| BIOGRAPHICAL SKETCH | | | |
|---------------------------------------|--------------------------------|------|--|
| NAME | Kharrazi, Hadi H.K. | | |
| eRA COMMONS USERNAME | Kharrazi | | |
| POSITION TITLE | Professor & Center Co-Director | | |
| EDUCATION/TRAINING | | | |
| INSTITUTION AND LOCATION | DEGREE | DATE | FIELD OF STUDY |
| Dalhousie University, Halifax, Canada | PhD | 2008 | Interdisciplinary (Clinical Informatics) |
| Dalhousie University, Halifax, Canada | МНІ | 2005 | Health Informatics |
| IUMS, Tehran, Iran | MD | 2003 | Medicine |

(A) Personal Statement

I am the co-director of the Johns Hopkins Center for Population Health IT (CPHIT) and the research director of the Johns Hopkins ACG (hopkinsacg.org). Since 2012, I have managed a research team of 15+ clinicians, analysts, and data scientists focusing on utilizing massive real-world clinical data (RWD) sources to develop predictive models of healthcare utilization. CPHIT faculty and scientists are considered international leaders in developing and evaluating advanced risk stratification models using large RWD, which are adopted by 30+ U.S. health plans, 200+ U.S. health systems, and 20+ international governments. More than 250 million patient records are risk stratified and managed annually using the predictive models developed by CPHIT and distributed through the Johns Hopkins ACG solution. Our RWD data sources often involve EHR and insurance claims records of 80+ million patients. Our other data sources include individual-level social determinants of health (SDOH) and massive geo-derived SDOH predictors extracted and harmonized from 25+ publicly available data sources. I serve on the data governance boards of Johns Hopkins Schools of Public Health and Medicine.

My personal research focuses on the application of informatics solutions to advance the science of population health predictive analytics. My recent research focuses on assessing the effect of real-world data quality and bias in predicting healthcare outcomes. I have meticulously evaluated the value of new data types such as SDOH for population health analytics and assessed the challenges of novel methods to extract usable data from EHRs (e.g., bias in NLP methods). I have served as PI or co-PI on many federal grants and contracts (e.g., NIH, ONC, AHRQ, FDA, VHA), totaling 56+ million, which often has required the development of large-scale platforms to acquire, collate, and harmonize centralized clinical data warehouse. Some of these platforms have leveraged HIE infrastructures to create population health analytic platforms. I have actively participated and contributed to the Hopkins' Economics of Alzheimer's Disease & Services (HEADS) center, and the N3C (National COVID Cohort Collaborative) effort. I am the associate editor for Medical Care and JAMIA Open journals; and serve as a standing member on the NIMH SERV study section.

Role in the Proposed Project: I will serve as an MPI for the A4C project and support the management of A4C. Given my expertise with RWDs, data science, and informatics, I will lead the efforts for Aim 3 and 4. In Aim 3, I will work closely with our University of Michigan collaborators to direct the assessment of CDE validity using a longitudinal format of CMS and LINKAGE data. My background in data quality will enable me to adjust the CDE validation framework as needed. In Aim 4, I will work with A4C data managers and data scientists to evaluate and refine the validation of CDEs using a series of research and real-world data sources. I have previously contributed to the design, development, or analysis of JHMI Enclave, N3C, and OCHIN RWD sources, which will be used in Aim 4. My experience in serving on multiple NIH review committees will help me to lead the scientific review committee of A4C for Aim 5. As an MPI, I will also serve on the steering committee (as a non-voting member). Considering my background in population health and multimorbidity risk analysis, I will co-lead Domain #2 (disease characteristics, including ADRD and MCCs) in Aim 1. In addition to overseeing the day-to-day operations of the project, I will directly supervise several data managers and data scientists involved in Aims 3 and 4 and collaborate with the A4C NLP experts. The following grants and publications support my expertise in data science, predictive modeling, and experience with RWD sources such as claims and EHR data:

NLM (1T15LM013979) H. Kharrazi & C. Chute (Co-PIs) Johns Hopkins Training Program in Biomedical Informatics and Data Science

FDA CERSI (5U01FD005942)

H. Kharrazi & J. Weiner (Co-PIs) Assessing disparities in occurrence and outcomes of type 2 diabetes ADEs in minority populations using real world administrative claims and electronic health records

