Clini-Compare: An Interactive Patient-Similarity Visualization Tool for Clinical Decision Support

Arpit Mathur
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA
arpitmam@andrew.cmu.edu

Adam Perer
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA
adamperer@cmu.edu

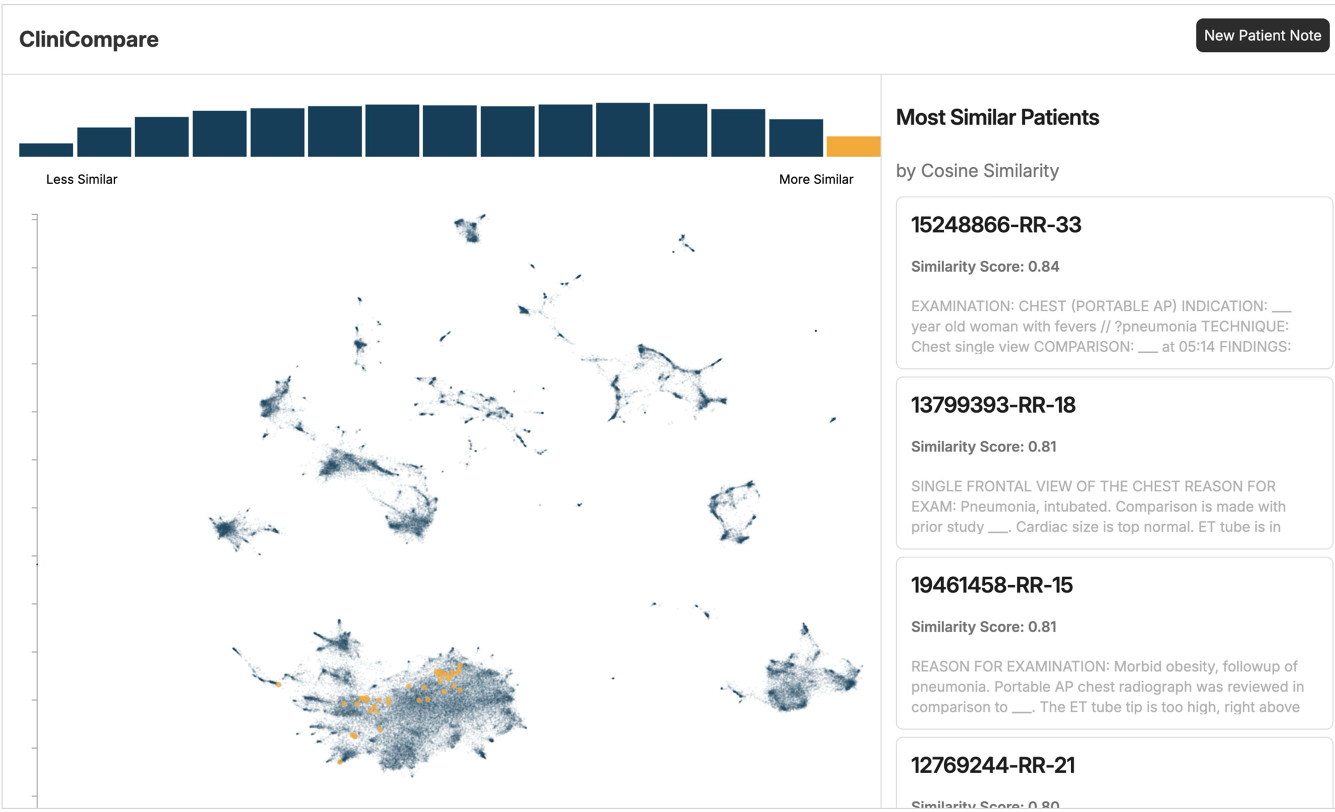


Figure 1: Clini-Compare interface visualizing patients similar to a given patient with a clinical note describing pneumonia.

