



CENTER FOR ARTIFICIAL INTELLIGENCE RESEARCH

**Wake Forest Center for
Artificial Intelligence Research**

2024 WORKSHOP
**AI in Healthcare: Transforming
Medicine for a Better Future**

Friday, June 28, 2024

Book of Abstracts



July 3, 2024

Dear Colleagues and Participants,

On behalf of the Center for Artificial Intelligence Research (CAIR) at Wake Forest University School of Medicine, I am delighted to present the proceedings of our recent workshop, held on June 28th, 2024. The workshop brought together researchers, practitioners, and trainees to explore the latest advancements in AI applications in healthcare within our Advocate Health family and Wake Forest University School of Medicine.

The workshop featured diverse presentations and discussions, covering a wide range of topics, including AI algorithms in health applications, emerging AI technologies for healthcare, equitable AI, and advances in AI for drug discovery. Our keynote speaker, Dr. Susan Gregurick, Associate Director for Data Science and Director of the Office of Data Science Strategy (ODSS) at the National Institutes of Health (NIH), provided invaluable insights that have undoubtedly enriched our collective understanding and inspired further research in these critical areas.

This Book of Abstracts serves as a comprehensive compilation of the innovative research presented at the workshop. Each abstract has been carefully reviewed and selected to reflect the high caliber of work being conducted in the field of AI and healthcare. We are proud to showcase the contributions of our participants, whose dedication and expertise are driving the future of medical technology and improving patient care.

We are also pleased to announce that Dr. V. Christian Sanderfer, MD, and Dr. Shekhar Singh, PhD, were named the winners of two new travel awards given by CAIR at this year's workshop. Both recipients will receive a \$1,500 prize to be used within a year for attending and presenting at an AI-related scientific conference or enrolling in an AI certificate program. We congratulate Dr. Sanderfer and Dr. Singh on their achievements and look forward to their continued contributions to the field of AI in healthcare.

Our center remains committed to fostering collaboration, advancing research, and promoting education in artificial intelligence. We believe that the exchange of ideas and knowledge at events like this workshop is essential to achieving our mission of creating AI-driven solutions that have a tangible impact on healthcare delivery.

I would like to extend my gratitude to all the speakers, attendees, session chairs, and organizing committee members whose efforts made this workshop a resounding success. Your engagement and contributions are instrumental in pushing the boundaries of what AI can achieve in medicine. Special thanks are due to Brittany Jones, whose tireless efforts and dedication were instrumental in making this workshop a resounding success. Her hard work and commitment ensured that every detail was meticulously planned and executed, and her contributions are deeply appreciated.

We look forward to continuing our work together and to the many future opportunities for collaboration and discovery. Please do not hesitate to reach out to us with any questions or for further information about our ongoing projects and initiatives.

Sincerely,

Metin N. Gurcan, PhD
Senior Associate Dean for Artificial Intelligence
Director, Center for Artificial Intelligence Research
Professor, Internal Medicine
Wake Forest University School of Medicine
CAIR@wakehealth.edu

Index:

Session I: Keynote from Susan K. Gregurick, PhD

Chaired by Metin Gurcan, PhD

Keynote: AI Activities, AI Education opportunities, and AI protection and security

Susan K. Gregurick, PhD

Associate Director for Data Science
Director of the Office of Data Science Strategy
National Institutes of Health

Session II: AI Applications in Healthcare Talks

Chaired by Martin Kohn, MD, MS, FAMIA, FACEP, FACPE

Predicting Time-to-Conversion for Dementia of Alzheimer's Type Using Multi-Modal Neuroimaging Genomics Deep Survival Analysis - Sophia Kyriacou, Ghazal Mirabnahrazam, Cédric Beaulac, Sieun Lee, Karteek Popuri1, Hyunwoo Lee, Jiguo Cao, Lei Wang, James E Galvin, Mirza Faisal Beg, Suzanne Craft, Metin Gurcan, Da Ma, PhD

AI-Derived Body Composition and Post-Operative Outcomes After Minimally-Invasive Esophagectomy – V. Christian Sanderfer MD, Ansley Beth Ricker MD, Alexis Holland MD, Erin Donahue PhD, Reilly Shea BS, Matigan Simpson BA, Katelyn Cockerham BS, Nicholas Mullis BS, Sophia Bellavia, Ella Schwarzen, M. Hart Squires MD, Kunal C. Kadakia MD, Jonathan C. Salo MD, Department of Surgery, Carolinas Medical Center, Atrium Health, Charlotte, NC, Department of Surgical Oncology, Levine Cancer Institute, Atrium Health, Charlotte

Towards Remote and Needleless Monitoring of NT-proBNP from ECG via IA - Ibrahim Karabayir, PhD, Melissa Hudson, Kristen Ness, Daniel Mulrooney, Stephanie Dixon, Turgay Celik, David McManus, Robert L. Davis, Oguz Akbilgic

Can AI-Powered Clinical Documentation Improve Clinicians' Efficiency and Satisfaction?

- Tsai-Ling Liu, PhD, MSPH, Timothy C. Hetherington, MS, Andrew McWilliams, MD, MPH, Ajay Dharod, MD, FACP, Tracey Carroll, RHIT, MHA, MBA, OCSA, ITIL, Richa Bundy, MPH, Hieu Nguyen, MS, McKenzie Isreal, MPH, Jeffrey A. Cleveland, MD, FAAP

Session III: AI-Driven Solutions in Medicine Talks

Chaired by Brian Wells, MD, PhD

Machine Learning Algorithm for Reduction of Pre-Settling Times of Human Body Models – Seth Mischo, Wade von Kleeck, Katherine Mahon, F. Scott Gayzik

Exploring the Host Response in Infected Lung Organoids Using NanoString Technology: A Machine Learning Analysis of Gene Expression Data – Mostafa Rezapour, PhD, Muhammad Khalid Khan Niazi, PhD; Stephen J. Walker, PhD, David A. Ornelles, PhD; Patrick M. McNutt, PhD; Anthony Atala, MD; Metin Gurcan, PhD

Deep Learning for AFib and Sleep Apnea Detection Using Multimodal Sleep Study Data – Shekhar Singh, PhD, Liam Butler, Ibrahim Karabayir, Oguz Akbilgic

Enhancing Clinical Documentation Efficiency Through AI for Hospitalist at Home Program – Usman Afzaal, Padageshwar Rao Sunkara, Sneha Chebrolu, Sumera Andleeb, Suneel Parvathareddy, Metin Gurcan

Poster Session Abstracts

Chaired by Arezoo Movaghar, PhD

Remote Maternal Health Monitoring: 9+IAI Co-Pilot App - Luke Patterson, MS1, Ibrahim Karabayir, PhD1, Robert L

Davis, MD, MPH², Oguz Akbilgic, DBA, PhD¹, Wake Forest School of Medicine, Winston-Salem, NC , University of Tennessee Health Science Center, Memphis, TN

Development of a multi-modal artificial intelligence model to detect degrees of hemorrhage - Fahim Usshibab Mobin¹, Antonio C. Renaldo¹, Micaela K. Gomez¹, Sandra Januszko¹, Jacob Dooley¹, James E. Jordan², Oguz Akbilgic³, Timothy K. Williams⁴, and Elaheh Rahbar^{1,5}. 1. Department of Biomedical Engineering, Wake Forest University School of Medicine 2. Department of Cardiothoracic Surgery 3. Center for Artificial Intelligence Research 4. Department of Vascular and Endovascular Surgery 5. Department of Biomedical Engineering, Texas A&M University

Development of a batch process to mine and analyze carotid duplex ultrasound - Fahim Usshibab Mobin¹, Micaela K. Gomez¹, Timothy K. Williams², Aarti Sarwal³, Oguz Akbilgic⁴, and Elaheh Rahbar^{1,5}. 1. Department of Biomedical Engineering, Wake Forest University School of Medicine 2. Department of Vascular and Endovascular Surgery 3. Department of Neurology 4. Center for Artificial Intelligence Research 5. Department of Biomedical Engineering, Texas A&M University

Enhancing Colorectal Cancer Tumor Bud Detection Using Deep Learning from Routine H&E-Stained Slides - Osama Sajjad^a, Ziyu Su^a, Mostafa Rezapour^a, Thomas Tavolara², Wendy L. Frankel³, Metin N. Gurcan¹, Wei Chen³, M. Khalid Khan Niazi¹. Center for Artificial Intelligence Research, Wake Forest University School of Medicine, Winston-Salem, NC, USA; Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA; Department of Pathology, The Ohio State University, Columbus, OH, USA;

