

EVALUATING THE USE OF ELECTRONIC MEDICAL RECORD SEARCH TOOLS IN MODERN HEALTHCARE

S. Narmatha¹, Dr. V. Maniraj²

¹Research Scholar, Dept. of Computer Science, AVVM Sri Pushpam College (Affiliated to Bharathidasan University, Tiruchirappalli), Poondi, Thanjavur, Tamilnadu, India.

²Research Advisor, PG and Research Dept. of Computer Science, AVVM Sri Pushpam College (Affiliated to Bharathidasan University, Tiruchirappalli)., Poondi, Thanjavur, Tamilnadu, India.

ABSTRACT

Healthcare organizations generate vast amounts of data at an increasingly rapid pace, leading to both positive and negative outcomes. The exponential growth of free-text clinical documents in the healthcare sector has resulted in the production of massive quantities of electronic health record (EHR) data. Patients' medical records are widely accessible, with a comprehensive medical chart containing vital clinical information, such as vital signs, prescriptions, demographics, exams, treatment plans, progress notes, risks, vaccination history, allergies, radiological images, and lab results. This information is sourced from various channels, including administrative databases used for billing or care management, patient surveys, and medical records. An electronic medical database stores these patient records.

Healthcare professionals often use terms like "electronic health record," "computerized patient record," and "digital medical record" to describe this data. However, retrieving medical and healthcare-related information is a complex task that requires both time and accuracy. Additionally, some patient information is restricted and requires authorization due to its confidential nature. As such, we have compiled numerous studies on EHR data retrieval, covering various search engines, information retrieval methods, and techniques for accessing healthcare data stored in medical databases. Lastly, we will discuss some of the limitations and challenges associated with these processes.

1. INTRODUCTION

A search engine is an online tool that assists users in locating information on the internet. The primary reasons individuals utilize search engines include shopping, conducting research, and seeking entertainment. Numerous studies have identified three main types of search engines: informational, transactional, and navigational. To gain a better understanding of users' search intentions, major search engines analyze queries and assess each user's goal in order to determine the most relevant results based on their selection.

2. RELATEDWORK

A significant amount of research has been dedicated to helping patients retrieve reliable health data from medical databases. The challenges and limitations of traditional paper-based data collection led to the development of current standards for managing drug-related information in clinical trials. However, numerous electronic tools have since been introduced to assist in the collection and analysis of medication data.

One key area of artificial intelligence (AI) is natural language processing (NLP), which enables computers to understand spoken language. Another important method is syntactic matching, which involves analyzing the words typed by the user to identify relevant keywords for a search, ensuring precise phrase matching. The third approach, semantic matching, focuses on understanding the intent behind the user's query and matching search terms to the underlying meaning of the request.

3. METHODOLOGY

Natural Language Processing (NLP): This study outlines the use of "notational language" input by ophthalmologists during patient encounters to extract structured data related to glaucoma diagnosis and progression [1]. Through NLP techniques, raw query input can be filtered and divided into useful analytical categories. These methods allow for a better understanding of users' comprehension levels and the complexity of the information they seek online, which can then be analyzed and interpreted [2]. An example of this is a health information website where natural language search engines could help users, especially those with limited knowledge or language proficiency, to easily access the information they need. Further research will be conducted to explore future requests.

Studies have demonstrated that electronic health records (EHRs) can aid clinicians in enhancing patient outcomes, expediting treatments, and reducing medication errors. Multiple studies reviewed in this paper highlight the benefits of EHRs, showing how NLP techniques can be applied to electronic data, aiding clinicians in monitoring post-surgical events more efficiently [3].