Abstract

Given a patient profile, the retrieval of health records of previously treated patients with similar characteristics has proven valuable across various healthcare applications, including recruitment for clinical trials, the education and training of clinicians, and the dissemination of healthcare information. Although many such models and algorithms are available, interactive patient similarity systems for complex disease decision support remain largely unexplored. This paper introduces Clini-Compare, an interactive system that evaluates patient similarity based on the semantic analysis of their clinical notes to support decision-making for treatment.

CCS Concepts

• **Human-centered computing → Visualization systems and tools**; **Interactive systems and tools**; • **Applied computing → Health care information systems**.

Keywords

Clinical Decision Support, Interactive Healthcare Systems, AI in Healthcare

ACM Reference Format:

Arpit Mathur and Adam Perer. 2025. Clini-Compare: An Interactive Patient-Similarity Visualization Tool for Clinical Decision Support. In *Proceedings of (CHI '25 Workshop on Envisioning the Future of Interactive Health).* ACM, New York, NY, USA, 5 pages.

CHI '25 Workshop on Envisioning the Future of Interactive Health, Yokohama, Japan

1 Introduction

Patient similarity is a valuable concept in healthcare, enabling physicians to refer to previously encountered cases with similar characteristics, by themselves or other physicians, to inform their decision-making [1, 23]. Prior research has explored patient similarity across multiple healthcare applications, including clinical trial matching [25], where past patient cases help identify suitable candidates for trials; education and training [3], where historical patient records are used for medical learning and simulation-based training; and information dissemination [10], where virtual patient profiles are used as a means to communicate healthcare awareness to the public.

However, patient similarity has not been explored as a method for complex disease decision support. While the machine learning (ML) community has developed numerous algorithms to identify similar patients using structured data [19, 23, 24, 27], unstructured data [8, 14, 25], or both [17, 26]; much of this work has remained algorithm-focused rather than user-driven. Clinical notes, in particular, have been recognized as a rich source of semantic information for capturing patient characteristics [9, 17], yet interactive systems that validate the utility of patient similarity using these notes in real-world decision-making remain absent. It is unknown how well such systems might fit into clinical workflows, how physicians interpret similarity-based recommendations, and how trust is calibrated in such approaches.

This paper presents Clini-Compare, an interactive tool designed to support clinical decision-making by retrieving semantically similar patients based on radiology notes. Given a new patient's radiology note, Clini-Compare searches the MIMIC-IV de-identified radiology clinical note dataset [12, 15], and presents the clinical note and discharge summary of patients that are semantically the most similar to the input patient based on their clinical notes.

Through a task-based user study, we aim to use Clini-Compare as a technology probe to explore the following research questions:

(1) **What is the perceived utility of a patient similarity system in clinical decision-making for physicians?**
(2) **How do physicians calibrate their trust in patient similarity systems compared to other AI interventions in healthcare?**
(3) **What are the key opportunities and challenges in designing patient similarity systems for real-world clinical settings?**

2 Related Work

2.1 Case-Based Reasoning

The practice of utilizing previous patient cases to inform clinical decision-making is known as Case-Based Reasoning (CBR). CBR involves making decisions for new cases by referencing solutions from previous cases that are similar to the current case. This approach has proven to be effective in multiple healthcare applications. Dussart et al. [11] argued that interactive CBR systems are valuable as they closely model the human decision-making process, that is, "I have seen a patient like this." Another study strengthens this argument by indicating that the adoption of evidence-based approaches can improve the reasoning of physicians and reduce biases [5]. Several controlled studies have also measured the performance of CBR, with results showing that it can improve patient outcomes for complex cases [2, 4, 21]. These findings suggest that patient similarity solutions in clinical settings could be beneficial in supporting physicians in complex decision making.

