hoffmeyer, A. Shaver, S. Sharma, **M. Lucas**, K.K. Zarrabi, S.J. Freedland, W.K. Kelly, G. Lu-Yao (2025), “Early unplanned hospitalization risk after novel hormonal therapy initiation in metastatic prostate cancer: Impact of disease type, treatment, and race”, *Journal of Geriatric Oncology*, Volume 16, Issue 8, Supplement, 2025, 102504, ISSN 1879-4068, https://www.geriatriconcology.net/article/S1879-4068(25)00320-0

Lucas, M.M., Yang, C.C. (2025). A Collaborative Learning Approach for Fairness in Prediction of Substance Use Disorder Treatment Completion. In: Bellazzi, R., Juarez Herrero, J.M., Sacchi, L., Zupan, B. (eds) *Artificial Intelligence in Medicine. AIME 2025. Lecture Notes in Computer Science()*, vol 15735. Springer, Cham. https://doi.org/10.1007/978-3-031-95841-0_44

Latrice Landry, **Mary Lucas**, Anietie Andy, Ebelechukwu Nwafor (2024). “Artificial Intelligence Assisted Curation of Population Groups in Biomedical Literature.” *International Journal of Digital Curation*, **18(1)**, pp.9. https://doi.org/10.2218/ijdc.v18i1.950

Chang, CH., Lucas, M.M., Lee, Y., Yang, C.C., Lu-Yao, G. (2024). “Beyond Self-consistency: Ensemble Reasoning Boosts Consistency and Accuracy of LLMs in Cancer Staging”. In: Finkelstein, J., Moskovitch, R., Parimbelli, E. (eds) *Artificial Intelligence in Medicine. AIME 2024*. Lecture Notes in Computer Science(), vol 14844. Springer, Cham. https://doi.org/10.1007/978-3-031-66538-7_23

M. M. Lucas, X. Wang, C. -H. Chang, C. C. Yang, J. E. Braughton and Q. M. Ngo (2024), “An ExplainableFair Framework for Prediction of Substance Use Disorder Treatment Completion,” 2024 IEEE 12th International Conference on Healthcare Informatics (ICHI), Orlando, FL, USA, 2024, pp. 157-166, https://doi.org/10.1109/ICHI61247.2024.00028

C. -H. Chang, **M. M. Lucas**, G. Lu-Yao and C. C. Yang (2024), “Classifying Cancer Stage with Open-Source Clinical Large Language Models,” *IEEE 12th International Conference on Healthcare Informatics (ICHI)*, Orlando, FL, USA, 2024, pp. 76-82, https://doi.org/10.1109/ICHI61247.2024.00018

Mary M. Lucas, Chia-Hsuan Chang, and Christopher C. Yang (2023). “Resampling for Mitigating Bias in Predictive Model for Substance Use Disorder Treatment Completion”, *11th IEEE International Conference on Healthcare Informatics (IEEE ICHI)*, Houston, TX, USA, pp. 709-711, https://doi.org/10.1109/ICHI57859.2023.00128

Ji, L., Li, Z., **Lucas, M.**, Vodenska, I., Chitkushev, L., Zhang, G.L. (2022). “Pregnancy Outcomes in Women with Pregestational Diabetes”, In: Zlateva, T., Goleva, R. (eds) *Computer Science and Education in Computer Science. CSECS 2022. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, vol 450. Springer, Cham. https://doi.org/10.1007/978-3-031-17292-2_10

Manuscripts Under Review / In Preparation

Development and validation of a prognostic risk score for one-year mortality in patients with metastatic prostate cancer initiating novel hormone therapies (VA EHR study)

Markov chain simulation study: racial differences in mortality risk associated with treatment sequences in prostate cancer patients, with attention to disparities (VA EHR study)

Explainable fairness enhancement in collaborative learning for healthcare prediction

Fairness and uncertainty in predictive modeling for prognostic prediction in SUD treatment completion and ICU mortality for sepsis patients