44	INTERNATIONAL JOURNAL OF PROGRESSIVE	e-ISSN :
IJPREMS	RESEARCH IN ENGINEERING MANAGEMENT	2583-1062
	AND SCIENCE (IJPREMS)	Impact
www.ijprems.com	(Int Peer Reviewed Journal)	Factor :
editor@ijprems.com	Vol. 05, Issue 03, March 2024, pp : 282-288	7.001

Due to the need for manual review of clinical notes, researchers and healthcare professionals often struggle to quickly and effectively analyze large volumes of clinical interactions. NLP enables the automated interpretation of EHRs by examining the context of medical terms and phrases, thus allowing access to high-speed computers for further analysis. From automated quality assessments to studies on comparative effectiveness, NLP has a wide range of potential applications. One key finding is that NLP techniques, such as Part-Of-Speech (POS) tagging, can enhance information retrieval (IR) models, improving identification rates within the biomedical field. Further testing is being done to evaluate the efficacy of machine learning techniques and POS categorization for IR.

Significant improvements from POS tagging have been observed in biological IR applications. However, the research notes that their approach has limitations, including the use of only one POS tagger. They suggest that further research should focus on comparing and reviewing various NLP techniques across different scenarios. Additionally, the generalizability of their results is uncertain, as they only used data from a single healthcare facility.

SYNTACTIC SEARCH:

Medical or clinical data refers to health-related information typically associated with standard patient care or part of clinical trial programs. Patient and disease registries are valuable tools for gathering and tracking clinical information for specific groups of patients. An electronic health record (EHR) is a digital, standardized record of a patient's medical history. The practice of using previously collected data for new purposes beyond its original intent is known as data reuse. EMERSE is an effective and dynamic search engine for accessing free-text documents in electronic medical records (EMRs). The article also discusses other research related to this topic.

SEARCH ENGINES:

In the University of Michigan Health System's registry, over 90% of patients are identified through a manual screening process. However, an automated computer system could perform this labor-intensive, time-consuming task more efficiently and accurately [6]. To automate the search for cancer-related terms in free-text medical data, a technology was developed that compiled over 800 SNOMED codes and more than 2,500 related words and phrases. This system, called the Case Finding Engine (CaFE), scans medical text for relevant terms and flags them for manual review. The registration team has praised the CaFE system for its accuracy and efficiency, and enhancements have already been made based on feedback from registrars. If further research focuses on specific areas, the system's reliability and usability could improve even more. The article explains the process and how the Star Tracker medical database was used, aiming to develop a search feature that integrates demographic, clinical, and diagnostic data to assist users in categorizing patients [7].

The most effective approach, though time-consuming and resource-heavy, is to build an enterprise-level data warehouse. An alternative approach is more compatible with older systems and doesn't require prior integration of those databases, yet it can still provide significant value in a cost-effective manner. The Star Tracker search engine allows for comprehensive population-level searches using existing hardware and database systems. The goal is to quickly gather and process large volumes of unstructured textual data using technical advancements, which is crucial in the medical field. Quick access to information is essential for patient care. The future of healthcare will rely on better methods to retrieve sensitive data hidden in millions of patient records. One potential improvement to the current method of keyword selection from Personal Health Records (PHR) is to use information retrieval (IR) techniques to extract relevant terms from record descriptions rather than solely relying on item titles.

DATA WAREHOUSING AND REUSE OF CLINICAL DATA:

To assist in the creation of Personal Health Records (PHRs) and assess the value of their features, a comprehensive framework has been proposed. While this framework is specific to PHRs, it has been adapted to evaluate other healthcare IT systems as well [9]. The proposed PHR framework, along with its related methods, offers a thorough evaluation of PHR value. The author also reviews the process of coding prescription data in multi-site research settings. Medication data classification, reporting, and analysis were suggested as areas for further exploration [10]. Future studies are needed to assess the practicality, accuracy, repeatability, and scalability of various classification methods, as well as to improve the integration of drug categorization into data management and analysis. To enable the reuse of data whenever necessary, methods to standardize, aggregate, and query data from EHRs are essential [11].