2.2 Patient Similarity Algorithms

In the Machine Learning (ML) and Natural Language Processing (NLP) communities, multiple models and algorithms have been developed to identify patient similarity. These models have been applied to various chronic diseases [18], such as diabetes and hypertension [6]. Studies have identified clinical notes as a rich source of information about patients [9, 17], making them a convenient modality for modeling similarity. For instance, one model was developed to identify patients eligible for clinical trials based on the unstructured clinical notes available in their Electronic Health Records (EHR) [25]. OncoLLM [13] is a large language model (LLM)-based approach to match patients based on unstructured EHR data for clinical trials. Although the focus of these studies has been on developing efficient algorithms and measuring their performance, they do not qualitatively assess interaction with these techniques in real-world clinical settings. There remains a gap for interactive systems that allow physicians to provide qualitative insights into the patient similarity use case, along with the unique challenges and opportunities they present.

2.3 Interactive Systems

The HCI community has made significant contributions in interactive healthcare systems, including those driven by artificial intelligence for decision support. However, as mentioned previously, there is a lack of systems and studies that examine the use of patient similarity methods within physician workflows. CareFlow [20] is one such interactive tool that analyzes patient cohorts and quantitatively measures outcomes based on different care plans for similar patients. This tool is based on structured data, which, while convenient for modeling purposes, is not always readily available in real-world settings. Clinical notes, on the other hand, are generated by physicians during visits, and thus ready for use as needed by the clinician. Given that structured and unstructured data vary significantly in nature and performance, it is worth evaluating qualitative differences in the interactions they afford.

3 System Overview

Clini-Compare [1] (Figure 1) is an interactive tool designed to assist physicians by providing patient similarity assessments using unstructured clinical radiology notes. The system computes semantic embeddings to assess similarity between clinical notes, facilitating comparison of past and current patient cases in complex disease scenarios.

The system is designed for use by physicians in the post-documentation workflow, after a clinical note describing the patient has been generated. At this stage, it provides notes on past patients most similar to the current one based on their description. Physicians can use this information to compare treatment plans and diagnoses, supporting their decision-making process. Clini-Compare can also facilitate

[1] Source code available at https://github.com/MrMathur/clini-compare

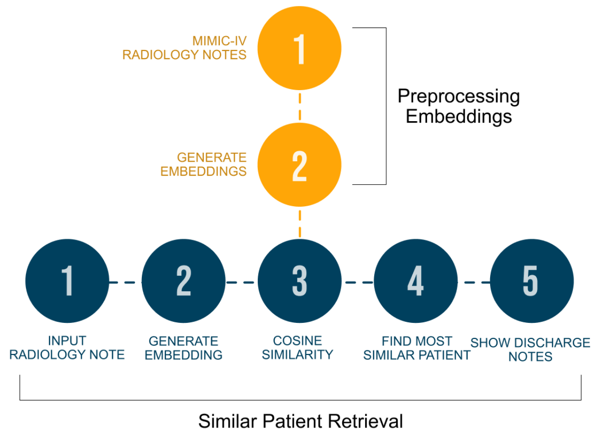


Figure 2: The system pipeline describing the (a) preprocessing stage of generating embeddings for the clinical note database, and (b) the core functionality of retrieving similar patients

communication with patients or assist in case discussions with other physicians.

3.1 Information Pipeline

The information pipeline for Clini-Compare is illustrated in Figure 2.

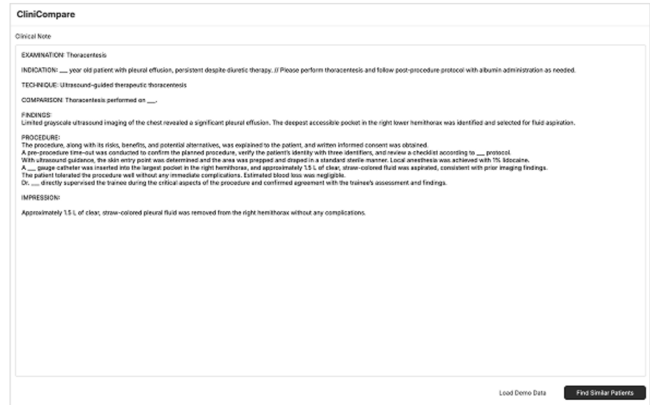
3.1.1 Preprocessing and Embedding Generation. Clini-Compare preprocesses a large dataset of de-identified radiology notes from the MIMIC-IV clinical dataset. It generates embeddings for each note using three models: all-MiniLM-L6-v2 (a general-purpose model) [22], ClinicalBERT (trained on MIMIC-III clinical notes) [14], and BioSentVec (trained on PubMed journals and MIMIC-III clinical notes [7].

