# Enhancing Clinical Decision-Making in Oncology with LLM Agents

Master's Thesis Proposal

Ayda Yilmazer
Supervisor: Xing David Wang

Humboldt-Universität zu Berlin

August 21, 2025

## 1 Introduction

Medicine is a prominent research field for Large Language Models (LLMs). LLMs offer significant potential for various applications in medicine, such as question-answering, medical education, electronic health record generation, assisting in scientific research, and decision-making [1]. Research in the field of clinical decision-making focuses on leveraging LLMs to assist in the interpretation of patient-related data, diagnosis, and treatment planning [2, 3, 4, 5], among other related areas.

A particularly challenging area of clinical decision-making is precision oncology. Precision oncology is a cancer treatment approach that analyzes the individual molecular profiles of each patient's tumors to offer personalized treatment. Cancer is a genomic disease that is characterized by uncontrolled cell growth and presents differently in each patient, resulting in an overall heterogeneous condition. This heterogeneity is not limited to differences between patients, but can also occur within a single patient, which can pose great challenges for cancer treatment [6]. This makes it inherently complex to deal with, often requiring a highly individualized, case-by-case approach, as seen in precision oncology. To address these challenges, precision oncology cases are typically discussed by interdisciplinary teams of medical experts, known as Molecular Tumor Boards (MTBs). MTBs focus on complex cancer cases where treatments are tailored to each patient's individual molecular profile [7]. This is a complex and time-consuming process that requires manually searching through extensive text-based resources. It is therefore considered a major bottleneck in precision medicine [8].

Because of their proficiency in understanding human-like text, LLMs excel at analyzing large volumes of unstructured information—often with a better performance than humans [9]—and applying the extracted knowledge for problem-solving. This ability indicates the potential of LLMs to support decision-making in oncology cases. Several studies have already explored this potential of LLMs; however they often report that foundational LLMs fall short of achieving the desired performance for implementation in clinical practice [5, 10, 11, 12].

One key problem that causes this poor performance is hallucinations, where LLMs generate factually incorrect outputs in a seemingly convincing manner. Hallucinations often arise when LLMs are asked questions without any supporting context. In contrast, when LLMs are given relevant context and tasked to extract information from it, hallucinations are considerably less common [13]. This limitation is attributed to the inherent knowledge gaps of LLMs, which are caused by their training being on general datasets rather than on domain-specific medical data and their limited exposure to up-to-date information.

Addressing this key limitation would greatly benefit general oncologists, who often operate from smaller, private clinics and do not have access to the same amount of resources and support as MTBs. General oncologists typically offer routine care based on established clinical guidelines, while MTBs and university hospitals are more involved in clinical trials and hard-to-treat cases. In addition to helping reduce hallucinations and speeding up the currently time-consuming, manual process of information retrieval, incorporating relevant clinical practice guidelines into LLMs can also support determine whether a case can be managed with standard treatment protocols or involves rare features requiring discussion in an MTB, which can speed up workflows.

In this master's thesis, we aim to enhance an LLM to serve as a more reliable clinical decision-support tool in oncology. By integrating established clinical practice guidelines, we address the

issue of hallucinations in current LLMs, a key limitation that affects their trustworthiness in clinical settings. With this integration, we aim to enable the model to provide guideline-adherent treatment recommendations for common cancer cases and to introduce a classification capability that allows the model to distinguish between routine cases, which can be managed using standard treatment protocols, and more complex ones that require referral to an MTB. Our goal is to streamline decision-making and accelerate workflows for general oncologists, particularly those with limited access to multidisciplinary expertise.

## 2 Background & Related Work

Multiple approaches can help reduce the risk of hallucinations: Retrieval-augmented generation (RAG), fine-tuning on domain-specific data, chain-of-thought (CoT) prompting, and the ReAct method are the most prominent ones [13, 14].

RAG [15] is a technique used to extend the knowledge access of LLMs by retrieving and providing relevant documents as context during generation. Many retrieval methods first convert the text into high-dimensional numerical representations, known as vector embeddings, to efficiently identify the most semantically relevant segments based on the user's query using similarity metrics, such as cosine similarity [16]. The retrieved segments are then used to provide additional context for generating more informed responses. Additionally, RAG has the benefit of requiring less computation than fine-tuning [14], which involves adjusting the weights of a pre-trained language model to adapt it to a specific task. In contrast, RAG only adds a retriever component without modifying the model's original weights. Figure 1 shows a basic overview of the RAG process.