The authors propose a Data Warehouse (DW) approach based on EHRs that promotes data reuse, interoperability, and the rapid aggregation of data across different contexts. This approach allows EHR data to be modeled, transformed, integrated, standardized, and aggregated for future use, leveraging the technologies discussed in the article. Recognizing the importance of clinical data modeling for both reusing clinical information and improving healthcare delivery, this method aims to streamline patient information retrieval, a typically labor-intensive and time-consuming process in healthcare data management. They introduce a novel data warehouse design for medical information systems [12] that

44	INTERNATIONAL JOURNAL OF PROGRESSIVE	e-ISSN :
IJPREMS	RESEARCH IN ENGINEERING MANAGEMENT	2583-1062
	AND SCIENCE (IJPREMS)	Impact
www.ijprems.com	(Int Peer Reviewed Journal)	Factor :
editor@ijprems.com	Vol. 05, Issue 03, March 2024, pp : 282-288	7.001

keeps clinical data up-to-date and simplifies the process for data analysts and clinical managers to extract and analyze the data. However, one significant drawback is the increased maintenance required due to the large volume of data stored.

USING EMERSE:

A study focused on aneurysms in patients who have undergone abdominal transplants aims to determine their prevalence and common symptoms. It specifically seeks to identify arterial aneurysms in patients who have had liver or kidney transplants for over eleven years [13]. The study was limited to reviewing electronic medical records (EMRs) within their institution, as some records were stored on paper or in non-digital formats, or had been transferred to other facilities. The study's findings suggest that daily use of antacids could significantly benefit patients with Head and Neck Squamous Cell Carcinoma (HNSCC), potentially leading to the development of low-toxicity treatment and prevention strategies [14]. By reviewing patient charts through the EMERSE application, the study identified medications taken before or after therapy, along with clinical, demographic, and histological data. Using this tailored search technology, complex queries were formed to extract necessary data.

The EMERSE system, developed by the University of Michigan, was instrumental in analyzing these medical records. It is a full-text search engine designed to retrieve information from EHRs and has shown enhanced sensitivity and specificity in numerous evaluation studies, significantly improving the effectiveness of chart reviews. However, integrating information retrieval (IR) capabilities into EHRs presents challenges, as it requires understanding the varying needs of those extracting information from archived medical records.

OTHERS:

The CER Hub is a platform designed for the systematic and scalable extraction, consolidation, and analysis of clinical data across multiple institutions. It helps address challenges related to comparative effectiveness research, allowing physicians to quickly identify and analyze patient groups similar to their own for clinical insights. This tool facilitates stratified survival analysis and physician-driven cohort selection, leading to insights such as the frequency of BRAF mutations and the survival rates of patients with BRAF-mutant tumors after treatment with BRAF inhibitors. Further research is needed to determine the best methods for incorporating this feature into the EMR clinical workflow to inform clinical decision-making.

OPEN-SOURCE SEARCH ENGINES:

While there is abundant information available through web search engines, these lack the functionality to automatically connect data to specific biological applications. CDAPubMed is an open-source, platform-independent tool designed to search literature based on keywords in EHRs. Its seamless integration with conventional CDAPubMed interfaces makes it appear as an extension of a web browser. The modular design of CDAPubMed allows the biomedical informatics community to contribute to its development, with future integration potential. It is compatible with multiple systems and is available for free for non-commercial purposes.

At the Leon Berard Cancer Treatment Centre in France, a full-text search engine has become standard practice. The application uses multi-level modeling within open EHR systems, significantly reducing the time required to process data for GastrOS. Software maintainability is improved through openEHR model-driven development, although one limitation of the study was that only one developer made updates to the software, and they were unaware of each other's changes. Prior to the research, the second author, responsible for developing the GastrOS and implementing the clinical rules (CR), lacked both domain knowledge and experience with open EHR deployments.

CLINICAL CORRELATION:

The authors' work primarily focuses on integrating unstructured data (like research databases, discharge letters, clinical notes, and diagnostic reports) with structured medical data (including vital statistics, medications, and test results). Their future work aims to combine these different types of data, such as diagnoses, medications, and test results, within a time-based framework for deeper analysis. For instance, they used a combination of structured and unstructured (free-text) data from Electronic Health Records (EHRs) to develop an algorithm for diagnosing exfoliation syndrome (XFS). This algorithm helped extract a likelihood score for XFS from patient EHRs, offering a more sensitive and specific method for detecting the condition compared to traditional diagnostic methods. They also highlight that the accuracy of this algorithm could improve with more EHR data.