Poster Presentations

Mary M. Lucas, Christopher C. Yang, and Mario Schootman (2022), "Investigating health disparities and AI bias in models to predict development of chronic kidney disease in patients with type II diabetes", presented at *NIH AIM-AHEAD Annual Meeting 2023*. Rockville, MD. https://doi.org/10.6084/m9.figshare.23929128.v1

Prachiti Aras, Guanglan Zhang, **Mary Lucas**, Reza Rawassizadeh, Irena Vodenska, and Lou Chitkushev. "Quality Assessment of Inpatient Medical Claim Data", presented at *IEEE International Conference on Bioinformatics and Biomedicine 2020 (BIBM 2020)*.

Earlier publications in Physics available upon request / on personal website.

TECHNICAL SKILLS

Programming Languages: Python, R, SQL

AI & Machine Learning: Python (scikit-learn, PyTorch), XGBoost, Hugging Face (Transformers, PEFT/LoRA), Large Language Models (LLMs), Conformal Prediction (MAPIE), Fairness Auditing & Bias Mitigation, Uncertainty Quantification

Data Engineering & Health Informatics: SQL, OMOP Common Data Model, FHIR Interoperability Standards, Clinical Coding Systems (ICD-10, CPT, SNOMED) for cohort definition and data interpretation, Clinical Data Workflows (EHR, Registries)

Infrastructure & Tooling: Docker, Git/GitHub, AWS, GCP

LLMs & Clinical NLP: Prompt Engineering, Inference with OpenAI / Hugging Face / Anthropic / Gemini, Self-Hosted LLM Deployment, Clinical Text Processing

AWARDS AND FELLOWSHIPS

Edith Peterson Mitchell, MD Health Equity Travel Scholarship, ECOG-ACRIN Fall Group Meeting (2023, 2024)

Research Fellowship, NIH AIM-AHEAD Consortium (2022)

Excellence in Graduate Study in Computer Information Systems, Boston University (2020)

CLINICAL DOMAIN EXPERTISE

Medical Content Reviewer & Clinical Advisor *(Remote)* | *hims & hers* | 2020-2024

- Reviewed and validated patient-facing medical content against current scientific literature and standards of care.
- Provided evidence-based revisions, reference verification, and clinical guidance prior to medical director approval.
- Contributed to quality assurance processes for large-scale digital health education materials.

Clinical Precertification & Care Coordination RN | *Independence Blue Cross*, Philadelphia, PA | 2017 - 2022

- **Utilization Management (Oncology & Specialty Infusions)**: Conducted rigorous clinical reviews for high-complexity authorization requests (Chemotherapy, Direct Ship Biologics), verifying medical necessity against FDA labelling and payer coverage policies.
- **Clinical Evidence Evaluation**: Analysed longitudinal patient records—including disease progression, treatment history, and pathology reports—to ensure therapeutic alignment with standard-of-care protocols.
- **Relevance to Research**: Gained direct insight into the "approval logic" that dictates which patients receive specific treatments in real-world datasets, informing the design of selection criteria for retrospective studies.

Emergency Department Registered Nurse | Multiple High-Acuity Centers | 2008 - 2019
Locations: *Our Lady of Lourdes Medical Center* (Comprehensive Stroke & CV Center, NJ); *Auckland City Hospital* (Major Regional Trauma Center, New Zealand)

- **High-Acuity Clinical Management**: Provided direct emergency care for complex patient populations, utilizing rapid assessment protocols and triage logic (ESI) to prioritize resource allocation in high-pressure environments across diverse tertiary environments, ranging from a regional cardiovascular referral center to a major trauma hub serving the Pacific region.
- **Domain Knowledge**: Deep expertise in the pathophysiology of acute deterioration, providing the clinical intuition necessary to validate features in clinical prediction models.
- **Clinical Data Generation**: Managed high-volume documentation across diverse EHR platforms (Epic, Cerner), gaining first-hand insight into clinical workflow bottlenecks that cause "data leakage" and informative missingness in electronic health records.
- **Nurse Educator** (Secondment, 2014-2015): Developed and delivered clinical training protocols for ER nurses, bridging the gap between theoretical guidelines and practical

bedside implementation.