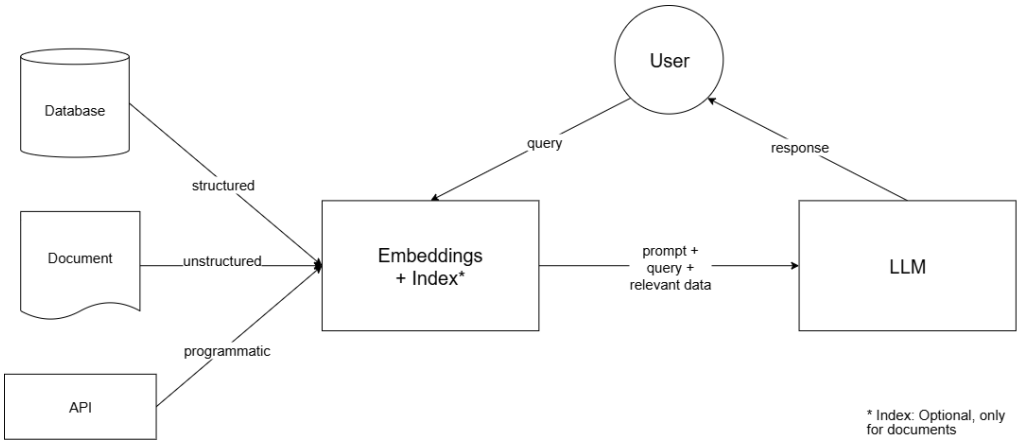


Figure 1: A basic overview of the RAG process

RAG is a widely discussed approach that can be used to address the knowledge gaps of LLMs. Ferber et al. [17] assessed GPT-4 in interpreting clinical guidelines from the American Society of Clinical Oncology (ASCO) and the European Society for Medical Oncology (ESMO). They used RAG to provide information on pancreatic cancer, metastatic colorectal cancer, and hepatocellular carcinoma to GPT-4 and evaluated its performance with 30 clinically relevant questions. GPT-4 without RAG served as the baseline. Manual comparison showed that RAG improved response correctness (84% vs. 57%) and helped identify 11 of 12 expected discrepancies between ASCO and ESMO guidelines, suggesting that model hallucinations can be reduced in this manner.

Khene et al. [18] developed a personalized chatbot (Uro_Chat) based on GPT-3.5-Turbo, using RAG to incorporate the European Association of Urology (EAU) guideline, with LlamaIndex serving as the interface. They evaluated it using 15 pre-existing uro-oncology questions, comparing its answers to those of ChatGPT-3.5 and ChatGPT-4. They reported that Uro_Chat gave adequate responses to all questions. An additional analysis with five recent literature-based questions showed that Uro_Chat answered all correctly, while ChatGPT-3.5 and GPT-4 failed.

Both Ferber et al. [17] and Khene et al. [18] demonstrate the benefits of enhancing LLMs with RAG for guideline retrieval. However, our approach differs in several aspects. Compared to our approach, both Ferber et al. and Khene et al. work with a smaller clinical scope, with Ferber et al. focusing on three cancer types and Khene et al. limiting their work to urology. Their evaluations are based on answering manually curated or literature-derived questions. In contrast, we generate synthetic patient cases using an LLM and prompt the model for treatment recommendations for a more realistic and clinically relevant evaluation. While both prior works

use RAG alone, we combine RAG with ReAct to enhance our model's reasoning capabilities. Our system also introduces a classification capability to distinguish cases that fall within or outside the scope of existing clinical guidelines. Additionally, unlike their models, which provide general responses, our system links recommendations directly to specific guideline sections to improve traceability and transparency of the recommendations. Finally, while Ferber et al. and Khene et al. use proprietary LLMs, we use a local, open-weight model for the advantage of having a transparent architecture and improved control over data privacy.

ReAct (Reason + Act) is another method effective against hallucinations [13]. Introduced by Yao et al. [19], it combines reasoning with external information retrieval to improve the structured and transparent thinking of LLMs. Similar to CoT prompting, which guides the model to generate intermediate reasoning steps before arriving at a final answer [20], ReAct also promotes step-by-step thinking but extends it by allowing the model to take external actions —such as querying a search tool—and incorporate the retrieved information into its reasoning. This is especially valuable in medicine, where the literature is constantly evolving. As an example, when answering a medical question, a ReAct model might first reason about what kind of information is needed, then search a medical database, and integrate the findings into its final answer. This continuous interaction between reasoning and acting helps reduce errors and improve understanding.

