362

Evaluating the educational quality of ChatGPT as a health information resource for patients with acute myeloid leukemia (AML)

Mihir Patel[1], Fadzai Chinyengetere[1], Sanghee Hong[2], Chenyu Lin[2], Michele Sainvil[1], Allison O. Taylor[1], Brenda Branchaud[1], Kris W. Herring[1] and Thomas W. LeBlanc[2]

[1]Duke University School of Medicine, Durham, NC and [2]Division of Hematologic Malignancies and Cellular Therapy, Department of Medicine, Duke University Medical Center, Durham, NC

OBJECTIVES/GOALS: Upon diagnosis, patients with acute myeloid leukemia (AML) have significant information needs. Given its recent increase in popularity, patients may use ChatGPT to access information about AML. We will examine the quality, reliability, and readability of information that ChatGPT provides in response to frequently asked questions (FAQs) about AML. METHODS/STUDY POPULATION: From FAQs on the top 3 patient-facing websites about AML, we derived 26 questions, written in lay terms, about AML diagnosis, treatment, prognosis, and functional impact. We queried ChatGPT-4o on 10/14/2024 using a new Google account with no prior history. We asked each question in a separate chat window once, verbatim, and without prompt engineering. After calibration, 5 oncologists independently reviewed ChatGPT responses. We assessed quality via the Global Quality Scale (GQS), scored from 1 (poor) to 5 (excellent) based on flow, topic coverage, and usefulness. For reliability, we assessed whether each response addresses the query and is factually accurate, elaborating on specific inaccuracies. For readability, we assessed Flesch-Kincaid Grade Level, Gunning Fog Index, and Simple Measure of Gobbledygook. RESULTS/ANTICIPATED RESULTS: This will be a descriptive analysis of ChatGPT responses. For quality and reliability assessments, we will report Fleiss' kappa for inter-rater reliability and expect substantial agreement or greater (≥0.61). Per prior studies in other domains, we hypothesize that ChatGPT responses will have good quality on average (i.e., GQS score near 4). We hypothesize that nearly all responses will address their query and will mostly be accurate; a minority of responses may have partial inaccuracies. Finally, we hypothesize that readability metrics will suggest that a higher educational level (e.g., college-level education) is required for comprehension. Overall, these findings will help elucidate strengths and limitations of ChatGPT for AML and guide discussion of factors patients should be aware of when using ChatGPT. DISCUSSION/SIGNIFICANCE OF IMPACT: No prior study has examined the educational quality of ChatGPT for AML. Our study will detail whether patients are receiving trustworthy and meaningful information, identify misinformation, and provide guidance to oncologists when recommending information resources to patients or fielding questions that patients may raise after using ChatGPT.

363

The art and science of data navigation for translational research

Ram Gouripeddi, Vanessa Amburgey, Mary Irion, Randy Madsen and Julio Facelli

University of Utah

OBJECTIVES/GOALS: Translational researchers spend significant amounts of time finding available datasets and other research data resources for their purposes. Objectives of this program are develop and evaluate a multipronged approach to supporting researchers with existing data resources. METHODS/STUDY POPULATION: We established a dedicated service with expertise in data resources to increase awareness, understanding, and utilization of existing data resources. This program assists investigators and trainees discover appropriate data resources, formulate scientific problems in computable formats, advise on state-of-the-art data analytics, data management, build collaborations, mentor data users, and develop a service pipeline for streamlined data resource project management. This is accomplished through these essential functions: (1) Discover, catalog, document, and manage metadata resources, (2) train and present data resources to the research community, (3) provide individual consultations, and (4) explore and assess novel data resources. RESULTS/ANTICIPATED RESULTS: In a phased approach, the data navigation program is performing outreach to the research community and integrating with existing data efforts on campus, presenting and demonstrating existing data resources, established a consultation service, and building core competencies into long-term usage and navigation of resources across campus. Evaluating the program monthly has shown an increase in various metrics for evaluating commitment and engagement including number of requests for access to data resource, consultations, publications and presentations, co-authorship, and proposals. Unawareness and inappropriate use of data resources leads to delays in performing research and potentially unnecessary duplications of efforts. DISCUSSION/SIGNIFICANCE OF IMPACT: Our data navigation program has increased use of data resources in research. Next steps are to continue evaluation and further streamline informatics approaches to data discovery, abstraction, formulation, and analysis. Harmonized data resource programs are important translational science approach to foster the next generation of research.

365

Childhood Exposure to Violence and Adolescent Depression: The Role of Epigenetic Aging in Risk Identification

Victoria Nguyen

Columbia University

OBJECTIVES/GOALS: This study examines associations between childhood violence exposure, accelerated biological aging, and adolescent depression using DNA methylation-based epigenetic clocks. Findings aim to identify biomarkers for early detection, guide interventions, and address youth mental health disparities. METHODS/STUDY POPULATION: Data from the Future of Families and Child Wellbeing Study (N = 4,898), a diverse urban U.S. cohort, were analyzed. Childhood violence exposure, assessed using the Parent-Child Conflict Tactics Scale, included measures of physical, emotional, and psychological aggression and neglect. Biological aging at age 15 was evaluated using second-generation epigenetic clocks derived from saliva DNA methylation patterns, while depressive symptoms were measured with the CES-D scale. Multiple linear regression models tested associations between early violence exposure, epigenetic aging, and depressive symptoms, adjusting for socioeconomic status, caregiver mental health, and other key covariates. RESULTS/ANTICIPATED RESULTS: Preliminary results suggest that early violence exposure may be linked to accelerated biological aging and depressive symptoms during adolescence, a critical developmental period. Epigenetic clocks offer an objective method for identifying high-risk youth, complementing mental health evaluations. With further validation and participatory action research, these findings could guide the development of biomarkers for longitudinal testing

https://doi.org/10.1017/cts.2024.990

in school-based screenings and community health programs. These tools aim to be accessible, culturally relevant, and tailored to diverse populations, enhancing early detection, informing personalized interventions, and supporting scalable clinical applications. DISCUSSION/SIGNIFICANCE OF IMPACT: This study explores links between early adversity, biological aging, and mental health, advancing understanding of adolescent depression. Epigenetic biomarkers could improve risk detection and guide tailored interventions in schools and community settings, enhancing access and reducing disparities.