CAIR AI in Healthcare Workshop: Revolutionizing Pathway Analysis with Deep Learning: The Gene PointNet Method - Hao Lu^{1}, Mostafa Rezapour¹, Haseebullah Baha², Muhammad Khalid Khan Niazi¹, Aarthi Narayanan³, Metin Nafi Gurcan¹. Center for Artificial Intelligence Research, Wake Forest University School of Medicine, Winston- Salem, NC, USA. 2.School of Systems Biology, College of Science, George Mason University, Fairfax, VA 22030, USA. 3.Department of Biology, George Mason University, Fairfax, VA, USA.*

*Automatic Tympanic Membrane Segmentation from Otoscopic Videos Using the SAM-Adapter
Author: Seda Camalan, PhD*

*Collective-Intelligence Recommender Systems: Advancing Health Communication into the 21st Century -Thomas Houston, MD MPH^a; Rajani Sadasivam, PhD^b; Peter Lindenauer, MD MSc^c; Daniel Ford, MD MPH^d
a Wake Forest University School of Medicine, Department of Internal Medicine; b UMass Chan School of Medicine, Worcester, MA; c UMass Chan Medical School - Baystate Regional Campus, Springfield, MA. USA; d Johns Hopkins University, Baltimore, MD*

Artificial Intelligence in Medical Physics and the AAPM AI Initiative - J. Daniel Bourland, PhD - Professor and Physicist, Department of Radiation Oncology, Biomedical Engineering and Physics Wake Forest University School of Medicine

*Multimodal Radio-Pathomics and Interpretable Machine Learning for Predicting Immunotherapy Response in Gastric Cancer - Yuming Jiang¹, James Edward Han¹, Qing Lyu², Fan Yang³
1Department of Radiation Oncology and 2Radiology, Wake Forest University School of Medicine; 3Department of Computer Science, Wake Forest University.*

*Deep-ODX: An efficient deep learning tool to risk stratify breast cancer patients from histopathology images - Ziyu Su¹, Amanda Rosen², Robert Wesolowski³, Gary Tozbikian⁴, M. Khalid Khan Niazi¹, Metin N. Gurcan¹
Center for Artificial Intelligence Research, Wake Forest University School of Medicine. 2. Department of Biomedical Engineering, Duke University. 3. Division of Medical Oncology, Comprehensive Cancer Center, The Ohio State University College of Medicine Department of Pathology, The Ohio State University.*

Session II Abstracts

CAIR AI in Healthcare Workshop: Predicting time-to- conversion for dementia of Alzheimer's type using multi-modal neuroimaging genomics deep survival analysis

Authors

Sophia Kyriacou^{1,2}, Ghazal Mirabnahrazam³, Cédric Beaulac⁴, Sieun Lee³, Karteeck Popuri^{1, 4}, Hyunwoo Lee⁵, Jiguo Cao¹, Lei Wang⁶, James E Galvin⁷, Mirza Faisal Beg³, Suzanne Craft², Metin Gurcan², Da Ma²

¹*Wake Forest University School of Medicine*

²*Georgia Tech Institute of Technology*

³*Simon Fraser University - Burnaby (Canada)*,

⁴*University Of Victoria - Victoria (Canada)*,

⁵*University Of Nottingham - Nottingham (United Kingdom)*,

⁶*Memorial University Of Newfoundland - St. John's (Canada)*

⁵*University Of British Columbia – Vancouver (Canada)*,

⁶*Ohio State University Wexner Medical Center - Columbus (United States)*,

⁷*University Of Miami Miller School Of Medicine - Miami (United States)*

Abstract

Background and Objectives

Dementia of Alzheimer's Type (DAT) is a complex disorder influenced by numerous factors, and it is difficult to predict individual progression trajectory from normal or mildly impaired cognition to DAT. An in-depth examination of multiple modalities of data may yield an accurate estimate of time-to-conversion to DAT for preclinical subjects at various stages of disease development. This study aims to predict subjects' time-to-conversion to DAT using multi-modal neuroimaging genomic input.

Methods

We used 543 subjects with 63 features from 3 different data modalities from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. The following modalities were used: 1) MRI, 2) genetic and 3) and CDC (Cognitive tests, Demographic, and CSF) data. Deep survival analysis was conducted using DeepSurv method using a multi-layer perceptron (MLP) model architecture as the feature extractor. The 63 most important features were automatically selected for the three modalities. We define the non-AD-progression as "survivor", and AD-progression as "non-survivor". The subjects were divided into two groups: progressive subjects (non-survivor), who were either healthy or diagnosed with Mild Cognitive Impairment (MCI) at the initial clinical visit and later developed AD, and non-progressive subjects ("survivor"), who were either healthy or MCI at the initial visit but did not develop AD later. We used 10 random sub-samples, selecting 80% of the subjects for training and 20% for testing each time; 20% of training data was used for internal validation.

Results

MRI data in predicting DAT time-to-conversion for subjects with Mild Cognitive Impairment (MCI). On the other hand, genetic data provided the most predictive power for subjects with Normal Cognition (NC) at the time of the visit. Furthermore, combining MRI and genetic features improved the time-to-event prediction over using either modality alone. Finally, adding CDC to any combination of features only worked as well as using only the CDC features.

Conclusions

Our study demonstrated that using powerful predictive models on multi-modal data can improve prediction of time-to-conversion. This not only leads to a better understanding of AD, but it also provides essential tools for practitioners who wish to follow their patients' disease progression.

Keywords

Alzheimer's disease, Deep Survival Analysis, Early detection, Multi-modal deep learning, Neuroimage genomics

Acknowledgments

Funding for this research is gratefully acknowledged from Wake Forest School of Medicine Center for Artificial Intelligence Research Biomedical Informatics Pilot Award from, Wake Forest Alzheimer's Disease Research Center with funding from the National Institute on Aging under award number (P30AG072947), the National Institute on Aging (R01 AG055121-01A1, R01 AG069765-01, R01 AG071514-01), National Institute of Neurological Disorders and Stroke (NINDS) (R01 NS101483- 01A1), Alzheimer Society Research Program, National Science Engineering Research Council (NSERC), Canadian Institutes of Health Research (CIHR), Fondation Brain Canada, Pacific Alzheimer's Research Foundation, the Michael Smith Foundation for Health Research (MSFHR), Precision Imaging Beacon, University of Nottingham, and Canadian Statistical Sciences Institute (CANSSI).

References:

1. Kvamme, H., Borgan, Ø., & Scheel, I. (2019), 'Time-to-event prediction with neural networks and Cox regression', *Journal of Machine Learning Research*, vol. 20, pp. 1-30.
2. Lee C, Yoon J, Van Der Schaar M. Dynamic-deephit: A deep learning approach for dynamic survival analysis with competing risks based on longitudinal data. *IEEE Transactions on Biomedical Engineering*. 2019 Apr 3;67(1):122-33.
3. Mirabnahrazam G, Ma D, Lee S, Popuri K, Lee H, Cao J, Wang L, Galvin JE, Beg MF, Alzheimer's Disease Neuroimaging Initiative. Machine learning based multimodal neuroimaging genomics dementia score for predicting future conversion to alzheimer's disease. *Journal of Alzheimer's Disease*. 2022 Jan 1;87(3):1345-65.
4. Ma D, Yee E, Stocks JK, Jenkins LM, Popuri K, Chausse G, Wang L, Probst S, Beg MF. Blinded clinical evaluation for dementia of Alzheimer's type classification using FDG-PET: A comparison between feature-engineered and non-feature-engineered machine learning methods. *Journal of Alzheimer's Disease*. 2021 Jan 1;80(2):715-26.
5. Lu P, Colliot O. Multilevel survival modeling with structured penalties for disease prediction from imaging genetics data. *IEEE Journal of Biomedical and Health Informatics*. 2021 Jul 30;26(2):798-808.

6. Nakagawa T, Ishida M, Naito J, Nagai A, Yamaguchi S, Onoda K, Alzheimer's Disease Neuroimaging Initiative. Prediction of conversion to Alzheimer's disease using deep survival analysis of MRI images. *Brain communications*. 2020;2(1):fcaa057.
7. Mirabnahrazam G, Ma D, Beaulac C, Lee S, Popuri K, Lee H, Cao J, Galvin JE, Wang L, Beg MF, Alzheimer's Disease Neuroimaging Initiative. Predicting time-to-conversion for dementia of Alzheimer's type using multi-modal deep survival analysis. *Neurobiology of Aging*. 2023 Jan 1;121:139-56.