EMERSE:

The EMERSE (Electronic Medical Record Search Engine) is a powerful tool for retrieving free-text information from medical records. It was developed to address the need for effective searching of medical records for research and data abstraction. EMERSE is user-friendly, enabling even those with limited computer knowledge to perform complex search queries. It can be particularly useful for identifying patients affected by issues like medication recalls. While it is

44	INTERNATIONAL JOURNAL OF PROGRESSIVE	e-ISSN:
IJPREMS	RESEARCH IN ENGINEERING MANAGEMENT	2583-1062
	AND SCIENCE (IJPREMS)	Impact
www.ijprems.com	(Int Peer Reviewed Journal)	Factor :
editor@ijprems.com	Vol. 05, Issue 03, March 2024, pp : 282-288	7.001

beneficial for direct patient care, EMERSE could also streamline the process of reviewing a patient's medical history for significant events, which is valuable in a busy clinical setting.

OTHERS:

The centralization of healthcare systems is a challenge, as it makes adapting to changing needs difficult. Real-world applications of medical informatics reveal untapped potential in current systems for better data collection and secondary use, such as for research and surveillance. Data quality remains a significant focus in clinical trials, as the increasing volume of electronically accessible patient data facilitates patient recruitment and trial data documentation. The study found that ICU data could be automatically used for research, reducing the time spent gathering data, although there is a risk of data quality issues, such as typographical errors. The authors advocate for allowing researchers to enter data in bulk via web-based platforms in future multi-center data collection systems.

In terms of clinical narrative text, a framework has been proposed to represent both textual and temporal aspects of data. This flexible framework could help with future research by integrating temporal structures within EHRs, and it aims to create more effective ways to combine textual and time-related data. The authors are also looking at more advanced methods of combining these similarities.

While using EHR data in research can have both advantages and disadvantages, the right integration can help achieve positive outcomes and therapeutic improvements. There is ongoing debate about the ideal methods to integrate research data collection into clinical practices. A challenge is the relative smallness of healthcare data sets due to privacy concerns, though advancements in medical visual information retrieval are making progress despite the scientific challenges. Since visual data is difficult to analyze without comprehensive context, it is important to merge multimodal data sources, with fusion techniques becoming increasingly important for future research. Low clinical use remains a challenge.

SEMANTIC SEARCH IN CLINICAL DATA:

The Bio Patent Miner tool is used to identify biomedical patents, analyzing physiologically relevant words and linking them to each other. This tool aims to improve its memory by adding more templates for identifying related patterns in phrases. Moreover, on-demand mobile healthcare services, which connect patients, healthcare professionals, and hospital records through secure wireless connections, can deliver healthcare directly to patients' homes. The study also explores different semantic similarity measures used in web-based research. However, assessing how closely two statements are semantically related is a challenging task. Traditional methods, like ontology-based approaches, assume that related concepts should exist within the same branch of a conceptual hierarchy, but this can be a complex task given the vast and ever-expanding nature of the internet.

SEMANTIC ANALYSIS:

Over the last two decades, the amount of digital information in various sectors, including healthcare, has rapidly increased. This has led to faster diagnoses through the extraction of essential information from medical records. The authors' work focuses on information extraction via semantic analysis, aiming to improve this area by integrating more complex tasks, such as correcting spelling errors, identifying negative phrases, and analyzing assertions based on probability and speculation. Their goal is to build a system that can handle multiple languages, provided the necessary database dictionaries are available. A new medical information retrieval system was proposed, utilizing a two-stage query expansion approach to enhance performance.

MEDICAL RECORDS SEARCH ENGINE:

Traditional search engines struggle to understand the user's intent, which can be particularly problematic when searching for health-related information. The authors propose a specialized search engine that tailors results to a patient's medical history, improving the relevance of the information retrieved. They tested this system with 18 volunteers, focusing on how information retrieval (IR) techniques can improve the search process by extracting relevant terms from patient records. The study shows that current search engines like Google and Yahoo often return irrelevant results, and a more targeted question-and-answer search engine can provide more accurate outcomes. The research also explored the benefits of using semantic search for better decision-making, helping physicians navigate large amounts of data and avoid information overload. The study suggests that a larger sample size would help assess the effectiveness of this semantic search tool.