3.1.2 Patient Similarity Matching. The system allows users to input a new clinical radiology note, which is then transformed into an embedding using one of the three selectable models. Cosine similarity is computed between the input note's embedding and the preprocessed embeddings. The results are visualized through a dashboard that displays the most similar notes based on similarity scores. Users can select a patient to review their discharge summary for relevant details that support clinical decision-making. In addition, common words between the notes are highlighted to provide insight into the basis of their similarity.

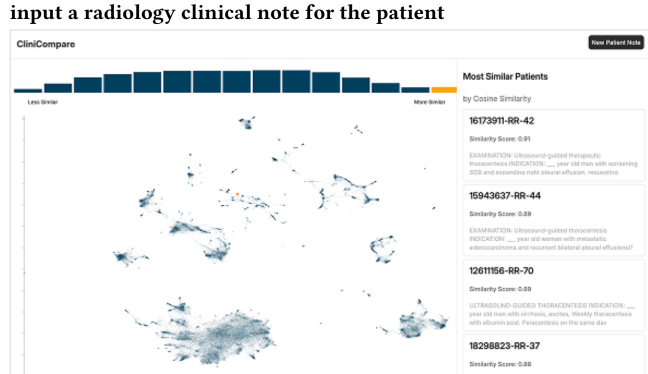
3.2 Visualization Techniques

Clini-Compare incorporates three key visualization techniques to help users gain insight into the structure and distribution of patient data in relation to the input patient.

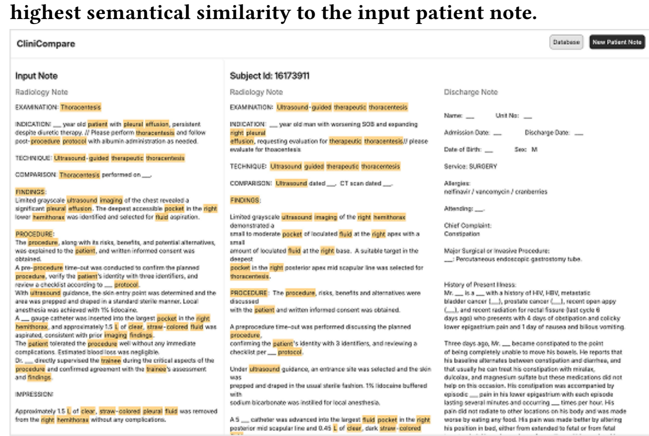
3.2.1 Scatterplot: UMAP Projection of Patient Notes. Clini-Compare features a UMAP [16] (Uniform Manifold Approximation and Projection) of the patient notes. UMAP is a dimensionality reduction technique that projects embeddings of all MIMIC-IV patient radiology notes into a 2D space. This visualization allows users to (i)



(a) The input page features a UI textbox that allows the clinician to input a radiology clinical note for the patient



(b) The dashboard features a histogram to show the distribution of cosine similarity values across the patient database, a UMAP of the clinical notes in the patient database, and a list of patients with the highest semantical similarity to the input patient note.



(c) The patient compare view features a text alignment visualization comparing the input patient and the selected patient side by side. The clinician is also presented with a discharge summary for the selected patient.