Figure

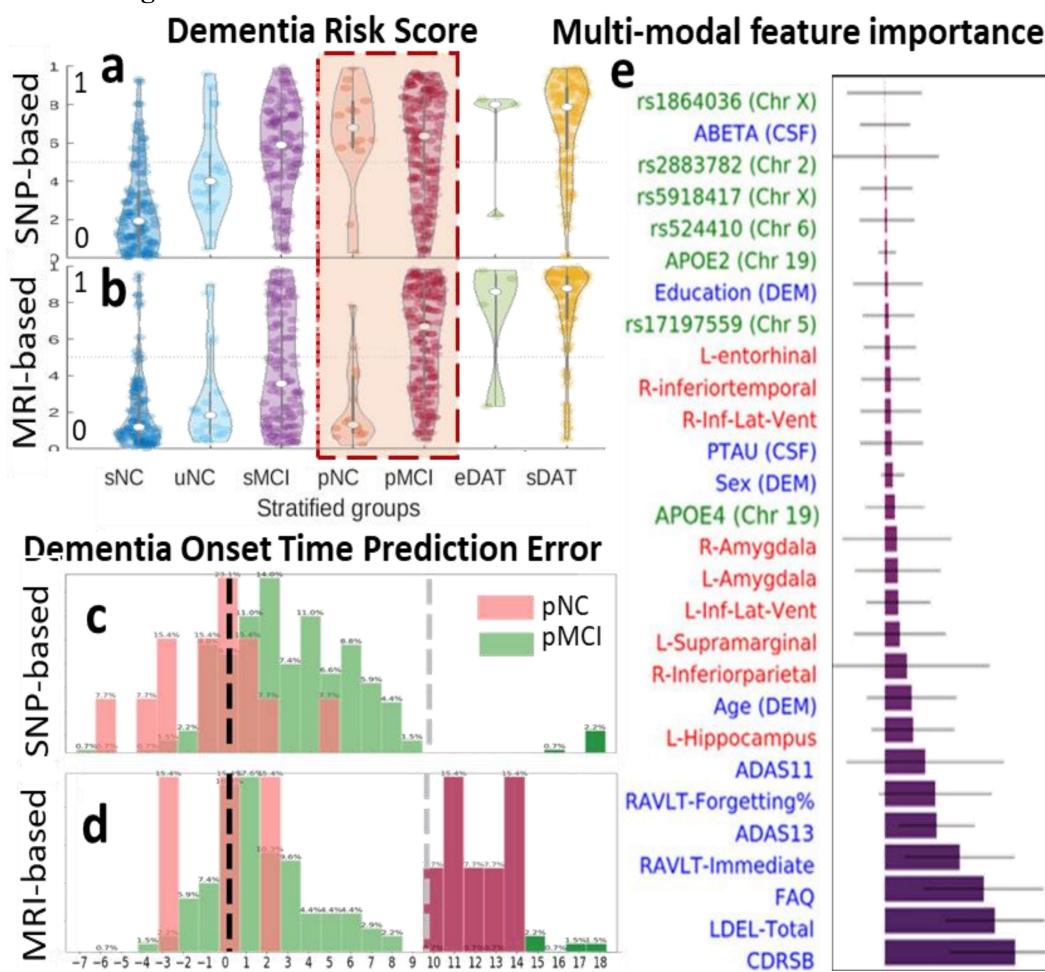


Figure: (a,b): Predicted AD risk from a) genomic; b) neuroimaging features. (c,d): Predictive error for time-to-dementia-onset (dark dashed line: error = 0 year). (e): Feature importance for AD prognosis. Red: MRI; Green: genomic; Blue: cognitive+CSF.

Green:

AI-Derived Body Composition and Post-operative Outcomes after Minimally-invasive Esophagectomy

Authors

V Christian Sanderfer MD¹, Ansley Beth Ricker MD¹ Alexis Holland MD¹, Erin Donahue PhD², Reilly Shea BS³, Matigan Simpson BA³, Katelyn Cockerham BS³, Nicholas Mullis BS³, Sophia Bellavia², Ella Schwarzen², M Hart Squires MD², Kunal C Kadakia MD², Jonathan C Salo MD²

¹ Department of Surgery, Carolinas Medical Center, Atrium Health, Charlotte, NC

² Department of Surgical Oncology, Levine Cancer Institute, Atrium Health, Charlotte

Abstract

Age and body composition were used to construct a predictive model for post-operative mortality after minimally-invasive esophagectomy. These factors efficiently predicted mortality risk.

Background and Objectives

Esophagectomy is a complex operation associated with significant morbidity for even fit patients. There is little prescriptive data regarding pre-operative patient selection. This study used CT-derived body composition measures and age to generate a risk model for the prediction of postoperative mortality after minimally-invasive esophagectomy.

Methods

Patients undergoing minimally-invasive esophagectomy from 2010-2022 were identified. Preoperative CT scans at the L3-vertebral level were analyzed using a U-Net convolutional neural network for automatic muscle and adipose tissue segmentation in order to calculate skeletal muscle gauge (SMG) as the product of skeletal muscle index and density. Patient demographics, clinical characteristics, and outcomes were collected. Patients were grouped based on SMG (low muscle = bottom quartile vs normal muscle = top 3 quartiles) and age (≥ 75 vs < 75 years). 90-day postoperative mortality was compared between risk groups.

Results

Of 399 patients, mean age was 62.8 years (SD 10.3). Anastomotic leak occurred in 8.3%, pneumonia in 17.0%, ventilation >48 hours in 21.1%, and discharge to home in 81.5% of patients. Thirty-day readmission was 17.3% and 90-day mortality was 5.5%. Ninety-day mortality by risk group showed 32% mortality among patients ≥ 75 years with low muscle, 9.6% for patients < 75 years with low muscle, 5.6% for patients ≥ 75 years with normal muscle and 1.8% for patients < 75 years with normal muscle.

Conclusions

Low-risk patients < 75 years with normal muscle had favorable outcomes with 1.8% 90-day postoperative mortality. Contrastingly, patients ≥ 75 years with low muscle measures were high risk with 32% mortality. Predictive models including SMG and age can identify high-risk groups and predict postoperative mortality after minimally-invasive esophagectomy.

Keywords

Body composition, Sarcopenia, pre-operative risk calculation

Acknowledgments

We thank Michelle L. Wallander PhD for writing support. Funding was received from the Levine Cancer Institute Advancement Fund through the generous gift of Duke Almon.

Towards Remote and Needleless Monitoring of NT-proBNP from ECG via AI

Authors

Ibrahim Karabayir¹, Melissa Hudson², Kristen Ness², Daniel Mulrooney², Stephanie Dixon², Turgay Celik¹, David McManus³, Robert L Davis⁴, Oguz Akbilgic¹

¹Wake Forest School of Medicine, ²StJude Children's Research Hospital, ³Umass Chan Medical School, ⁴UTHSC

Abstract

In this work, we aim to develop an electrocardiographic artificial intelligence (ECG-AI) model that can detect elevated N-terminal pro b-type natriuretic peptide (NT-proBNP) from ECG alone.

Background and Objectives

NT-proBNP is a protein elevated when a heart needs to work harder than normal to pump blood[1]. Therefore, NT-proBNP is a major biomarker used for heart failure (HF) diagnosis but also associated with HF readmission. Yet, NT-proBNP test is typically based on an expensive blood test, and not available for continuous monitoring of HF patients after discharge.

Methods

We used data from University of Tennessee Health Science Center (UTHSC), Memphis, TN to develop two deep learning models to detect clinical HF[2, 3]; 1) using 12 lead 10 seconds ECG and 2) using only lead I of 12 lead 10 seconds ECG. Next, we applied these two ECG-AI models, without any adjustment, to paired ECG and NT-proBNP data from St Jude Lifetime Cohort Study (SJLIFE) to detect elevated NT-proBNP. We presented the elevated NT-proBNP detection accuracies at different NT-proBNP levels using area under the receiver operating characteristics curve (AUC) statistics with 95% confidence interval.

Results

UTHSC cohort included 202,945 ECG from 77,769 patients. 43% of them were male, 61% African American, 36% white with average age \pm std of 53.8 ± 17.7 while 16,643 of them had HF. SJLIFE analytical cohort included 2,435 paired ECG and NT-proBNP from 1987 childhood cancer survivors including 50% male, 82% white, 15% African American with average age \pm std of 32.7 ± 9.5 . The ECG-AI models to detect HF was derived from UTHSC data using 75% for training, 15% for validation, and 10% for holdout test. A convolutional neural network based ECG-AI model yielded an AUC of 0.91 (0.90-0.91) on the holdout data when using all 12 leads of ECG and AUC of 0.89 (0.89-0.90) when using only lead I ECG to detect HF. When this 12 lead ECG-AI model was applied at SJLIFE data for a different purpose, to detect elevated NT proBNP, it provided AUCs ranging from 0.79 to 0.93 for different NT-proBNP thresholds from 200 to 500 pg/ML (Table 1). For the threshold of 500 pg/ML, the lead I model provided an AUC of 0.77 (0.76-0.78).