OTHERS:

The authors plan to expand their model to integrate other health information systems. In a subsequent experiment with 10 participants, they presented two scenarios requiring the retrieval of patient information. The first scenario provided accurate, patient-specific information, while the second introduced a semantic search tool that could identify relevant terms within a patient's electronic health record (EHR). Participants were interviewed afterward to assess their

A4	INTERNATIONAL JOURNAL OF PROGRESSIVE	e-ISSN :
IJPREMS	RESEARCH IN ENGINEERING MANAGEMENT	2583-1062
	AND SCIENCE (IJPREMS)	Impact
www.ijprems.com	(Int Peer Reviewed Journal)	Factor :
editor@ijprems.com	Vol. 05, Issue 03, March 2024, pp : 282-288	7.001

experience with EHR systems and how semantic search can reduce cognitive load in therapeutic contexts. They also suggest that providing a list of searchable terms could improve the trust and perceived accuracy of the system, and tracking how these perceptions change over time would be an interesting study.

The authors propose a system that analyzes free-text clinical reports for patterns and outcomes, correlates them with external knowledge sources, and tailors the user's record accordingly. One challenge with current search engines is their failure to explain why returned results are relevant to the query, leading to confusion about how to refine searches to get better results. To address this, the authors recommend an information retrieval system that uses domain ontology and provides graphical explanations of the query results, aiming to improve user interaction and comprehension.

COLLABORATIVE SEARCH:

The authors explore a joint search technique using the Unified Medical Language System (UMLS) and electronic health data. UMLS is a set of files and programs that integrate various biological and health vocabularies and standards, helping to facilitate communication between computer systems and EHRs. In their study, they used a collaborative human-computer technique to identify common data elements (CDEs) across eligibility criteria for clinical trials studying diseases like cardiovascular problems and breast cancer. This approach highlights how data from different clinical trials can be integrated for more effective research.

CRITERIA CODES AND SEMANTIC ANNOTATION:

In this research, the authors describe a method for identifying and analyzing terms from the Unified Medical Language System (UMLS) within clinical trial eligibility criteria. They use a semantic annotator to extract relevant terms from free-text eligibility criteria, although it generates a small number of false positives (e.g., "arrange" and "repair"). Their method improves retrieval speed by effectively incorporating medical domain knowledge, including tools like MetaMap and UMLS to expand and enrich the external concept space. This approach illustrates the advantages of domain knowledge in improving the performance of medical search systems.

ELECTRONIC HEALTH RECORD (EHR) SEARCH SYSTEMS:

The authors highlight a limitation in current EHR search systems—users' lack of familiarity with search engines or medical domain knowledge. To address this, they created a collaborative search tool to support teamwork in documenting, improving, and sharing EHR search knowledge. This tool facilitates more accurate and efficient healthcare information retrieval by encouraging social information-foraging mechanisms that enhance the search process. They also discuss incorporating temporal data (e.g., timestamps) into the search system to improve performance by correlating EHR time distributions with search results. Future work will explore combining temporal analysis with medical ontologies for better EHR search efficiency.

QUERY-BASED SEARCH AND LOG ANALYSIS:

The research also explores query-based search and the use of query logs to understand user information-seeking behavior. By analyzing the queries, researchers can explore trends, system performance, and the nature of information-seeking. One limitation of query log analysis is the lack of information about the user's history, which can be supplemented with survey and observational studies to gain deeper insights into user behavior. The authors aim to improve search results by providing intelligent query suggestions using the UMLS knowledge base, which could support the development of more effective methods for retrieving data from electronic medical records (EMRs).

QUERY FORMULATION:

This research emphasizes the importance of user autonomy in querying EHR data. By allowing biomedical researchers to conduct reference interviews and define search criteria based on predetermined topics, the authors aim to give users greater control over their search experience.