In our approach, we adopt the ReAct framework to enhance the model's reasoning capabilities when working with complex clinical cases. By enabling step-by-step reasoning combined with access to external guideline content, we aim for our system to increase the accuracy in its treatment recommendations by reducing hallucinations.

# 3 Objective

The primary goal of this thesis is to develop a LLM-based decision-support system integrating and adhering to established oncology guidelines. We will enhance Llama3.3-70B [21], an open-weight and local LLM, to distinguish between common and rare cancer cases and provide treatment recommendations for common cases. To achieve this, we will integrate relevant clinical practice guidelines using RAG combined with the ReAct framework within an agent-based system using the LlamaIndex framework [22]. By incorporating clinical practice guidelines and utilizing the reasoning capabilities of the ReAct framework, we aim to address the common limitation of hallucinations, which often arise from a lack of access to up-to-date information. The enhanced model will be able to decide whether a patient case falls within the scope of existing clinical guidelines or involves rare molecular aberrations that are not covered by them. This classification capability helps assess whether standard treatment protocols are applicable to a given case or if the case should be referred to an MTB for further evaluation. For cases within the scope of the guidelines, the model will provide treatment recommendations based on the corresponding clinical guideline. It will also cite the specific guideline and relevant section as a source for clinicians to refer back to the original documentation. Our approach aims to improve oncologists' workflows and ensure alignment with the latest clinical guideline updates.

To assess our enhanced model, we will generate fictional cancer patient cases using an LLM-based approach and use these to obtain treatment recommendations. First, we will evaluate the model's ability to classify the cases as either within or outside the scope of the guidelines. For the cases that were correctly classified as in-scope, we will then evaluate the model's treatment recommendations. This evaluation will consist of two parts: First, we will use an LLM-based assessment to compare the treatment recommendations from the enhanced model with the corresponding guideline, evaluating our model's adherence to current standard-of-care. Here, we will also assess the accuracy of the model's citations verifying that it references the correct section of the guideline. Second, we will query baseline models with the same fictional patient cases and compare their recommendations to those of the enhanced model to determine whether the enhancements led to improved performance.

# 4 Approach

## 4.1 Agent-based Approach with LlamaIndex

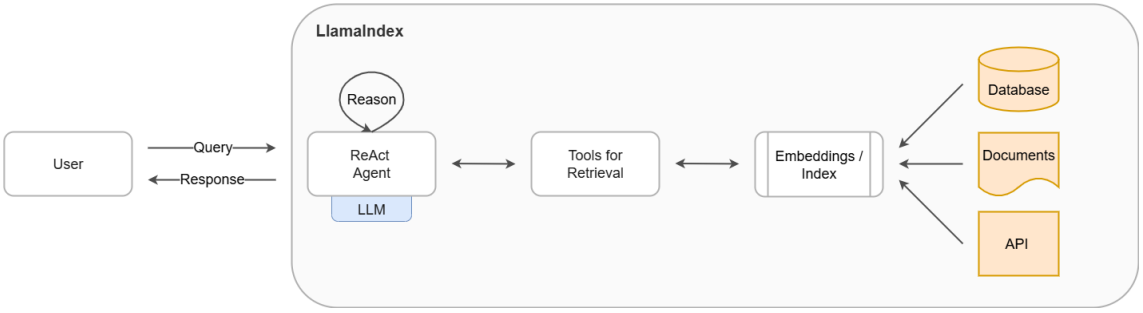


Figure 2: Overview of an agent-based system using LlamaIndex

LlamaIndex is a data framework for building LLM-powered applications that offers the ability to chain multiple processes together, along with built-in tools for data retrieval, indexing, and integration. We chose to base our approach on LlamaIndex because of its native support for agent-based workflows, which support ReAct-style iterative reasoning by coordinating interactions between the language model and various tools or external data sources. This functionality is particularly effective for handling complex tasks, such as clinical decision-making.

We will use the built-in RAG functionality of LlamaIndex to provide the LLM with the relevant guidelines. LlamaIndex supports this by handling document import, chunking, embedding, indexing, and retrieval parts of the RAG process. When responding to a query, the indexing allows the system to quickly retrieve the most relevant information from the available data sources. Figure 2 shows an example diagram demonstrating this system. Here, we see that after the user prompts the agent, it can choose either to reason by thinking through the problem step by step or to act by accessing external knowledge. In our approach, we will develop a custom tool using LlamaIndex to provide the LLM with the clinical practice guidelines.