Conclusions

Using our ECG-AI model originally developed to detect clinical HF, we detected elevated NT-proBNP with very high accuracy using ECG-data alone at an external cohort. While the lead I results are promising, they are suboptimal to those obtained from 12-lead ECG. Considering that the original model was not specifically developed for NT-proBNP, and no patient level information was used except ECG, future studies focused on developing lead I ECG specific NT- proBNP detection models also using clinical data may improve accuracy. Such a tool can offer potential to facilitate early HF diagnosis and management with remote monitoring integrated with telemedicine.

Keywords

ECG, artificial intelligence, N-terminal pro b-type natriuretic peptide, heart failure

Table 1. ECG-AI based elevated NT-proBNP detection accuracies.

NT-proBNP Threshold	Number of participants with NTpro-BNP		AUC (95% CI)
	≤ Threshold	> Threshold	
200*	2329	106	0.79 (0.76-0.83)
300	2383	52	0.83 (0.79-0.89)
400	2399	36	0.87 (0.83-0.92)
500	2413	22	0.93 (0.90-0.97)

References:

- [1] J.Doust, R. Lehman, and P.Glasziou, "The role of BNP testing in heart failure," *Am Fam Physician*, vol. 74, no. 11, pp. 1893-8, Dec 1 2006. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pubmed/17168346>.
- [2] I. Karabayir *et al.*, "Single Lead Wearable ECG Simulation Augmented With AI For Heart Failure Identification," *Journal of the American College of Cardiology*, vol. 83, no. 13_Supplement, pp. 2393-2393, 2024/04/02 2024, doi: 10.1016/S0735-1097(24)04383-3.
- [3] I. Karabayir *et al.*, "Development and validation of an electrocardiographic artificial intelligence model for detection of peripartum cardiomyopathy," *American Journal of Obstetrics & Gynecology MFM*, vol. 6, no. 4, p. 101337, 2024/04/01/2024, doi: <https://doi.org/10.1016/j.ajogmf.2024.101337>.

CAIR AI in Healthcare Workshop: Can AI-powered clinical documentation improve clinicians' efficiency and satisfaction?

Authors

Tsai-Ling Liu, PhD¹, Timothy C. Hetherington, MS^{1,2}, Andrew McWilliams, MD, MPH^{2,3}, Ajay Dharod, MD, FACP⁴, Tracey Carroll, RHIT, MHA, MBA, OCSA, ITIL⁹, Richa Bundy, MPH⁴, Hieu Nguyen, MS¹, McKenzie Isreal, MPH¹, Jeffrey A. Cleveland, MD, FAAP¹⁰

¹Center for Health System Sciences, Atrium Health, Charlotte, NC

²Information Technology, Atrium Health, Charlotte, NC ³Department of Internal Medicine, Atrium Health, Charlotte, NC

⁴Informatics and Analytics, Department of Internal Medicine, Wake Forest University School of Medicine, Winston-Salem, NC

⁵General Internal Medicine, Department of Internal Medicine, Wake Forest University School of Medicine, Winston-Salem, NC

⁶Department of Implementation Science, Division of Public Health Sciences, Wake Forest University School of Medicine, Winston-Salem, NC

⁷Wake Forest Center for Healthcare Innovation, Wake Forest University School of Medicine, Winston- Salem, NC

⁸Wake Forest Center for Artificial Intelligence, Wake Forest University School of Medicine, Winston- Salem, NC

⁹Information Technology, Wake Forest University School of Medicine, Winston-Salem, NC

¹⁰Department of Pediatrics, Atrium Health, Charlotte, NC

Abstract

Background and Objectives

In 2019, Atrium Health created the Best Place to Care strategy, to eliminate points of frustration and interruptions to patient care, ultimately, to reduce burnout among clinicians. Administrative tasks like documentation outside work hours contribute to clinician burnout and dissatisfaction. Nuance's Dragon Ambient eXperience (DAX) Copilot is a comprehensive AI-driven ambient clinical documentation software platform designed to generate an initial clinical note by "listening" to the interaction between clinicians and patients which aims to improve electronic health record (EHR) efficiency and reduce burnout. The objective of this study is to assess the performance of participating clinicians across various practices while utilizing DAX Copilot.

Methods

In this non-randomized controlled trial, primary care clinicians were recruited between June and August 2023. A control group of clinicians not utilizing DAX was recruited from similar practices. Primary outcomes of EHR use and financial impact were assessed across 180-days using linear mixed models. Secondary analyses were performed to compare the control group with the user subgroups, who transferred ≥ 25% and ≥ 60% of DAX notes, respectively. A secondary outcome included a satisfaction survey, leveraging the American Medical Association (AMA) Organizational Biopsy electronic health record (EHR)-specific questions.²

Results

In the metrics analysis, high DAX users had 6.4% ($p=0.013$, 95% CI: 1.4%-11.2%) decreased hours of documentation compared to the control group after controlling for length of intervention, age, gender, provider type, years of practice, and baseline outcome. We did not find statistical significance in any other outcomes between the two groups. In the survey results, approximately 45% of DAX users report a positive experience in EHR experiences: Among intervention group respondents, 47% (40/85) reported decreased time on the EHR at home and 45% (38/85) reported decreased weekly time on the EHR outside normal work hours, compared to 15% (8/55) and 20% (11/55) among the control group ($p<0.001$ and $p=0.003$, respectively). Moreover, 44% (37/85) of intervention respondents reported decreased time on documentation after the visits and 45% (38/85) reported less frustration using the EHR, compared to 18% (10/55) and 15% (8/55) among the control group ($p=0.002$ and $p<0.001$, respectively). Conversely, an average of 45% participants in the intervention group and 68.7% in the control group reported their EHR experiences were comparable to pre- and post- intervention.

Conclusions

Our study identified a slight positive trend in EHR and financial metrics, however, except for time spent on documentation these differences did not achieve statistical significance compared to the control group. Additionally, we observed a positive impact of DAX Copilot on the experience of approximately half of the participants, suggesting the potential for this and similar technologies to alleviate burnout among certain users. AI-powered ambient clinical documentation software offers a promising solution to alleviate the documentation workload experienced by outpatient clinicians. Nonetheless, our findings indicate that, overall, the tool did not make clinicians as a group more efficient.

Keywords

DAX Copilot; Clinical documentation; EHR use; financial; clinician efficiency

Acknowledgments

The study is Funded by Wake Forest University Health Sciences (ClinicalTrials.gov number, NCT06329427). The authors thank Marc Kowalkowski, PhD, of Department of Internal Medicine, Wake Forest University School of Medicine for his assistance on conducting power analysis for this study.

Cayla Mansel from Operations Performance Improvement, Medical Group and Brittani M. Porter from Medical Group, Finance, helped pulling operational and financial metrics from Atrium Health. Todd M. Banks from Ambulatory Services Department helped pulling operational and financial metrics from Atrium Health Wake Forest Baptist Health. David Harold Russ from Human Resources assisted providing clinicians' demographics for analyses. Teammates mentioned above did not receive compensation beyond their usual salary for their contribution.

References:

1. Sinsky CA, Rule A, Cohen G, et al. Metrics for assessing physician activity using electronic health record log data. *J Am Med Inform Assoc*. Apr 1 2020;27(4):639-643. doi:10.1093/jamia/ocz223
2. Organizational Well-Being Assessment. American Medical Association. Accessed February 15, 2024. <https://www.ama-assn.org/system/files/org-well-being-assessment.pdf>

Session III Abstracts

CAIR AI in Healthcare Workshop: Machine Learning Algorithm for Reduction of Pre-Settling Times of Human Body Models

Authors

Seth Mischo¹, Wade von Kleeck¹, Katherine Mahon¹, F. Scott Gayzik¹

1. Center for Injury Biomechanics, Wake Forest School of Medicine

Abstract

A machine learning algorithm (MLA) created a predictive model using a Human Body Model (HBM).

Background and Objectives

HBMs are typically settled under the force of gravity for as long as one second to properly couple them to the seat. The use of an MLA trained on gravity settled models can reduce the settling time of HBMs.

Methods

A decision tree was trained on 100 simulations in a seating environment with a Simplified 50th Percentile HBM. The model consists of 299216 nodes representing all body structures and spatially oriented into baseline posture. Six features were analyzed. Parameter values not explicitly defined in the LHS were used to analyze the training data, and the nodal output was compared to an explicit pre-settling simulation.

Results

We found that the MLA output a predictive node set of the HBM with <5 mm of surface deviation. Computational time for analysis took approximately 75% less time than the standard modality.

Conclusions

We have successfully developed a first-of-its-kind MLA which can predict HBMs' nodal coordinates.