They acknowledge that some relevant topics might be overlooked if researchers rely solely on pre-selected search topics and suggest that a more thorough and inclusive approach to query formulation could yield better results. They also discuss the challenges in clinical research query mediation and communication, noting that more focused research is needed to address these issues effectively.

DUPLICATE DOCUMENT DETECTION:

Lastly, the study explores the concept of duplicate document detection, which is used to identify documents with similar content. This is accomplished through "document transformation," where documents are converted into more analytically-friendly formats for easier comparison.

The framework they developed lays the groundwork for identifying and eliminating duplicate documents in medical and clinical research data, which can improve the overall quality and efficiency of information retrieval.

44	INTERNATIONAL JOURNAL OF PROGRESSIVE	e-ISSN :
IJPREMS	RESEARCH IN ENGINEERING MANAGEMENT	2583-1062
and the second	AND SCIENCE (IJPREMS)	Impact
www.ijprems.com	(Int Peer Reviewed Journal)	Factor :
editor@ijprems.com	Vol. 05, Issue 03, March 2024, pp : 282-288	7.001

DUPLICATE DOCUMENT DETECTION AND FILE TRANSFORMATION:

This section emphasizes the importance of identifying and eliminating duplicate documents in electronic health records (EHRs) to improve search quality and efficiency. Documents may have identical content, size, or name, and when duplicates are included in search results, it can diminish search quality. To address this issue, the research focuses on duplicate document detection using algorithms that scan for similarities during the indexing process. These algorithms help to identify and remove duplicates, which in turn enhances search accuracy, speeds up indexing, and reduces runtime. The study presents methods for detecting near-duplicates, utilizing techniques like sentence-level detection algorithms, fingerprinting, and shingling approaches. These methods help identify plagiarism, potential citations, and duplicated documents. By using these techniques, the study offers a framework for efficient duplicate document identification, even with large datasets, and improves content-based automated file renaming.

CHARACTER SHAPE CODING (CSC) AND DETECTION METHODOLOGIES:

The study also introduces the Character Shape Coding (CSC) method for mapping character images according to their position and shape, which is used to detect duplicate documents. This method efficiently eliminates non-textual information by focusing only on textual components. With datasets containing up to 100,000 records, this approach was found to be cost-efficient and effective. Additionally, the study explored various models for duplicate detection, including full-content, partial-layout, full-layout, and document content duplication. These models provide resilience in detecting duplicate documents, improving accuracy when compared to older Optical Character Recognition (OCR) models. The fingerprint approach and signature stability assessment were used to identify identical or highly similar documents, with testing conducted on datasets as large as 50 million papers.

FINGERPRINTING AND SIGNATURE STABILITY:

The research demonstrates how digital signatures, such as fingerprints, can be used to detect duplicate documents in large datasets, like those ranging from 30 MB to 2 GB. The study compares the efficiency of fingerprinting against other techniques like syntactic filtering and shingles, highlighting how the fingerprint method provides accurate results in a fraction of the time. Furthermore, the study examines the effectiveness of duplicate detection in document image databases, including those with false duplicates. This method established an empirical connection between duplication models, comparison measures, and duplicate detection.

TRIE-TREE STRUCTURE AND SPAM DETECTION:

The paper introduces a new approach for detecting duplicate documents using a trie-tree structure, storing 64-bit fingerprints of documents downloaded from the internet. This method was shown to improve accuracy in detecting spam emails, which is another potential application for duplicate detection. In terms of content-based file management, a methodology for automatic file renaming was suggested, involving content analysis, semantic proof of file names, and file categorization.

EHR SEARCH CHALLENGES AND SOLUTIONS:

In the context of electronic health records (EHRs), the complexity of data search is addressed. EHRs contain diverse patient information, from test results and diagnoses to treatments and follow-up appointments. The challenge arises because medical terms may have multiple definitions, and misspellings, acronyms, and synonymous terms often create difficulties for efficient searches. Clinicians frequently use different expressions for the same condition, further complicating searches. A lack of standardized language and punctuation exacerbates the issue. To overcome these challenges, a powerful search engine for EHR data, such as the free-text EHR search engine EMERSE, is necessary. This search engine is designed to help researchers and clinicians efficiently find relevant information within EHRs for research and health-related tasks. It enables faster retrieval of data by addressing context-sensitive interpretations, synonymous terms, and privacy concerns, although it faces limitations related to data security and privacy.