Figure 3: Overview of the system interface, showing (a) the input screen, (b) the exploratory dashboard, and (c) the patient compare view.

identify clusters of similar patients in the dataset, (ii) identify outliers that are semantically distant from the rest of the population, and (iii) understand the coverage of how well the patient population is represented across various types of clinical notes.

3.2.2 Histogram: Cosine Similarity. The histogram of cosine similarity represents the distribution of similarity scores between the input clinical note and all patient notes in the database. The cosine similarity metric is used to quantify the similarity between two clinical notes based on their embedding vectors, with values ranging from -1 (completely dissimilar) to 1 (identical). The distribution of similarity scores helps users quickly assess whether most patients in the database are highly similar or dissimilar to the input case.

3.2.3 Highlighting: Text-Alignment. The patient view presents a side-by-side comparison of the input clinical note and the selected patient's clinical note. Common terms between the two notes are highlighted, providing an overview of their similarity. This allows the user to assess whether the system's similarity measurement aligns with meaningful clinical characteristics in the context of their usage.

4 Future Work

The proposed study will use Clini-Compare as a technology probe to explore the use of patient similarity systems in the context of clinical decision-making. A task-based user study will be conducted in which physicians make treatment decisions about a mock patient persona while receiving decision support from the tool. Participants will be provided with a list of symptoms, health conditions, and a patient history, based on which they would draft a clinical note describing the fabricated patient persona. They will then interact with Clini-Compare to input the note and surface relevant information regarding similar patients to determine the diagnosis and treatment plan for the assigned patient. A think-aloud protocol will allow us to capture their decision-making processes and challenges. We will conclude the study with a semi-structured interview to understand the opportunities and challenges associated with the potential use of the system to enhance decision-support and communication in real clinical settings.

4.1 Data Collection and Analysis

Data will be collected through both interaction logs and interviews. Qualitative data from the think-aloud sessions and interview transcripts will be analyzed using bottom-up affinity mapping to identify key themes related to:

(1) **Perceived Utility:** How useful do clinicians find patient similarity systems for decision-making in complex cases?
(2) **Trust:** How do clinicians trust the system compared to other AI interventions in healthcare?
(3) **Opportunities and Challenges:** What are the opportunities and challenges in integrating these systems into clinical workflows?

The tool is currently built on MIMIC-IV radiology clinical notes, but in the future can be extended to other types of clinical notes. Future work should also explore differences in patient similarity workflows for clinicians when using structured data versus unstructured data.

5 Conclusion

In this paper, we introduce Clini-Compare, an interactive system that leverages semantic embeddings of clinical radiology notes to find information about similar patients that can help support complex disease decision-making. Through a user study using Clini-Compare as a technology probe, our aim is to explore the perceived utility, trust, and challenges of patient similarity systems in real-world clinical settings. The contribution from this study will be twofold: (i) providing an open source interactive tool for patient similarity workflows, and (ii) providing insights into how such systems can enhance clinical decision-making.