Keywords

Human Body Modeling, Crash Injury Biomechanics, Global Human Body Model Consortium

Acknowledgments

Simulations were run on the Wake Forest School of Medicine High Performance Cluster. Work for this project was sponsored by the Global Human Body Model Consortium.

References:

1. Gayzik, F.S.: Development of a Full Human Body Finite Element Model for Blunt Injury Prediction Utilizing a Multi-Modality Medical Imaging Protocol

Exploring the Host Response in Infected Lung Organoids Using NanoString Technology: A Machine Learning Analysis of Gene Expression Data

Mostafa Rezapour, PhD^{1*}; Muhammad Khalid Khan Niazi, PhD¹; Stephen J. Walker, PhD²; David A. Ornelles, PhD³; Patrick M. McNutt, PhD²; Anthony Atala, MD²; Metin Nafi Gurcan, PhD¹

¹Center for Artificial Intelligence Research, Wake Forest University School of Medicine, Winston-Salem, NC, USA

²Wake Forest Institute for Regenerative Medicine, Wake Forest University School of Medicine, Winston-Salem, NC, USA

³Department of Microbiology and Immunology, Wake Forest University School of Medicine, Winston-Salem, NC, USA

Abstract

Background

Traditional methods for studying viral infections in human airways are limited by their inability to accurately mimic the complex 3D environment of the respiratory tract. In response, we utilized a three-dimensional airway “organ tissue equivalent” (OTE) model at an air-liquid interface (ALI) to closely replicate human airway dynamics.

Objective

The primary goal of this study was to assess gene expression changes in OTE models infected with Influenza A virus (IAV), Human metapneumovirus (MPV), and Human Parainfluenza virus type 3 (PIV3) using the NanoString platform. We aimed to identify genes capable of differentiating between live infection conditions and their respective post-infection times at 24- and 72-hour intervals, focusing on a panel of 773 specific genes.

Methods

To refine our analysis and ensure the biological relevance of our findings, we introduced two novel algorithms: the Magnitude-Altitude Score (MAS) and Optimized Gene Selection based on MAS Score (OGS-MAS). The MAS algorithm integrates biological significance, as indicated by fold changes in gene expression, with statistical rigor through adjusted p-values. The OGS-MAS algorithm is designed to solve classification challenges by identifying distinctive genes that exhibit time-related variations in expression, effectively distinguishing between different viral treatments and control conditions.

Results

Our analysis unveiled distinct gene expression patterns in response to each virus, demonstrating the utility of the MAS algorithm in identifying differentially expressed genes that are not mere statistical artifacts. Using the OGS-MAS method, we successfully differentiated between infected and non-infected samples with high accuracy. This method identified three distinctive genes for each virus, distinguishing between all treated conditions and the baseline across both assessed time points.

Conclusion

The application of our novel analytical frameworks on the NanoString platform has provided a more holistic understanding of the airway tissue response to viral infections. By combining detailed gene expression analysis with innovative algorithmic approaches, we have enhanced the capability to discern subtle yet significant changes in the host response to different viral challenges.

CAIR AI in Healthcare Workshop: Deep Learning for AFib and Sleep Apnea Detection Using Multimodal Sleep Study Data

Authors

Shekhar Singh, Liam Butler, Ibrahim Karabayir, Oguz Akbilgic

Wake Forest University School of Medicine

Abstract

Background

Sleep apnea has been associated with atrial fibrillation. Patients with sleep apnea are typically recommended to use continuous positive airway pressure (CPAP) machine partially due to reduce cardiovascular disease risk. However, there has been conflicting evidence whether CPAP is associated with reduced cardiovascular disease risk or not. This research aims to study the association of atrial fibrillation and sleep apnea by utilizing very rick continuous physiologic data collected during sleep studies.

Goals

- 1) Predicting risk for atrial fibrillation using short (30 seconds) and long term continues (up to 9 hours) sleep ECG to predict 5-year risk for atrial fibrillation.
- 2) Assess whether sleep apnea status of people lead to improved performance or not.
- 3) Detect sleep apnea from first 30 seconds of the recording to assess feasibility of sleep apnea detection from wearable without a need for sleep study.

Data and Methods

This study leverages the Sleep Heart Health Study (SHHS) dataset, encompassing data from 5793 participants. Our research utilizes advanced AI techniques, specifically deep convolutional neural networks (CNNs), to analyze a comprehensive set of physiological features from the SHHS dataset. The data is divided into training, validation, and testing subsets to ensure robust model performance and generalizability. The CNN model architecture is designed to manage the complexities of multi-dimensional physiological signals, extracting key features indicative of AFib.

Results

We could successfully download very large SHHS data, each sleep study including 6 to 9 hours of sleep data of fourteen waveforms sampled a 125Hz in EDF format. The waveforms included SaO₂, H.R., EEG (sec), ECG, EMG, EOG(L), EOG(R), EEG, THOR RES, ABDO RES, POSITION, LIGHT, AIRFLOW, OX stat. Two separate clinical datasets have been linked to these waveforms to drive patient characteristics and clinical outcomes such as incident atrial fibrillation and sleep apnea status. We are thoroughly examining the SHHS dataset to gain insights into the characteristics of AFib and sleep apnea patterns. The initial steps include data preprocessing, feature extraction, and exploratory data analysis to identify relevant patterns and trends in the physiological data.

Next Steps and Conclusions

To run CNN models for the three outcomes mentioned above.

Keywords

AFib classification, sleep apnea detection, deep learning, convolutional neural networks, SHHS dataset, medical diagnostics, AI in healthcare

Acknowledgements

This study is funded by NHLBI (R21HL167126, Younghan, Akbilgic, Azarbarzin).

Note: I would like to be considered for the award.

CAIR AI in Healthcare Workshop: [Enhancing Clinical Documentation Efficiency through AI for Hospitalist at Home Program]

Authors

Usman Afzaal¹, Padageshwar Rao Sunkara², Sneha Chebrolu², Sumera Andleeb², Suneel Parvathareddy², Metin Nafi Gurcan¹

¹Center for Artificial Intelligence Research, Wake Forest University School of Medicine, Winston-Salem, North Carolina

²Department of Internal Medicine, Wake Forest School of Medicine, Winston-Salem, North Carolina
Email: uafzaal@wakehealth.edu

Abstract

Clinical documentation is crucial for various purposes, including billing substantiation, information relay, and understanding clinical decision-making. The advent of electronic health records (EHRs) has facilitated the capture and interpretation of extensive data, albeit with challenges in accessibility and usability. Additionally, the burden of documentation has escalated, with more time spent on chart computer/medical records review and documentation, detracting from bedside patient care while contributing to provider burnout [1]. Various interventions, including AI, have been explored to alleviate this burden. The adoption of AI offers promise in streamlining this process, but its accuracy, potential biases, and safeguards need rigorous evaluation.

Background and Objectives

This study aims to assess the efficacy of AI in generating meaningful and actionable summaries for patients under the care of the hospitalist at-home team following discharge from hospital, using the information documented during their hospital stay documented in form of a discharge summary.

Methods

Patient EHR data is used to generate summaries using different AI algorithms. Multiple open Large-language models (LLMs) will be evaluated on their ability to generate accurate summaries based on the EHRs, whereas the LLMs will be run entirely locally. The choice of LLMs include state-of-the-art models, such as Llama-3 [2], Mixtral [3], Meditron [4] and Med-Alpaca [5].

The experimental data includes anonymized EHR from 100 patients. These records are summarized by an expert clinician to obtain the ground-truth summary. The records are then fed to each of the selected LLM to get the respective AI-generated summaries, which are then evaluated against the ground-truth summaries. The generated summaries are evaluated for completeness, correctness, and conciseness, using standard NLP metrics including BLEU, ROUGE-L, BERTScore and MEDCON [1]. Additionally, Llama-3 is employed to offer a semantic-based comparison between the predictions and the actual summaries.

Results

Initial results for 15 predicted patient summaries using Llama-3 are present in Table 1.

Table 1 Model performance for different metrics.

Model\Score	BLEU	ROUGE-L
Llama-3	0.2589	0.1883

Conclusions

This study seeks to contribute to the ongoing discourse on optimizing clinical documentation efficiency through AI. By evaluating the effectiveness of AI-generated summaries, we aim to provide insights into enhancing provider workflows and decreasing burnout related to clinical documentation.

Keywords

Clinical documentation; Electronic Health Records; Large Language Models; Patient care; Hospitalist at Home

References

1. Van Veen, D., et al., Clinical text summarization: adapting large language models can outperform human experts; 2023. Preprint at`(` <https://doi.org/10.48550/arXiv.2309>.
2. AI@Meta. Llama 3 Model Card. 2024; Available from: https://github.com/meta-llama/llama3/blob/main/MODEL_CARD.md.
3. Jiang, A.Q., et al., Mixtral of experts. arXiv preprint arXiv:2401.04088, 2024.
4. Bosselut, A., et al., MEDITRON: Open Medical aFoundation Models Adapted for Clinical Practice. 2024.
5. Han, T., et al., MedAlpaca--an open-source collection of medical conversational AI models and training data. arXiv preprint arXiv:2304.08247, 2023.