366

Universal representation of human diseases using large language models

Geoffrey Siwo, Ellen R. Bowen and Akbar K. Waljee
University of Michigan Medical School

OBJECTIVES/GOALS: Understanding the interconnections among over 20,000 human diseases spanning organ systems could inform more precise diagnosis and treatment of diseases. Here, we examine whether the ability of large language models (LLMs) to learn universal representations of concepts can be leveraged to discover complex relationships across human diseases. METHODS/STUDY POPULATION: To address the challenge of computationally representing thousands diseases spanning multiple organ systems, we used internal representations of concepts by LLMs to encode diseases based on their descriptions from standard disease ontologies (ICD10 and Phecodes). To do this, we leveraged application programming interfaces (APIs) of three LLMs-GPT3.5, Mistral and Voyage to encode disease relationships. We then performed unsupervised clustering of the diseases using their encodings (embeddings) from each LLM to determine whether the resulting clusters reflect disease relationships. To enable deeper exploration of disease relationships, we developed interactive plots that provide a system level view of the relationships between thousands of diseases and their association with specific organ systems. RESULTS/ANTICIPATED RESULTS: We found that unsupervised analysis of disease relationships using the LLM encodings reveal high similarities among diseases based on organ systems they affect. All the LLMs clustered diseases into groups largely defined by the organ systems they affect without being trained to specifically classify diseases into their corresponding organ system classification. An exception to this was tumors in which we observed that most tumors cluster together as a group irrespective of the organs they affect. Interestingly, we found that tumors affecting anatomically related organs show higher similarity to each other than to those affecting distantly related organs. In addition to anatomical relationships between diseases, we found that the LLM embeddings capture genetic relationships between diseases. DISCUSSION/SIGNIFICANCE OF IMPACT: Overall, we found that the LLM-derived encodings uphold biologically and clinically significant relationships across organ systems and disease types. These results suggest that LLM encodings could provide a universal framework for representing diseases as computable phenotypes and enable the discovery of complex disease relationships.

367

Localization of critical speech areas in glioma-infiltrated brain cortex using local neuronal field potentials via electrocorticography (ECOG)*

Vardhaan Ambati[1], Alexander Aabedi[1], Youssef Sibih[1], Sanjeev Herr[1], Sena Oten[1], Hunter Yamada[1], Jasleen Kaur[1], David Brang[2] and Shawn Hervey-Jumper[1]
[1]University of California, San Francisco and [2]University of Michigan

OBJECTIVES/GOALS: The standard care for malignant gliomas includes maximal tumor resection, but challenges arise near functional (speech) areas. Direct cortical stimulation (DCS) identifies functional (nonresectable) cortex. We aim to identify electrophysiologic (via subdural electrode recordings [ECOG]) biomarkers of DCS-positive (functional) areas. METHODS/STUDY POPULATION: Our lab maintains one of the largest datasets of electrophysiology analysis of glioma infiltrated brain cortex in the USA. Recordings of intraoperative brain mapping were analyzed to identify cortical sites that were found to be positive (functional) during DCS. DCS positive and negative (nonfunctional) sites were aligned to corresponding subdural electrodes. Future analysis: We plan to compare the temporal and spectral electrophysiologic variations associated with cortical sites found to be DCS positive versus negative during brain mapping. We plan to train machine learning classifiers that utilize these electrophysiologic biomarkers to discriminate between DCS positive and negative sites. RESULTS/ANTICIPATED RESULTS: In total, our database comprised of 110 resections with brain mapping (DCS) and ECOG, including 4 patients who underwent a second procedure for resection. Eight patients were excluded as their resections were for brain metastases, not glioma. Our final cohort was comprised of 98 glioma resections, including 4 patients who underwent surgery twice for recurrence. During these resections, a total of 1393 sites were mapped via DCS for language function (including picture naming, word reading, and sentence syntax tasks). Of these 1393 sites, 100 sites were found to be DCS positive (7.1% positivity rate). (Currently in the process of conducting analysis comparing electrophysiologic features and biomarkers of DCS positive versus negative sites.) DISCUSSION/SIGNIFICANCE OF IMPACT: This research is ongoing. Identifying electrophysiologic biomarkers of critical DCS-positive regions may provide a durable alternative to stimulation mapping. Due to its resource intensity, DCS has access barriers. Future neurosurgeons may use biomarkers from subdural electrode recordings to plan safer cortical resections.

368

Uncovering links between innate immunity, DNA repair, and cognitive health in aging populations

Jamie Leiriao[1], Hannah Lords[2] and Paola Sebastiani[3]
[1]Tufts University; [2]Candidate Bioinformatics and Computational Biology Program, Boston University and [3]Clinical and Translational Science Institute, Tufts University

OBJECTIVES/GOALS: Neurodegenerative diseases involve progressive neuronal loss or dysfunction, often due to accumulated

https://doi.org/10.1017/cts.2024.990