NIMH (R01MH124724)

H. Kharrazi (PI) Advancing Maryland's Statewide Suicide Data Warehouse to Improve Individual and Population-level Mortality Prediction and Prevention

AHRQ ACTION-IV (K72AS19.25505)

H. Kharrazi (PI)

Developing and Assessing the Validity of Claims-based Indicators of Frailty & Functional Disabilities and Testing their Use in Other Data (EHRs and linked EHR-claims)

Citations:

- a. <u>Kharrazi H</u>, Ma X, Chang H-Y, Richards TR, Jung C. Comparing the predictive effects of patient medication adherence indices in EHR and claims-based risk stratification models. *Popul Health Manag.* 2021; 24(5): 601-609. PMID: 33544044.
- b. <u>Kharrazi H</u>, Lasser E, Yasnoff WA, Loonsk J, Advani A, Lehmann H, Chin D, Weiner JP. A proposed national research and development agenda for population health informatics: summary recommendations from a national expert workshop. *J Am Med Inform Assoc.* 2017; 24 (1): 2-12. PMID: 27018264.
- c. <u>Kharrazi H</u>, Chi W, Chang HY, Richards TM, Gallagher JM, Knudson SM, Weiner JP. Comparing population-based risk-stratification model performance using data extracted from electronic health records versus administrative claims. *Med Care*. 2017; 55 (8): 789-796. PMID: 28598890.
- d. <u>Kharrazi H</u>, Wang C, Scharfstein D. Prospective EHR-based clinical trials: the challenge of missing data. *J Gen Intern Med.* 2014; 29(7) 976-978. PMID: 24839057.

(B) Positions, Scientific Appointments and Honors

Positions and Scientific Appointments

| 2024 – Now | Professor Johns Hopkins School of Public Health, Health Policy and Management (HPM) Johns Hopkins School of Medicine, Biomedical Informatics and Data Science (BIDS) Co-Director, Johns Hopkins Center for Population Health IT (CPHIT) Research Director, Johns Hopkins ACG Solution Director, DrPH Informatics training program Director, PhD Health Informatics training program |
|----------------|---|
| 2021 – Now | Deputy Editor, Medical Care |
| 2021 – Now | Associate Editor, JAMIA Open |
| 2019 – 2023 | Associate Professor, Johns Hopkins School of Public Health |
| 2012 – 2019 | Assistant Professor, Johns Hopkins School of Public Health |
| 2012 – Now | Member, Academy Health (AH) (HIT Interest Group Advisory Committee) |
| 2005 – Now | Member, American Medical Informatics Association (AMIA) (PHI Executive Committee) |
| <u>Honors</u> | |
| 2022 – present | Fellow of the American College of Medical Informatics (FACMI) |
| 2020 – present | Fellow of the American Medical Informatics Association (FAMIA) |

08/2020 - 08/2025

09/2020 – 09/2025

09/2020 - 09/2025

(C) Contribution to Science

1. Population Health Informatics (PHI) & Risk Stratification Using Real World Data

My research focuses on the application and evaluation of informatics solutions within the context of population health. This emerging and rapidly growing domain of research and development is called "Population Health Informatics" (PHI). A key role of PHI is to improve the population health analytic cycle, which starts with data collection, followed by data preparation, data mining, model development & validation, knowledge sharing, and finally closing the process by a learning health system that applies these models and feeds new data back in the loop. I am specifically interested to assess the opportunities and challenges of integrating non-traditional data sources to improve population health analytics and eventually enhance public health interventions and outcomes.