Poster Session

CAIR AI in Healthcare Workshop:

Remote Maternal Health Monitoring: 9+1AI Co-Pilot App

Authors

Luke Patterson, MS¹, Ibrahim Karabayir, PhD¹, Robert L Davis, MD, MPH², Oguz Akbilgic, DBA, PhD¹

1. Wake Forest School of Medicine, Winston-Salem, NC

2. University of Tennessee Health Science Center, Memphis, TN

Abstract

Background and Objectives

Effective management and treatment of adverse pregnancy outcomes involves timely diagnosis and early risk assessment. 9+1AI Co-Pilot is an innovative mobile and web application aiming to be a device-agnostic toolkit leveraging artificial intelligence to remotely monitor maternal health and prevent adverse pregnancy outcomes by identifying user specific risks. The application accurately detects and predicts risk for preeclampsia, peripartum cardiomyopathy, elevated brain natriuretic peptide (BNP), and estimates blood pressure from a single-lead electrocardiogram (ECG).

Methods

We developed and validated electrocardiographic artificial intelligence (ECG-AI) models using 6 million ECGs from the University of Tennessee Health Science Center and Wake Forest School of Medicine. The ECG-AI models, designed as modified ResNet for 1D signal data, use only Lead I of standard 12-lead ECG as input. The 9+1AI Co-Pilot was developed using the cross-platform framework Flutter to execute the ECG-AI models on the user's device or in the cloud, depending on data privacy preferences.

Results

The ECG-AI models achieved high external validation accuracies, with Area Under the Receiver Operating Characteristics Curve (AUC) of 0.85 for preeclampsia, 0.92 for peripartum cardiomyopathy, and 0.89 for elevated BNP (>500 pg/ml). Systolic and diastolic blood pressure estimation had mean absolute errors < 9 mmHG and < 7 mmHG, respectively. 9+1AI Co-Pilot, tested on iOS and Android smartphones, successfully retrieved raw digital ECGs from Apple Watch, Withings ScanWatch, and Fitbit Charge 6. The application executes all four ECG-AI models within 1 second on the user's phone or via cloud computing.

Conclusions

9+1AI Co-Pilot is a flexible platform that accurately predicts preeclampsia, peripartum cardiomyopathy, elevated BNP risk, and estimates blood pressure from a single-lead ECG. The application's compatibility with popular smart watch devices and its ability to operate offline or using cloud computing make it accessible for remote maternal health monitoring, even in low-resource settings. Once prospectively validated in real-world remote monitoring scenario, it can help reducing maternal health mortality and morbidity by assisting pregnant women on how to interpret their symptoms.

Keywords

Maternal health, artificial intelligence, mobile application, electrocardiogram, remote monitoring, preeclampsia, peripartum cardiomyopathy, brain natriuretic peptide

Acknowledgments

This research is funded by Wake Forest Innovations. Akbilgic and Davis are co-founders of 9+1AI LLC.

Development of a multi-modal artificial intelligence model to detect degrees of hemorrhage

Authors

Fahim Ussihab Mobin¹, Antonio C. Renaldo¹, Micaela K. Gomez¹, Sandra Januszko¹, Jacob Dooley¹, James E. Jordan², Oguz Akbilgic³, Timothy K. Williams⁴, and Elaheh Rahbar^{1,5}

1. Department of Biomedical Engineering, Wake Forest University School of Medicine **2.** Department of Cardiothoracic Surgery **3.** Center for Artificial Intelligence Research **4.** Department of Vascular and Endovascular Surgery **5.** Department of Biomedical Engineering, Texas A&M University

Abstract

Early and accurate detection of hemorrhage level is paramount for effective intervention to optimize patient outcomes. This study presents a novel approach where two distinct deep learning models, namely multilayer perceptron (MLP) and multi-head 1D CNN (MHC), were developed to process hemodynamic data obtained from swine subjected to controlled hemorrhage.

Background and Objectives

Hemorrhagic shock is the leading cause of potentially preventable deaths in people aged 1-44 years. Early detection of hemorrhage is needed to ensure favorable patient outcomes. In this study, we leveraged multi-modal deep learning to process hemodynamic data for the detection and classification of blood volume loss from an established porcine model of hemorrhagic shock.

Methods

Yorkshire swine subjected to 10%, 20%, and 30% (n=18/group) total blood volume hemorrhage over 30 minutes. Pressure waveforms and flow hemodynamics were recorded in real-time. The MLP model was trained on mined cardiac P-V metrics while the MHC model was trained on blood pressure waveforms.

Results

In the MLP model, hemorrhage level detection accuracy increased significantly with the incorporation of features like stroke work and arterial elastance mined from the PV. In the MHC model, a similar trend was noticeable with left subclavian and distal arterial pressures yielding >80% model accuracy.

Conclusions

This study leverages CNNs for time-series analysis, presents an innovative approach that may prove to be effective and require less tuning compared to traditional waveform analyses in the context of hemorrhage.

Keywords

controlled hemorrhage, multi-head 1D CNN, multi-modal AI, hemodynamics, pressure waveforms

Development of a batch process to mine and analyze carotid duplex ultrasound

Authors

Fahim Usshab Mobin¹, Micaela K. Gomez¹, Timothy K. Williams², Aarti Sarwal³, Oguz Akbilgic⁴, and Elaheh Rahbar^{1,5}

1. Department of Biomedical Engineering, Wake Forest University School of Medicine **2.** Department of Vascular and Endovascular Surgery **3.** Department of Neurology **4.** Center for Artificial Intelligence Research **5.** Department of Biomedical Engineering, Texas A&M University

Abstract

Background and Objectives

Carotid arterial stenosis (CAS) is a common etiology for ischemic stroke, with > 50% of stroke patients having varying degrees of ipsilateral carotid stenosis. Refined methods to improve risk stratification in patients with asymptomatic CAS are needed. Carotid duplex ultrasound is used clinically to classify CAS patients using pulse wave Doppler velocity (PWV), but there is not yet a quantifiable way to leverage them. We postulate deep learning analysis will improve stroke risk predictions.

Methods

Carotid duplex ultrasounds were extracted (N=3500) and deidentified utilizing a custom web scraping bot to extract DICOM images from Intellispace Cardiovascular at WFBMC. A natural language processing (NLP) algorithm was developed for interpreting physician reports to provide data labels. Image segmentation utilized a Faster R-CNN algorithm, and a custom algorithm digitized PWV into 3D.

Results

DICOM images were effectively de-identified and the NLP algorithm has insofar interpreted physician reports with high fidelity. Image segmentation using Faster R-CNN was performed, resulting with PWV graphs isolated and digitized in 3D. Clinically correlation analysis for stroke prediction is pending.

Conclusions

The project presents a multi-faceted approach to analyze CAS data, integrating AI algorithms and data extraction techniques. Successful completion of these initial stages demonstrates study feasibility. Further analyses will reveal clinical correlations for the data to clinical events and other stroke prediction models.

Keywords

carotid arterial stenosis, DICOM, pulse wave velocity, deep learning, natural language processing, faster R-CNN, 3-D CNN

Enhancing Colorectal Cancer Tumor Bud Detection Using Deep Learning from Routine H&E-Stained Slides

Authors

Usama Sajjad^a, Ziyu Su^a, Mostafa Rezapour^a, Thomas Tavolara², Wendy L. Frankel³, Metin N. Gurcan¹, Wei Chen³, M. Khalid Khan Niazi¹

1. Center for Artificial Intelligence Research, Wake Forest University School of Medicine, Winston-Salem, NC, USA;
2. Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA;
3. Department of Pathology, The Ohio State University, Columbus, OH, USA;

Abstract

Colorectal cancer (CRC) is the third most widespread cancer, with an estimated 153,000 new cases projected for 2024 [1]. Tumor buds, small clusters of tumor cells (one to four), are a valuable prognostic factor but are challenging to count and grade. We propose a weakly-supervised deep learning method and Bayesian Multiple Instance Learning (BMIL) to improve tumor bud detection from H&E-stained images.

Background and Objectives

Our study addresses the challenge of detecting tumor buds in colorectal cancer (CRC), which are valuable prognostic factors [2, 3] but difficult to evaluate due to time-consuming and intra reader variability among pathologists. The main objective is to develop a weakly-supervised deep learning method for detecting tumor buds from routine H&E-stained images, enhancing the process's accuracy and efficiency. Additionally, we propose using Bayesian Multiple Instance Learning (BMIL) to improve the generalizability and stability of tumor bud detection algorithms.