## 4.2 Incorporating Clinical Practice Guidelines

We will incorporate clinical guidelines from Onkopedia [23] into the LLM. We will download the guidelines in PDF format, then chunk, embed, and index the data using the built-in functions in LlamaIndex. To make the data accessible to the LLM, we will create a custom LlamaIndex tool.

To enable the LLM to cite the specific guideline and sections it referred for the treatment recommendations we will chunk the guidelines structurally. Each chunk will correspond to a specific section of the guideline and will be augmented with metadata, so that the LLM can later reference the exact section the information was retrieved from.

## 4.3 LLM Specification

For this thesis, we will use Meta's Llama-3.3 model with 70 billion parameters. The Llama models have the advantage of being open-weight and locally deployable, which makes them better suited for scenarios involving sensitive personal data, such as patient information in medicine. To run the model, we will use Ollama [24], which is a framework for building and running language models on the local machine.

## 4.4 Generating Fictional Patients as our Dataset

We will assess the enhanced LLM using fictional patient cases. We will generate the fictional patient cases using Onkopedia guidelines. While generating the patients, we will aim to have full coverage on every treatment-branch of a guideline. We will adopt an LLM-based approach for the creation of the fictional cases by using OpenAI's [25] gpt-4o model via the OpenAI API. We aim to reduce model-specific bias, by using a different LLM for generating the patient cases

Onkopedia guidelines typically include images depicting the treatment structure for the disease at various stages. These images are usually found at the beginning of the Treatment section in the guideline. We have prioritized choosing the guidelines with simpler treatment structure trees to avoid overcomplicating the patient generation process. We selected five cancer types based

on their corresponding treatment decision trees. These include Small-Cell Lung Cancer (SCLC) [26], Gastric Cancer [27], Ovarian Cancer [28], Renal Cell Carcinoma (Hypernephroma) [29], and Gastrointestinal stromal tumors (GIST) [30]. Before generating the patient cases, we will first extract the therapy structure image from the relevant guideline and use the LLM to extract all defined treatment pathways. We will review the extracted pathways and refine them, if necessary. When prompting the LLM to generate patient cases, we will provide it with a specific treatment pathway to generate patients accordingly. As an example, Figure 3 shows a therapy structure image from Onkopedia's gastric cancer guideline and Listing 1 provides the extracted JSON tree.

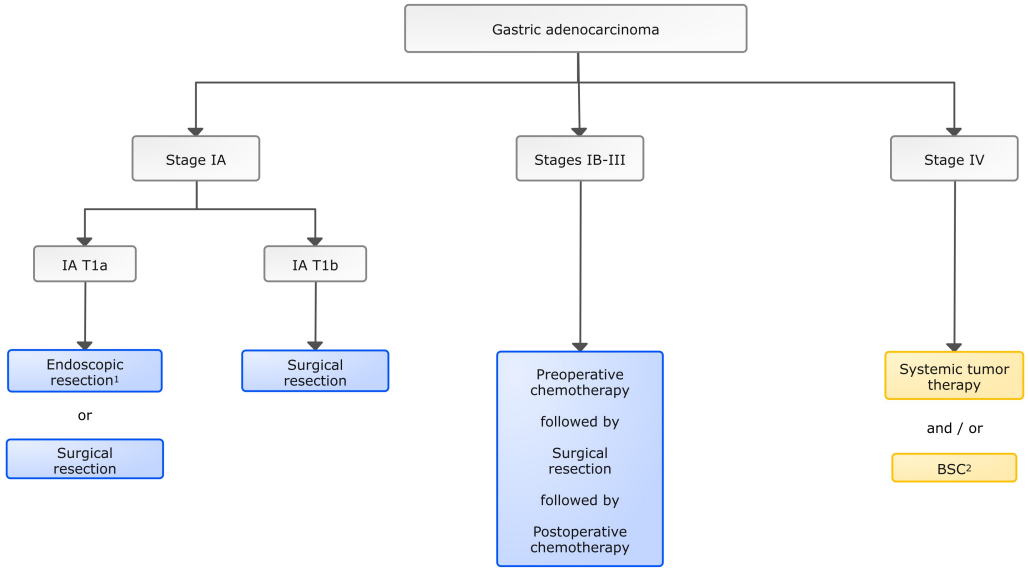


Figure 3: The therapy structure tree from Onkopedia's gastric cancer guideline

There are generally multiple images describing therapy structures in Onkopedia guidelines. Among them, one or more typically refer to the metastatic stage and describing systemic drug therapies intended for non-curative treatment. Figure 4 shows such an image from the Onkopedia gastric cancer guidelines as an example. To generate fictional cases at the metastatic stage, we will also process these images using the approach described above.