References

[1] Agnar Aamodt and Enric Plaza. 1994. Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches. *AI Communications* 7, 1 (1994), 39–59. doi:10.3233/AIC-1994-7104
[2] Mobyen Uddin Ahmed, Shahina Begum, and Peter Funk. 2012. Case studies on the clinical applications using case-based reasoning. In *2012 Federated Conference on Computer Science and Information Systems (FedCSIS)*. IEEE, 3–10. https://ieeexplore.ieee.org/abstract/document/6354404
[3] Naomi B. Berman, Steven J. Durning, Mark R. Fischer, Sören Huwendiek, and Marc M. Triola. 2016. The Role for Virtual Patients in the Future of Medical Education. *Academic Medicine* 91, 9 (2016), 1217–1222. doi:10.1097/ACM.0000000000001146
[4] Isabelle Bichindaritz. 2008. Case-Based Reasoning in the Health Sciences: Why It Matters for the Health Sciences and for CBR. In *Advances in Case-Based Reasoning*, Klaus-Dieter Althoff, Ralph Bergmann, Mirjam Minor, and Alexandre Hanft (Eds.). Springer, Berlin, Heidelberg, 1–17. doi:10.1007/978-3-540-85502-6_1
[5] Jennifer S. Blumenthal-Barby and Hannah Krieger. 2012. Rationality in medical decision making: a review of the literature on doctors' decision-making biases. *Journal of Evaluation in Clinical Practice* 18, 5 (2012), 937–947. doi:10.1111/j.1365-2753.2012.01832.x
[6] Sherry-Ann Brown. 2016. Patient Similarity: Emerging Concepts in Systems and Precision Medicine. *Frontiers in Physiology* 7 (Nov. 2016). doi:10.3389/fphys.2016.00561 Publisher: Frontiers.
[7] Qingyu Chen, Yifan Peng, and Zhiyong Lu. 2019. BioSentVec: creating sentence embeddings for biomedical texts. In *2019 IEEE International Conference on Healthcare Informatics (ICHI)*. IEEE. doi:10.1109/ichi.2019.8904728
[8] Dongjin Choi, Andy Xiang, Ozgur Ozturk, Deep Shrestha, Barry Drake, Hamid Haidarian, Faizan Javed, and Haesun Park. 2023. WellFactor: Patient Profiling using Integrative Embedding of Healthcare Data. arXiv:2312.14129 [cs.LG] https://arxiv.org/abs/2312.14129
[9] Shaika Chowdhury, Chenwei Zhang, Philip S. Yu, and Yuan Luo. 2019. Hierarchical Semantic Correspondence Learning for Post-Discharge Patient Mortality Prediction. doi:10.48550/arXiv.1910.06492 arXiv:1910.06492 [cs].
[10] Cook Da and M.M. Triola. 2009. Virtual patients: a critical literature review and proposed next steps. *Medical Education* 43, 4 (2009), 327–332. doi:10.1111/j.1365-2923.2008.03286.x
[11] Claude Dussart, Pascal Pommier, Valérie Siranyan, Gilles Grelaud, and Sophie Dussart. 2008. Optimizing clinical practice with case-based reasoning approach. *Journal of Evaluation in Clinical Practice* 14, 5 (Oct. 2008), 718–720. doi:10.1111/j.1365-2753.2008.01071.x
[12] Ary L. Goldberger, Luis A. N. Amaral, Leon Glass, Jeffrey M. Hausdorff, Plamen Ch. Ivanov, Roger G. Mark, Joseph E. Mietus, George B. Moody, C-K. Peng, and H. Eugene Stanley. 2000. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation* 101, 23 (2000), e215–e220. doi:10.1161/01.CIR.101.23.e215
[13] Shashi Kant Gupta, Aditya Basu, Mauro Nievas, Jerrin Thomas, Nathan Wolfrath, Adhitya Ramamurthi, Bradley Taylor, Anai N. Kothari, Regina Schwind, Therica M. Miller, Sorena Nadaf-Rahrov, Yanshan Wang, and Hrituraj Singh. 2024. PRISM: Patient Records Interpretation for Semantic Clinical Trial Matching using Large Language Models. doi:10.48550/arXiv.2404.15549 arXiv:2404.15549 [cs].
[14] Kexin Huang, Jaan Altosaar, and Rajesh Ranganath. 2019. ClinicalBERT: Modeling Clinical Notes and Predicting Hospital Readmission. *arXiv:1904.05342* (2019).
[15] Alistair Johnson, Tom Pollard, Steven Horng, Leo Anthony Celi, and Roger Mark. 2023. MIMIC-IV-Note: Deidentified free-text clinical notes (version 2.2). doi:10.13026/1n74-ne17
[16] Leland McInnes, John Healy, and James Melville. 2018. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv preprint arXiv:1802.03426* (2018).
[17] Hoda Memarzadeh, Nasser Ghadiri, Matthias Samwald, and Maryam Lotfi Shahreza. 2022. A Study into patient similarity through representation learning

from medical records. *Knowledge and Information Systems* 64, 12 (Dec. 2022), 3293–3324. doi:10.1007/s10115-022-01740-2 arXiv:2104.14229 [cs].