Methods

Multiple Instance Learning (MIL) [4] operates on groups (bags) of instances instead of individual instances. Each 512x512 pixel region around pathologist tumor bud annotations forms a bag, divided into patches to construct instances. MIL computes the attention scores of each instance and an aggregation network generates a bag representation for classifying pathologist annotations as tumor or non-tumor bud ROIs. Bayesian Multiple Instance Learning (BMIL) improves MIL classification by merging K-annotated regions from the same slide, each consisting of n patches and sharing the same label.

Results

Table 1: Results of tumor bud detection methods on test set.

Method		AUC	Precision	Recall
1	Attention MIL	0.95 ± 0.04	0.84 ± 0.11	0.90 ± 0.06
2	BMIL (K = 2)	0.96 ± 0.01	0.83 ± 0.06	0.91 ± 0.04
3	BMIL (K = 3)	0.95 ± 0.02	0.80 ± 0.10	0.89 ± 0.04

Conclusions

We presented BMIL, a weakly supervised method achieved high performance with an AUC of 0.96 ± 0.01 , precision of 0.83 ± 0.06 , and recall of 0.91 ± 0.04 across 70 slides, promising improved diagnostic capabilities for CRC in H&E images.

Acknowledgements

Partial funding for the research was provided by NIH grants R01CA276301, R21EB029493, R21CA273665,

R01DC020715, and Alliance Clinical Trials in Oncology GR125886. The authors bear full responsibility for the content, which may not represent the official views of NIH institutes.

References

1. Siegel, R.L.; Giaquinto, A.N.; Jemal, A. Cancer statistics, 2024. *CA: A Cancer Journal for Clinicians* **2024**, *74*, 12-49, doi:<https://doi.org/10.3322/caac.21820>.
2. Chen, K.; Collins, G.; Wang, H.; Toh, J.W.T. Pathological Features and Prognostication in Colorectal Cancer. *Current oncology (Toronto, Ont.)* **2021**.
3. Zlobec, I.; Berger, M.D.; Lugli, A. Tumour budding and its clinical implications in gastrointestinal cancers. *British journal of cancer* **2020**.
4. Ilse, M.; Tomczak, J.; Welling, M. Attention-based deep multiple instance learning. In Proceedings of the International conference on machine learning, 2018; pp. 2127-2136.

CAIR AI in Healthcare Workshop: Revolutionizing Pathway Analysis with Deep Learning: The Gene PointNet Method

Authors

[List of Authors] *Hao Lu¹**, *Mostafa Rezapour¹*, *Haseebullah Baha²*, *Muhammad Khalid Khan Niazi¹*, *Aarthi Narayanan³*, *Metin Nafi Gurcan¹*.

1.Center for Artificial Intelligence Research, Wake Forest University School of Medicine, Winston-Salem, NC, USA. 2.School of Systems Biology, College of Science, George Mason University, Fairfax, VA 22030, USA. 3.Department of Biology, George Mason University, Fairfax, VA, USA.

Abstract

Background and Objectives

Pathway analysis is a cornerstone in bioinformatics for elucidating gene interactions within biological processes and diseases. Traditional methods, such as Gene Set Enrichment Analysis (GSEA)[1] and Over-representation Analysis (ORA)[2], often overlook complex gene interactions within pathways. This study introduces Gene PointNet (GPNet), a deep learning-based method that models intricate gene relationships and improves pathway analysis accuracy.

Methods

GPNet employs neural networks to capture detailed biomarker-pathway relationships. A novel p-value computation method within GPNet bypasses the need for permutation tests, enhancing computational efficiency. The effectiveness of GPNet was validated through comparative studies using both a simulated dataset and The Cancer Genome Atlas (TCGA) breast cancer RNA-Seq data.

Results

Our results demonstrate that GPNet outperforms traditional methods in detecting small expression level changes in pathways, especially with large sample sizes. Key contributions include the development of the GPNet model, an innovative p-value computation method based on the confusion matrix, and a comprehensive comparative study, shown as Figure 1.

GPNet's superior performance indicates its potential as a powerful tool for pathway analysis, capable of overcoming limitations of existing methods and providing more precise insights into gene-pathway interactions.

Conclusions

GPNet represents a significant advancement in pathway analysis, offering enhanced computational efficiency and accuracy. This method holds promise for widespread application in bioinformatics, enabling more nuanced understanding of gene interactions within biological pathways.

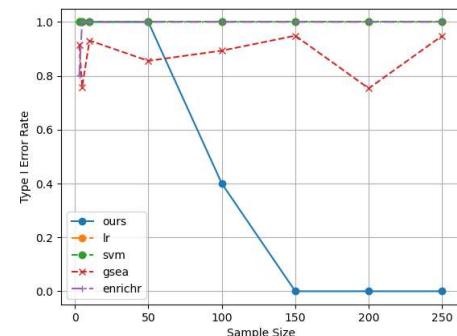


Figure 1. Comparison of Type I error Rates across different sample size. SNR=0.01, Type II error < 0.1

Keywords

Pathway analysis, Gene PointNet, deep learning, bioinformatics, gene interactions, p-value computation.

Acknowledgments

This research was supported by the U.S. Government under HDTRA 12310003 “Host signaling mechanisms contributing to endothelial damage in hemorrhagic fever virus infection.” The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements of the U.S. Government.

References:

1. J. Shi and M. G. Walker, "Gene set enrichment analysis (GSEA) for interpreting gene expression profiles," *Current Bioinformatics*, vol. 2, no. 2, pp. 133-137, 2007.
2. M. V. Kuleshov et al., "Enrichr: a comprehensive gene set enrichment analysis web server 2016 update," *Nucleic acids research*, vol. 44, no. W1, pp. W90-W97, 2016.

Automatic Tympanic Membrane Segmentation from Otoscopic Videos Using the SAM-Adapter

Author

Seda Camalan, PhD

Abstract

Hearing loss is a critical global health issue with significant impacts on healthcare costs, education, and productivity. Middle ear infections, particularly otitis media (OM), are major contributors to preventable childhood hearing loss, affecting speech and language development. Traditional otoscopic methods for diagnosing middle ear diseases, such as OM, often suffer from subjectivity and inconsistency, leading to a demand for computer-aided diagnosis (CAD) systems. Accurate segmentation of the tympanic membrane (TM) is essential for effective diagnosis and early intervention in these conditions.

This study introduces a novel approach utilizing the Segment Anything Model Adapter (SAM-Adapter) for the automatic segmentation of TMs from video-otoscopic frames. The SAM-Adapter, leveraging a Vision Transformer (ViT)-based architecture, demonstrates superior performance compared to traditional convolutional neural network (CNN) models, such as U-Net. Using a dataset of 765 video frames from 36 otoscopic videos, the SAM-Adapter model achieved a Dice similarity coefficient of 0.95 and pixel accuracy of 0.99, outperforming traditional U-Net-based models both qualitatively and quantitatively.

The study's dataset comprises frames extracted from high-definition video otoscopes and manually annotated by otolaryngologists to create ground truth masks for training the SAM-Adapter model. Data augmentation techniques were applied to enhance the training process, resulting in a comprehensive dataset split into training, validation, and testing sets. The SAM-Adapter model incorporates domain-specific prompts to improve segmentation performance without fine-tuning the pre-trained model weights.

The model effectively segments the TM, even in frames with blurriness and glare, which are common challenges in otoscopic images. The SAM-Adapter's architecture, which includes adapters for task-specific prompts, enhances its ability to handle the unique characteristics of TM segmentation in otoscopic images. This study provides compelling evidence of the SAM-Adapter's applicability and effectiveness in medical image segmentation, marking a significant advancement in CAD systems for middle ear disease diagnosis. The findings underscore the potential of ViT-based models in medical image analysis, contributing to improved diagnostic accuracy and early intervention strategies.

Collective-Intelligence Recommender Systems: Advancing Health Communication into the 21st Century

Authors

Thomas Houston, MD MPH ^a; Rajani Sadasivam, PhD; ^b Peter Lindenauer, MD MSc ^c; Daniel Ford, MD MPH ^d

^a Wake Forest University School of Medicine, Department of Internal Medicine; ^b UMass Chan School of Medicine, Worcester, MA; ^c UMass Chan Medical School - Baystate Regional Campus, Springfield, MA. USA; ^d Johns Hopkins University, Baltimore, MD

Abstract

Background and Objectives

What is the next frontier for computer-tailored health communication (CTHC) research? In current CTHC systems, designers select the variables and develop if-then rules that specify how content should be tailored. In collective-intelligence recommender systems (hereafter recommender systems) used by Web 2.0 companies (e.g.: Netflix and Amazon), machine learning algorithms combine user profiles and content ratings for tailoring. In a series of experiments over the past 10 years, we have assessed the feasibility of using recommender systems in health communication and the comparative effectiveness of a 21st-century recommender systems versus standard if-then logic strategies to tailor messages.