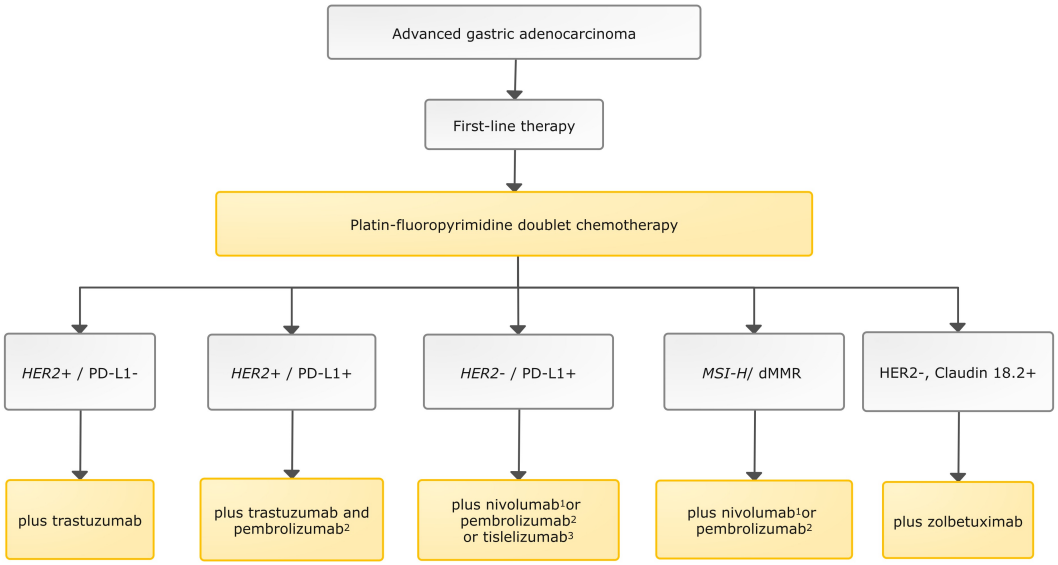


Figure 4: The therapy structure tree for metastatic gastric cancer from Onkopedia

We will construct a prompt template to query the model for generating patient cases, which will include placeholders to later incorporate both the relevant treatment pathway and excerpts from the guideline. We will first flatten the extracted JSON tree into a list of individual treatment

```json
{
  "Gastric adenocarcinoma": {
    "Stage IA": {
      "IA T1a": {
        "treatment_options": [
          "Endoscopic resection",
          "Surgical resection"
        ]
      },
      "IA T1b": {
        "treatment": "Surgical resection"
      }
    },
    "Stages IB-III": {
      "treatment_sequence": [
        "Preoperative chemotherapy",
        "Surgical resection",
        "Postoperative chemotherapy"
      ]
    },
    "Stage IV": {
      "treatment_options": [
        "Systemic tumor therapy",
        "BSC"
      ]
    }
  }
}
```

Listing 1: The extracted treatment tree in JSON format

paths so that each prompt can be supplied with a distinct treatment path. We will use RAG to dynamically insert the relevant guideline excerpts into each prompt based on the corresponding treatment path. Table 1 lists the attributes included in a patient case, along with examples and descriptions for each attribute. We will use few-shot prompting to provide the LLM with real world cases, to generate more realistic cases. Each prompt will include five cases from an anonymized dataset of various MTB cases from Charité [31]. To get incidence rates for each cancer type, we will use epidemiological data from SEER [32].

We will ensure that the generated patient cases are as precise as possible. To achieve this, we will consult with clinicians to manually review and validate the fictional cases. This step is crucial to ensure the evaluation dataset remains clinically realistic and thus more useful for evaluating potential application in clinical practice. For this reason, we will prioritize the quality of the generated patient cases over full coverage of the guidelines. With our fictional patient generation approach, we aim to (1) cover every therapy branch in the guidelines and (2) optimize the data generation process by utilizing an LLM. Figure 5 shows the diagram for fictional patient generation process.