The paragraph addresses the challenges and potential solutions for retrieving information from electronic health records (EHRs). It highlights several important aspects:

- 1. FUTURE DATA INTEGRATION: As it's impossible to visually represent all data, the future of EHR search will rely on multimodal data for better retrieval. There will be a need for approaches that can integrate data from various sources to enhance information retrieval.
- 2. UNDERUTILIZATION OF CLINICAL DATA: Clinical data is often underused, which limits its potential value. To address this, medical search engines will need to prioritize security to protect sensitive patient information from unauthorized access.
- **3. DATA STANDARDIZATION AND ORGANIZATION**: Techniques like natural language processing (NLP) will be crucial in organizing and standardizing clinical notes. This will make the data easier to understand and more accessible for use in research and healthcare.



- 4. EFFICIENT, ADAPTABLE SOLUTIONS: The search engines or information retrieval (IR) systems need to provide scalable solutions to handle disorganized clinical data. They must be able to make sense of complex medical texts with minimal difficulty.
- 5. ACCURATE PATIENT DATA: In order to make effective use of EHRs, accurate and complete data about individual patient groups is necessary for informed decision-making.

4. CONCLUSION

The literature review discusses various strategies for retrieving information from EHRs, highlighting both their advantages and drawbacks. With the increasing volume of clinical data, extracting relevant information for user queries becomes challenging. To address these obstacles, the research considers improvements in EHR search engine technology. However, it notes the lack of understanding regarding user needs for information mining from clinical documents, and warns about the risks involved in incorporating information recovery features into EHRs.

5. REFERENCES

- [1] Ethun, Cecilia G., et al. "Frailty and cancer: implications for oncology surgery, medical oncology, and radiation oncology."*CA: a cancer journal for clinicians*67.5 (2017): 362-377
- [2] Calabresi, Paul, Philip S. Schein, and Saul A. Rosenberg. "Medical oncology: basic principles and clinical management of cancer." (1985).
- [3] Schrag, Deborah, and Morgan Hanger. "Medical oncologists' views on communicating with patients about chemotherapy costs: a pilot survey." *Journal of Clinical Oncology*25.2 (2007): 233-237
- [4] McDermott, Ultan, and Jeff Settleman. "Personalized cancer therapy with selective kinase inhibitors: an emerging paradigm in medical oncology." Journal of Clinical Oncology 27.33 (2009): 5650-5659.
- [5] Muscaritoli, Maurizio, et al. "Prevalence of malnutrition in patients at first medical oncology visit: the PreMiO study." Oncotarget 8.45 (2017): 79884.
- [6] Tong, Ho, Elisabeth Isenring, and Patsy Yates. "The prevalence of nutrition impact symptoms and their relationship to quality of life and clinical outcomes in medical oncology patients." Supportive care in Cancer 17 (2009): 83-90.
- [7] Newell, Girgis, and Ackland. "The physical and psycho-social experiences of patients attending an outpatient medical oncology department: a cross-sectional study." European journal of cancer care 8.2 (1999): 73-82.
- [8] Fadul, Nada, et al. "Supportive versus palliative care: What's in a name? A survey of medical oncologists and midlevel providers at a comprehensive cancer center." Cancer 115.9 (2009): 2013-2021.
- [9] Braun, Ilana M., et al. "Medical oncologists' beliefs, practices, and knowledge regarding marijuana used therapeutically: a nationally representative survey study." Journal of Clinical Oncology 36.19 (2018): 1957.
- [10] Stoffel, Elena M., et al. "Hereditary colorectal cancer syndromes: American Society of Clinical Oncology clinical practice guideline endorsement of the familial risk–colorectal cancer: European Society for Medical Oncology clinical practice guidelines." Journal of clinical oncology 33.2 (2015): 209.