Methods

In our portfolio of work in health promotion (e.g.: smoking cessation), we first demonstrated the feasibility of collecting ratings from users (i.e.: current smokers). Users ($n = 1,000$) provided demographic information and rated a corpus of 261 motivational text messages including messages written by experts and other messages written by former smokers. Then, in two randomized trials, we compared delivery of the message content using standard if-then logic based on motivational readiness to select messages versus message selection driven by the Patient Experience Recommender System for Persuasive Communication Tailoring (PERSPeCT). PERSPeCT used a novel hybrid machine learning algorithm incorporating a content-based approach (machine learning matches users to items based on the provided user profile data and item metadata) and collaborative filtering (uses past and current user's item ratings). The more ratings a user provides, the better the algorithm can select a next message.

Results

In Trial 1, ($n = 120$ current smokers randomized), the proportion of days when smokers agreed/strongly agreed (daily rating ≥ 4) that the messages influenced them to quit was significantly higher for PERSPeCT (73%, 23/30) than standard CTHC (44%, 14/30, $P=.02$). Trial 2 is ongoing. In Trial 2, six-month smoking cessation outcomes are complete for 568 participants. Overall, those randomized to the recommender intervention versus standard CTHC comparison had similar cessation rates (19% versus 18%). Interestingly, among African-American participants, those randomized to intervention had higher cessation rates versus the comparison (22.5% vs. 13.9%).

Conclusions

Compared to standard CTHC with proven effectiveness, the machine learning recommender system outperformed in terms of influence ratings and resulted in similar cessation rates. We will discuss the pathways by which the recommender system may increase effectiveness in diverse populations.

Keywords

Collective Intelligence; Recommender System; Collaborative Filtering; Health Promotion; Health Communication; Cancer Prevention

Acknowledgments

Funding for this work was from the PCORI (PI12-001), NSF (CAREER award (1350522)), and NCI

(R01CA129091, K07CA172677, R01CA240551).

Artificial Intelligence in Medical Physics and the AAPM

Author

J. Daniel Bourland, PhD

Professor and Physicist

Department of Radiation Oncology, Biomedical Engineering and Physics Wake Forest University School of Medicine

bourland@wakehealth.edu

Abstract

Approximately 700 AI-based technologies, classified as medical devices, have been approved by the US FDA. In the last 2-3 years a majority of these AI-based devices, approximately 75%, have been approved for applications in Radiology, including Radiation Oncology. This presentation describes the uses of artificial intelligence in radiology and radiation oncology with a summary of FDA approvals, and the initiative of the American Association of Physicists in Medicine for equipping medical physicists, who serve with physician colleagues in these fields, with the principles and knowledge for the safe and ethical implementation of artificial intelligence-based technologies in the clinical environments of radiology and radiation oncology. AI continues a rapid growth that will impact the healthcare community, and all fields of medicine must be preparing for the appropriate and safe implementation of AI-based technologies in the clinical and research environments.

Multimodal Radio-Pathomics and Interpretable Machine Learning for Predicting Immunotherapy Response in Gastric Cancer

Authors

Yuming Jiang¹, James Edward Han¹, Qing Lyu², Fan Yang³

¹Department of Radiation Oncology and ²Radiology, Wake Forest University School of Medicine; ³Department of Computer Science, Wake Forest University.

Background

Immunotherapy has advanced from first-line treatment to neoadjuvant therapy in gastric cancer (GC) due to its efficacy. However, it remains challenging to accurately predict the response to immunotherapy. This study aims to develop an interpretable radio-pathomics signature (RPS) by integrating features from CT and digital hematoxylin and eosin (H&E) - stained images to predict the immunotherapy response, supported by a biological rationale.

Methods

Data from 298 patients with both CT and digital H&E images were retrospectively collected from three medical centers between January 2017 and March 2023. To capture the compact and informative features from the whole slide H&E images, we extracted three kinds of pathomics features, including tumor nucleus features, single-cell spatial distribution features, and deep-learning features of the tumor microenvironment (TME). For the CT image analysis, we extracted radiomics features from the peritumor and intratumor area, respectively, and extracted deep-learning features using the ResNet50. Then, the integrated model was developed using seven machine learning methods, including Logistic Regression, Support Vector Machine, Decision Tree, K-Nearest Neighbors, Naive Bayes, Random Forest, and eXtreme gradient boosting (XGBoost). Its predictive performance was assessed using the area under the receiver-operating-characteristic curve (AUC). Interpretability was evaluated using the SHapley Additive exPlanations (SHAP) method, and biological explanations were provided through image-genomics analyses of 30 patients undergoing RNA sequencing.

Results

The overall response rate was 24.5% (N = 73). The AUCs of the RPS model to predict immunotherapy response were 0.844 (95%CI, 0.745-0.917), and 0.813 (95%CI, 0.653-0.920) in the internal, and external validation cohorts, respectively, surpassing the clinical biomarkers of CPS score, MSI-H, EBV, and HER-2. Multivariate Cox analysis revealed the RPS model was the independent prognostic factor of progression-free survival and overall survival (Hazard ratio, ranges from 0.051-0.297, P < 0.05). Genetic analyses indicated that the RPS model correlates with enhanced immune regulation pathways and increased infiltration of naive B cells and M1 macrophages.

Conclusion

The interpretable RPS predicts immunotherapy response and survival outcomes, which can be implemented to create more precise patient-specific treatment strategies. Additionally, the RPS model appears to be linked with innate immune pathways.

Keywords

Cancer immunotherapy response, multimodal integration, interpretation, prediction.

Deep-ODX: An efficient deep learning tool to risk stratify breast cancer patients from histopathology images

Authors

Ziyu Su¹, Amanda Rosen², Robert Wesolowski³, Gary Tozbikian⁴, M. Khalid Khan Niazi¹, Metin N. Gurcan¹

1. Center for Artificial Intelligence Research, Wake Forest University School of Medicine. 2. Department of Biomedical Engineering, Duke University. 3. Division of Medical Oncology, Comprehensive Cancer Center, The Ohio State University College of Medicine Department of Pathology, The Ohio State University.

Abstract

Our novel deep learning approach, achieving a 0.862 AUC, predicts ODX risk from routine H&E histopathology images, potentially offering a cost-effective and accessible prognostic solution.

Background and Objectives

Breast cancer is the most common cancer diagnosed in women in the United States. In early-stage, HR+, HER2- invasive breast cancer, the Oncotype DX (ODX) Breast Cancer Recurrence Score Test predicts the risk of recurrence and the benefit of chemotherapy. However, this gene assay is costly and time-consuming, making it inaccessible to many patients. This study proposes a novel deep-learning approach, Deep-ODX, which performs ODX recurrence risk prediction based on routine H&E histopathology images.

Methods

Our dataset included H&E-stained digital WSIs from 151 anonymized HR+, HER2- breast cancer patients. We classified the patients into two classes: low-risk (ODX score ≤ 25) and high-risk (ODX score > 25). We cropped the tumor bulk region of each WSI into small patches at $40 \times$ magnification. Then, we encoded all the patches into embeddings that were inputs to our MIL model using CTransPath model¹. To aggregate all patch embeddings to a slide-level decision, we utilized a novelly designed MIL model—DeepODX. In short, given an input slide, we compared each of its embedding with the low/high-risk key embeddings to learn saliency-informed attention scores. Finally, we performed weighted summation on the input embeddings based on the attention scores. A one-layer neural network is used to map the aggregated embedding into slide-level prediction.

Results

We achieved an AUROC (area under the ROC curve) of 0.862 ± 0.034 and an F1-score of 0.761 ± 0.086 .

Conclusions

In conclusion, we present an accurate ODX recurrence risk prediction model based on WSIs. This technology holds the potential of offering a cost-effective alternative to the expensive ODX test, thereby enhancing the accessibility of high-quality prognosis tools for breast cancer patients.

Keywords

Deep Learning, Breast Cancer, Computational Pathology, Oncotype-DX

Acknowledgments

The research was partially funded by a National Institutes of Health R01CA276301 (PIs: Niazi, Chen), Trailblazer award R21EB029493 (PIs: Niazi, Segal), R21CA273665 (PI: Gurcan), R01DC020715 (PIs: Gurcan and Moberly), and Alliance Clinical Trials in Oncology GR125886 (PIs: Frankel and Niazi). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References:

1. Wang, X. *et al.* Transformer-based unsupervised contrastive learning for histopathological image classification. *Medical Image Analysis* **81**, 102559 (2022).