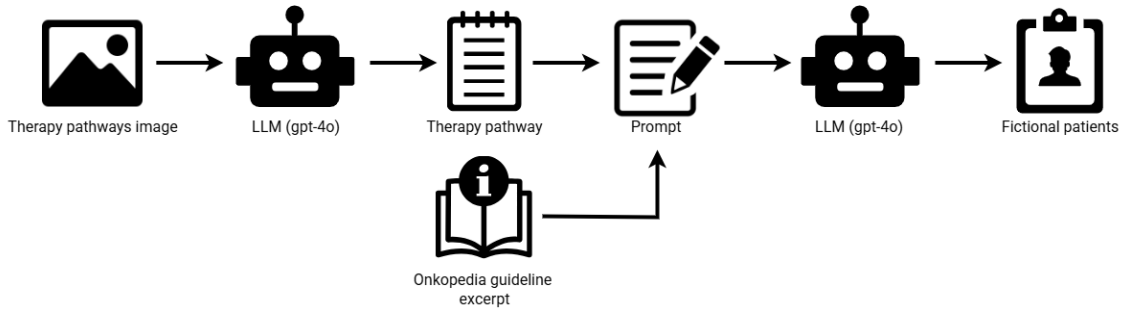


Figure 5: Process diagram for generating fictional patient cases

| Attribute | Example | Description |
|---|---|---|
| Age | 55 | Patient's age in years |
| Gender | Male | Patient's gender |
| Diagnosis | Non-small cell lung cancer | Type of cancer diagnosed |
| Stage | Stage II (T3N0M0) | The extent of a cancer's spread, including codes like TNM and stage groupings such as Stage II |
| Tumor Characteristics | Tumor size: 5 cm; Mitosis rate: high | Physical and biological features of the tumor |
| Histology | Adenocarcinoma | Type of the tumor |
| Driver Mutations | EGFR mutation positive | Presence of genetic mutations |
| PD-L1 Expression | 50% | The percentage of PD-L1 found in tumor cells |
| Treatment Recommendation | Surgery | Suggested treatment plan |
| Follow-Up | 6 months CT scan | Schedule and type of follow-up examinations |
| Classification | True | Whether the case falls within the guideline's scope |

Table 1: Attributes of fictional patient cases

### 4.5 Evaluation

When prompting the enhanced and baseline models, we will provide them with the complete set of correct treatment taken from all patients. The models will then be instructed to select the appropriate treatment from the given set for the presented case. This process will then be repeated for all patient cases. We chose this approach for evaluation because it simplifies the task by relying on a controlled set of treatment options and therefore allows automated assessment of model responses. Furthermore, since we base our work on established clinical guidelines, the complete set of valid treatment recommendations is already defined. This removes the need for the model to generate treatments from scratch. Figure 6 shows a diagram for the evaluation process.

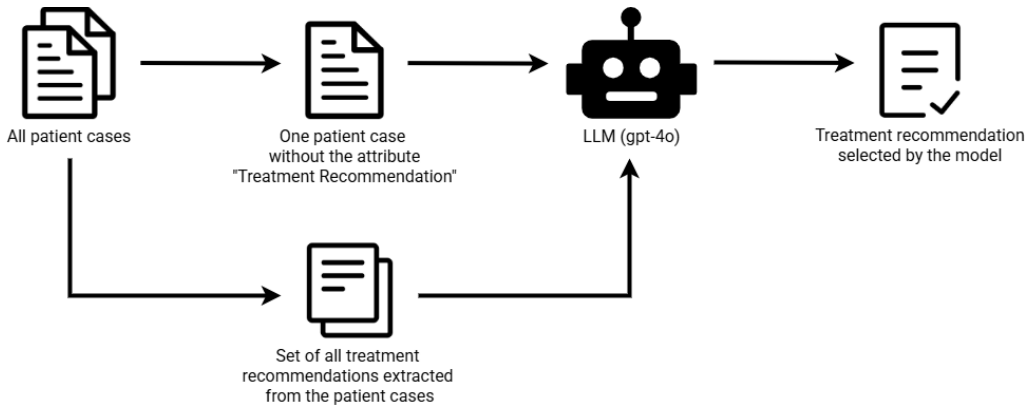


Figure 6: Diagram for the evaluation process

We will use two additional models as baselines to evaluate the enhanced model. We will query these models and compare the enhanced model's performance with theirs to assess whether it leads to any improvements in the treatment recommendations. The first baseline model will be a standard language model without any additional capabilities. The second will only incorporate standard RAG to see whether including ReAct provides additional benefits.