[18] Kenney Ng, Uri Kartoun, Harry Stavropoulos, John A. Zambrano, and Paul C. Tang. 2021. Personalized treatment options for chronic diseases using precision cohort analytics. *Scientific Reports* 11, 1 (Jan. 2021), 1139. doi:10.1038/s41598-021-80967-5 Publisher: Nature Publishing Group.

[19] Ronald Wihal Oei, Wynne Hsu, Mong Li Lee, and Ngiap Chuan Tan. 2023. Using similar patients to predict complication in patients with diabetes, hypertension, and lipid disorder: a domain knowledge-infused convolutional neural network approach. *Journal of the American Medical Informatics Association* 30, 2 (Feb. 2023), 273–281. doi:10.1093/jamia/ocac212

[20] Adam Perer and David Gotz. 2013. Data-driven exploration of care plans for patients. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems* (Paris, France) *(CHI EA '13)*. Association for Computing Machinery, New York, NY, USA, 439–444. doi:10.1145/2468356.2468434

[21] Timothy M. Rawson, Bernard Hernandez, Luke S. P. Moore, Pau Herrero, Esmita Charani, Damien Ming, Richard C. Wilson, Oliver Blandy, Shiranee Sriskandan, Mark Gilchrist, Christofer Toumazou, Pantelis Georgiou, and Alison H. Holmes. 2021. A Real-world Evaluation of a Case-based Reasoning Algorithm to Support Antimicrobial Prescribing Decisions in Acute Care. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America* 72, 12 (June 2021), 2103–2111. doi:10.1093/cid/ciaa383

[22] Nils Reimers and Iryna Gurevych. 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. arXiv:1908.10084 [cs.CL] https://arxiv.org/abs/1908.10084

[23] Jimeng Sun, Fei Wang, Jianying Hu, and Shahram Edabollahi. 2012. Supervised patient similarity measure of heterogeneous patient records. *SIGKDD Explor. Newsl.* 14, 1 (Dec. 2012), 16–24. doi:10.1145/2408736.2408740

[24] Ni Wang, Muyu Wang, Yang Zhou, Honglei Liu, Lan Wei, Xiaolu Fei, and Hui Chen. 2022. Sequential Data–Based Patient Similarity Framework for Patient Outcome Prediction: Algorithm Development. *Journal of Medical Internet Research* 24, 1 (Jan. 2022), e30720. doi:10.2196/30720 Company: Journal of Medical Internet Research Distributor: Journal of Medical Internet Research Institution: Journal of Medical Internet Research Label: Journal of Medical Internet Research Publisher: JMIR Publications Inc., Toronto, Canada.

[25] Michael Wornow, Alejandro Lozano, Dev Dash, Jenelle Jindal, Kenneth W. Mahaffey, and Nigam H. Shah. 2024. Zero-Shot Clinical Trial Patient Matching with LLMs. doi:10.48550/arXiv.2402.05125 arXiv:2402.05125 [cs].

[26] Dongdong Zhang, Changchang Yin, Jucheng Zeng, Xiaohui Yuan, and Ping Zhang. 2020. Combining structured and unstructured data for predictive models: a deep learning approach. *BMC Medical Informatics and Decision Making* 20, 1 (Oct. 2020), 280. doi:10.1186/s12911-020-01297-6

[27] Ping Zhang, Fei Wang, Jianying Hu, and Robert Sorrentino. 2014. Towards Personalized Medicine: Leveraging Patient Similarity and Drug Similarity Analytics. *AMIA Summits on Translational Science Proceedings* 2014 (April 2014), 132–136. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4333693/