LICENSURE: Registered Nurse - Pennsylvania (active); New Jersey (inactive); New Zealand (inactive)

TEACHING EXPERIENCE

Instructor of Record | Drexel University, College of Computing & Informatics | 2024 - 2025

Course: *INFO 896: Health Informatics Capstone (Graduate)*

Project Supervision: Led and mentored semester-long capstone research projects, advising master's students through topic development, proposal writing, methodological planning, technical implementation, analysis, and presentation of results.

Research Guidance: Advised students on research methodology, ensuring projects aligned with current industry standards in health data interoperability and system design.

Graduate Teaching Assistant (Data Science & Health IT) | Boston University (Metropolitan College) | 2021 - Present

Courses: *Data Science with Python (CS 677)*; *Information Structures with Python (CS 521)*; *Biomedical Sciences & Health IT (CS 570)*; *Health Informatics (CS 580)*; *Ethical & Legal Issues in Health Informatics (CS 584)*

Technical Mentorship: Provided code reviews and debugging support (Python/SQL/R) for a diverse cohort of working professionals in an asynchronous online environment.

Instructional and Domain Support: Supported instruction across programming, biomedical foundations, health informatics, and ethical/legal governance of health data. Translated complex concepts - including data structures, ML pipelines, database logic, clinical data workflows, and health data policy - into accessible learning frameworks for non-traditional students, effectively bridging clinical, technical, and policy domains.

Graduate Teaching Assistant (Information Science) | Drexel University | 2025 - Present

Course: *INFO 102 - Introduction to Information Systems*

Instructional Support: Assisted in teaching foundational concepts of information systems, including systems thinking, data and information flow, organizational use of technology, and core principles of information management.

Student Mentorship: Supported students in applying information science principles to real-world contexts through discussions, assignments, and guided feedback.

Lecturer and Teaching Assistant, Physics & Math | Moi University; University of Rhode Island (2000-2004)

Foundational STEM Instruction: Delivered lectures and lab supervision for undergraduate physics and mathematics, establishing a rigorous foundation in quantitative pedagogy and scientific communication.

ACADEMIC LEADERSHIP AND SERVICE

Editorial Positions

Managing Editor: Journal of Healthcare Informatics Research (2022-present)

Scientific Grant Review

NIH AIM-AHEAD Consortium: Reviewer for Fellowships and Training Programs (2023-2025).

Conference Leadership & Organization

Co-Chair: JHIR Special Session, for the 12th IEEE International Conference on Health Informatics (2024)

Organizing Committee:

The First Workshop on Applying LLMs in LMICs for Healthcare Solutions, for the 12th IEEE International Conference on Health Informatics (2024)

Women in Healthcare Informatics Event, for the 12th IEEE International Conference on Health Informatics (2024)

Community Leadership:

Event Organizer: DataPhilly (5000+ member Data Science Meetup Group), Philadelphia, PA (2019-2020)

Ad Hoc Peer Review

Journals: Journal of the American Medical Informatics Association (JAMIA); Information Processing and Management; Journal of Healthcare Informatics Research (JHIR)

Conferences: IEEE International Conference on Healthcare Informatics (ICHI) (2026); The EUSSET Conference on Computer-Supported Cooperative WorkECSCW (2026)

Institutional Service:

Curriculum Committee: Undergraduate Information Science - Drexel University (2025)

AFFILIATIONS

Professional Societies

- American Medical Informatics Association (AMIA) - Member (2022-present)
- American Statistical Association - Member (2025-present)
- Association for Computing Machinery (ACM) - Member (2022-present)
- Institute of Electrical and Electronics Engineers (IEEE) - Member (2022-present)

Research Consortia and Networks

- ECOG-ACRIN - Associate Member (2023-present)
- Cancer and Aging Research Group (CARG) - Associate Member (2025-present)