### 4.5.1 Metrics

We will use accuracy to evaluate the correctness of the treatment recommendations selected by the models for patient cases that fall within the scope of the clinical guidelines. For the treatment selection task, we define the accuracy as:

$$Accuracy = \frac{\text{Number of correct selections}}{\text{Total number of cases}}$$

To evaluate the classification performance of the model, we will also use accuracy, which is defined as:

$$Accuracy_C = \frac{\text{Number of correctly classified cases}}{\text{Total number of cases}}$$

# References


[1] Dandan Wang and Shiqing Zhang. Large language models in medical and healthcare fields: applications, advances, and challenges. *Artificial Intelligence Review*, 57(11):299, September 2024.

[2] Haritha Atluri and Bala Siva Prakash Thummisetti. Enhancing Antibiotic Prescribing in Urgent Care by Leveraging Large Language Models for Optimized Clinical Decision Support. *International Research Journal of Modernization in Engineering Technology and Science*, January 2024.

[3] Jasmine Chiat Ling Ong, Liyuan Jin, Kabilan Elangovan, Gilbert Yong San Lim, Daniel Yan Zheng Lim, Gerald Gui Ren Sng, Yuhe Ke, Joshua Yi Min Tung, Ryan Jian Zhong, Christopher Ming Yao Koh, Keane Zhi Hao Lee, Xiang Chen, Jack Kian Chng, Aung Than, Ken Junyang Goh, and Daniel Shu Wei Ting. Development and Testing of a Novel Large Language Model-Based Clinical Decision Support Systems for Medication Safety in 12 Clinical Specialties, February 2024. arXiv:2402.01741 [cs].

[4] Siru Liu, Aileen P Wright, Barron L Patterson, Jonathan P Wanderer, Robert W Turer, Scott D Nelson, Allison B McCoy, Dean F Sittig, and Adam Wright. Using AI-generated suggestions from ChatGPT to optimize clinical decision support. *Journal of the American Medical Informatics Association*, 30:1237–1245, July 2023.

[5] Shan Chen, Benjamin H. Kann, Michael B. Foote, Hugo J. W. L. Aerts, Guergana K. Savova, Raymond H. Mak, and Danielle S. Bitterman. Use of Artificial Intelligence Chatbots for Cancer Treatment Information. *JAMA Oncology*, 9(10):1459–1462, October 2023.

[6] Kimberly H. Allison and George W. Sledge. Heterogeneity and cancer. *Oncology*, 28(9):772–772, September 2014.

[7] Mario Lamping, Manuela Benary, Serge Leyvraz, Clemens Messerschmidt, Eric Blanc, Thomas Kessler, Moritz Schütte, Dido Lenze, Korinna Jöhrens, Susen Burock, Konrad Klinghammer, Sebastian Ochsenreither, Christine Sers, Reinhold Schäfer, Ingeborg Tinhofer, Dieter Beule, Frederick Klauschen, Marie-Laure Yaspo, Ulrich Keilholz, and Damian T. Rieke. Support of a molecular tumour board by an evidence-based decision management system for precision oncology. *European Journal of Cancer*, 127:41–51, March 2020.

[8] Benjamin M. Good, Benjamin J. Ainscough, Josh F. McMichael, Andrew I. Su, and Obi L. Griffith. Organizing knowledge to enable personalization of medicine in cancer. *Genome Biology*, 15(8):438, August 2014.

[9] Vicente J. Bermejo, Andres Gago, Ramiro H. Gálvez, and Nicolás Harari. LLMs Outperform Outsourced Human Coders on Complex Textual Analysis, November 2024.

[10] Vera Sorin, Eyal Klang, Miri Sklair-Levy, Israel Cohen, Douglas B. Zippel, Nora Balint Lahat, Eli Konen, and Yiftach Barash. Large language model (ChatGPT) as a support tool for breast tumor board. *npj Breast Cancer*, 9(1):1–4, May 2023.

[11] Shan Zhou, Xiao Luo, Chan Chen, Hong Jiang, Chun Yang, Guanghui Ran, Juan Yu, and Chengliang Yin. The performance of large language model powered chatbots compared to oncology physicians on colorectal cancer queries. *International Journal of Surgery*, page 10.1097/JS9.0000000000001850, 2024.

[12] Manuela Benary, Xing David Wang, Max Schmidt, Dominik Soll, Georg Hilfenhaus, Mani Nassir, Christian Sigler, Maren Knödler, Ulrich Keller, Dieter Beule, Ulrich Keilholz, Ulf Leser, and Damian T. Rieke. Leveraging Large Language Models for Decision Support in Personalized Oncology. *JAMA Network Open*, 6(11):e2343689, November 2023.