My work has contributed to community health IT solutions, population-based decision support systems, and population-wide predictive models to forecast utilization and specific health outcomes. Some of my work has been devoted to the research and development of the Johns Hopkins ACG (hopkinsacg.org), which is considered one of the top patient-level population-wide risk prediction solutions nationally and globally. The ACG system offers a unique approach to measuring morbidity that improves accuracy in evaluating provider performance, identifying patients at high risk, forecasting healthcare utilization, and setting equitable payment rates. Billions of dollars per year are now routinely exchanged using ACGs in almost every U.S. State and in 20+ nations. Over 1200+ peer reviewed articles have been published that apply and evaluate ACGs. The following are recent publications led by me further enhancing population risk stratification models:

- a. Howson SN, McShea MJ, Ramachandran R, Burkom HS, Chang HY, Weiner JP, <u>Kharrazi H</u>. Improving the prediction of persistent high healthcare utilizers: using an ensemble methodology. *JMIR Med Inform*. 2022; 10(3): e33212. PMID: 35275063.
- b. Ferris LM, Weiner JP, Saloner B, <u>Kharrazi H</u>. Comparing person-level matching algorithms to identify risk across disparate datasets among patients with a controlled substance prescription: retrospective analysis. *JAMIA Open*. 2022; 5(1): ooac020. PMID: 35571361.
- c. Ramachandran R, McShea MJ, Howson SN, Burkom HS, Chang H-Y, Weiner JP, <u>Kharrazi H</u>. Assessing the value of unsupervised clustering in detecting key classes of diagnostic and medication codes to improve the prediction of persistent high healthcare utilizers. *JMIR Med Inform.* 2021. 9(11): e31442. PMID: 34592712.
- Marrazi H, Chang HY, et al. Enhancing the prediction of healthcare costs and utilization by including outpatient BMI values to diagnosis-based risk models. *Med Care*. 2018; 56(12): 1042-1050. PMID: 30339574.

2. Social Determinants of Health (SDOH)

I have extensively collaborated with experts in the social determinants of health (SDOH) domain. SDOH explains a large variation of healthcare outcomes including utilization and costs. My research has mainly focused on integrating individual or geographical-level SDOH with clinical data (e.g., EHRs or claims data) to improve the prediction of various outcomes such as hospitalization, emergency room admission, and medical cost. These efforts have led our research team to investigate potential biases in our predictive models of care (e.g., models being biased toward a specific race or socio-economic status of patients). We have also started to examine potential bias in our modeling techniques, which may lead to unrecognized differences in healthcare access. The following are recent publications supporting my work in SDOH and population health:

- a. Wang HE, Landers M, Adams R, Subbaswamy A, <u>Kharrazi H</u>, Gaskin DJ, Saria S. Identifying bias in health care predictive models: Development of a modeling bias screening checklist and its pilot application for 30-day hospital readmission models. *J Am Med Inform Assoc.* 2022; 29(8): ocac065. PMID: 35579328.
- b. Lasser E, Gudzune K, Lehmann H, <u>Kharrazi H</u>, Weiner JP. Trends and patterns of social and behavioral determinants of health data collection within an electronic health record. Popul Health Manag. 2022; 26(1): 13-21. PMID: 36607903.
- c. Hatef E, Nau C, …, Lyons LJ, <u>Kharrazi H</u>, Weiner JP, Abu-Nasser M, Roblin D. Assessing the documentation of social needs in electronic health records' unstructured data: A comparison of three integrated healthcare delivery systems. *JAMIA Open.* 2022; 5(1) ooac006. PMID: 35224458.
- d. Tan M, Hatef E, Taghipoor E, Vyas K, <u>Kharrazi H</u>, Gottlieb L, Weiner JP. Including social and behavioral determinants in predictive models: trends, challenges, and opportunities. *JMIR Med Inform.* 2020; 8(9) e18084. PMID: 32897240.

3. New Methods & Data Quality

A major contribution of PHI to population health analytics is providing new methods to extract novel types of data from various data sources. For example, the free text of EHR includes ample information about individual patients that can be used in risk stratification efforts; however, this extra information is often missing in encoded fields of EHRs. In a prior study we assessed the value of free text in identifying geriatric syndromes (which are predictable for utilization). We also assessed how the physician's mention of frailty in the free text is associated with this information. However, the variable data quality levels of EHRs makes the use of such data for population health analytics a challenging task. We are currently evaluating the added value of free text for social determinants of health as well as predictors of various clinical outcomes (e.g., suicide death, adverse drug events). Following is a select list of recent publications supporting my contributions to the methods of clinical and population health informatics:

- a. Madlock-Brown C, …, <u>Kharrazi H</u>, Liu F, Lemon C, Adams WG. Clinical, social, and policy factors in COVID-19 cases and deaths: methodological considerations for feature selection and modeling in county-level analyses. *BMC Public Health*. 2022; 22(1) 747. PMID: 35421958.
- b. Chen T, Dredze M, Weiner JP, Hernandez L, Kimura J, <u>Kharrazi H</u>. Extraction of geriatric syndromes from electronic health record clinical notes: assessment of statistical natural language processing methods. *JMIR Med Inform*. 2019; 7(1): e13039. PMID: 30862607.
- Pandya CJ, Hatef E, Wu J, Richards T, Weiner JP, <u>Kharrazi H</u>. Impact of social needs in electronic health records and claims on health care utilization and costs risk-adjustment models within Medicaid population. *Popul Health Manag.* 2022 Oct; 25(5): 658-668. Epub 2022 Jun 23. PMID: 35736663.
- d. Kan HJ, <u>Kharrazi H</u>, Chang HY, et al. Exploring the use of machine learning for risk adjustment: a comparison of standard and penalized linear regression models in predicting health care costs in older adults. PLoS One. 2019; 14(3): e0213258. PMID: 30840682.

4. Public Health Informatics

My research also focuses on the application of health IT innovation within the traditional public health research. My recent work in this area has propelled the notion of "bridging" population health IT efforts and public health IT efforts:

- a. <u>Kharrazi H</u>, Gamache R, Weiner JP. Role of Informatics in Bridging Public and Population Health. Magnuson JA & Dixon BE (Eds.) *Public Health Informatics and Information Systems*; 3rd edition. 2020; 59-79. London UK: Springer-Verlag. ISBN 978-3-030-41215-9
- b. Hatef E, Weiner JP, <u>Kharrazi H</u>. A public health perspective on using electronic health records to address social determinants of health: The potential for a national system of local community health records in the United States. *Int J Med Inform.* 2019 Apr; 124:86-89. PMID: 30784431.
- c. <u>Kharrazi H</u>, Lehmann HP. Role of Population Health Informatics in Understanding Data, Information and Knowledge. In Joshi A (Ed.) *Population Health Informatics*. 2017; 65-89. Jones and Bartlett Learning. ISBN: 978-1-28-410396-0
- d. Gamache R, <u>Kharrazi H</u>, Weiner JP. Public health and population health informatics: the bridging of big data to benefit communities. *Yearb Med Inform*. 2018; 27(1): 199-206. PMID: 30157524.

5. ADRD & Geriatrics

Since 2012, I have studied the measurement and identification of geriatric frailty using real world data sources such as insurance claims and EHR data. More recently, through the collaboration with the Hopkins' Economics of Alzheimer's Disease & Services (HEADS) center, I have piloted several projects extracting ADRD markers from EHR's free text and understanding the potential racial disparities in such measurements (i.e., bias in the quality of clinical notes leading to racial bias in identifying ADRD patients). The following includes a sample of my publications focusing ADRD RWD and the extraction of geriatric frailty indicators from RWD such as EHRs and insurance claims:

- a. Socal MP, Odouard IC, <u>Kharrazi H</u>. Ownership and Interoperability Challenges of Alzheimer Monoclonal Antibody Registries. *JAMA Neurol.* 2024; 81(2): 109-110. PMID: 38079170.
- Bery A*, Anzaldi LJ, Boyd CM, Leff B, <u>Kharrazi H</u>. Potential value of electronic health records in capturing data on geriatric frailty for population health. *Arch Gerontol Geriatr.* 2020; 91:104224. PMID: 32829083.
- c. <u>Kharrazi H</u>, Anzaldi L, Hernandez L, Davison A, Boyd CM, Leff B, Kimura J, Weiner JP. Measuring the value of electronic health record's free text in identification of geriatric syndromes. *J Am Geriatr Soc.* 2018; 66(1) 1499-1507. PMID: 29972595.
- d. Chen T, Dredze M, Weiner JP, <u>Kharrazi H</u>. Identifying vulnerable older adult populations by contextualizing geriatric syndrome information in clinical notes of electronic health records. *J Am Med Inform Assoc.* 2019; 26(8-9): 787–795. PMID: 31265063.

Published Work (135+ papers; H-index: 42)

o PubMed: https://www.ncbi.nlm.nih.gov/myncbi/10eHcnhXecnkr/bibliography/public/