[13] Florian Putz, Marlen Haderlein, Sebastian Lettmaier, Sabine Semrau, Rainer Fietkau, and Yixing Huang. Exploring the Capabilities and Limitations of Large Language Models for Radiation Oncology Decision Support. *International Journal of Radiation Oncology*Biology*Physics*, 118(4):900–904, March 2024.

[14] Dyke Ferber and Jakob Nikolas Kather. Large Language Models in Uro-oncology. *European Urology Oncology*, 7(1):157–159, February 2024.

[15] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, Sebastian Riedel, and Douwe Kiela. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks, April 2021.

[16] Tolga Şakar and Hakan Emekci. Maximizing RAG efficiency: A comparative analysis of RAG methods. *Natural Language Processing*, 31(1):1–25, January 2025.

[17] Dyke Ferber, Isabella C. Wiest, Georg Wölflein, Matthias P. Ebert, Gernot Beutel, Jan-Niklas Eckardt, Daniel Truhn, Christoph Springfeld, Dirk Jäger, and Jakob Nikolas Kather. GPT-4 for Information Retrieval and Comparison of Medical Oncology Guidelines. *NEJM AI*, 1(6):AIcs2300235, May 2024.

[18] Zine-Eddine Khene, Pierre Bigot, Romain Mathieu, Morgan Rouprêt, and Karim Bensalah. Development of a Personalized Chat Model Based on the European Association of Urology Oncology Guidelines: Harnessing the Power of Generative Artificial Intelligence in Clinical Practice. *European Urology Oncology*, 7(1):160–162, February 2024.

[19] Shunyu Yao, Jeffrey Zhao, Dian Yu, Nan Du, Izhak Shafran, Karthik Narasimhan, and Yuan Cao. ReAct: Synergizing Reasoning and Acting in Language Models, March 2023. arXiv:2210.03629 [cs].

[20] Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed Chi, Quoc Le, and Denny Zhou. Chain-of-Thought Prompting Elicits Reasoning in Large Language Models, January 2023.

[21] Meta. Introducing Meta Llama 3: The most capable openly available LLM to date. https://ai.meta.com/blog/meta-llama-3/. Accessed: 2025-05-06.

[22] LlamaIndex. LlamaIndex. https://docs.llamaindex.ai/en/stable/. Accessed: 2025-04-17.

[23] Deutsche Gesellschaft für Hämatologie und Medizinische Onkologie e.V. Onkopedia Gudeilines. https://www.onkopedia.com/de/onkopedia/guidelines. Accessed: 2025-04-29.

[24] Ollama. Ollama. https://ollama.com/. Accessed: 2025-05-06.

[25] OpenAI. Openai. https://openai.com/. Accessed: 2025-07-05.

[26] Onkopedia. Small-Cell Lung Cancer (SCLC). https://www.onkopedia-guidelines.info/en/onkopedia/guidelines/small-cell-lung-cancer-sclc/@@guideline/html/index.html, January 2023. Accessed: 2025-07-14.

[27] Onkopedia. Gastric Cancer. https://www.onkopedia-guidelines.info/en/onkopedia/guidelines/gastric-cancer/@@guideline/html/index.html, February 2025. Accessed: 2025-07-14.

[28] Onkopedia. Ovarian Cancer. https://www.onkopedia-guidelines.info/en/onkopedia/guidelines/ovarian-cancer/@@guideline/html/index.html, July 2023. Accessed: 2025-07-14.

[29] Onkopedia. Renal Cell Carcinoma (Hypernephroma). https://www.onkopedia-guidelines.info/en/onkopedia/guidelines/renal-cell-carcinoma-hypernephroma/@@guideline/html/index.html, February 2024. Accessed: 2025-07-14.

[30] Onkopedia. Gastrointestinal stromal tumors (GIST). https://www.onkopedia-guidelines.info/en/onkopedia/guidelines/gastrointestinal-stromal-tumors-gist/@@guideline/html/index.html, April 2025. Accessed: 2025-07-14.

[31] Charité. Charité. https://www.charite.de/. Accessed: 2025-07-10.

[32] National Cancer Institute (NIH). Surveillance, Epidemiology, and End Results (SEER). https://seer.cancer.gov/. Accessed: 2025-07-04.