
CLINICAL DECISION SUPPORT BASED ON BEVACIZUMAB CANCER TRIALS AND PUSHING THE LIMITATIONS OF ADVANCED LLMs

Kevin Kawchak 

Chief Executive Officer

ChemicalQDevice

San Diego, CA

January 27, 2025

kevink@chemicalqdevice.com

ABSTRACT

An exhaustive study was needed to test the limits of leading Large Language Models (LLMs) using numerous real-world clinical trial outcomes. It was also necessary to provide extensive hallucination studies based on both extracting main points and providing novel AI clinical decision support. Here, 100 Bevacizumab cancer therapy articles representing over 900K words were summarized by the 3.5 Sonnet model into 49K words, which completed a detailed and complex problem of several cancer and study types to press the capabilities of the ChatGPT o1 reasoning model. Report summaries in general followed an effective format, with guardrails to de-identify patient information and numerical data sources attributions to ground the output to the input. The main takeaway was that both LLMs typically remained in-context with the input data when more structured prompts were used, while precise quotations and author name citations were less prominent. These errors were likely due to LLM pressure towards achieving coherence vs. exact recall based on manufacturer inference-time compute settings. Overall, ChatGPT o1 provided state-of-the-art evidence-based Bevacizumab insights regarding clinical efficacy across indications, dosing recommendations, combination therapies, and biomarker-driven selections.

Keywords Clinical decision support · Bevacizumab clinical trials · Cancer · LLMs

1 Introduction



Figure 1: Two part LLM study

2025 LLM clinical decision support systems (CDSS) have pressed forward with application improvements based on newer AI models. In 2025, Buhr C. et al. provided decision-making assessments with local and web-based LLMs versus human board recommendations. ChatGPT 4o yielded a mean score of 4.7 (IQR: 4–6), while Llama 3 [likely the 8B parameter model] achieved a score of 4.3 (IQR: 3–5) [1]. Kim B. et al. proposed a CDSS for mental health diagnosis that utilized GPT 4o, Gemini 1.5 Flash, and a Llama 3.2 model with constraint logic programming to translate diagnostic manuals to a logic program. GPT 4o had "9 correct diagnoses out of 10, followed by Gemini with 8 and Llama with 7" [2]. In addition, Ensina L. et al. utilized GPT 4 Turbo, 4, and 3.5 to analyze 969 electronic medical records in Brazilian and Portuguese regarding the detection of anaphylaxis. GPT 4 Turbo had the best performance, detecting anaphylaxis cases "with 90.6% precision, 100% sensitivity, 99.5% specificity, 99.5% accuracy, and a Cohen kappa coefficient of 0.95 [3]."

Gao Y., et al. assessed instruction-tuned versions of Mistral 7B and Llama 3 70B models in predicting binary outcomes for sepsis, arrhythmia, and congestive heart failure; however a XGB classifier outperformed the smaller sized LLMs across all tasks [4]. In a *Nature Communications* article by Griot M et al. titled "Large Language Models lack essential metacognition for reliable medical reasoning", the authors found that GPT 4o, GPT 3.5 turbo, and several open source models typically did not possess "a robust capacity to provide higher accuracy". GPT 4o was found to be the best performing model, with a 'missing answer recall metric' of $M = 46.1\%$, $SEM = 4.7\%$, which was significantly more accurate ($p < 0.0001$) than GPT 3.5 turbo's score of $M = 11.3\%$, $SEM = 2.9\%$ [5]. Steybe D. et al. used a retrieval-augmented generation (RAG) GPT 4 system "to provide answers on the prevention, diagnosis, and treatment of medication-related osteonecrosis of the jaw." Their GuideGPT framework using the Mann–Whitney U test showed significantly better ratings for GuideGPT than without using RAG "regarding content ($p = 0.006$), scientific explanation ($p = 0.032$), and agreement ($p = 0.008$), though not for language ($p = 0.407$) [6]."

Chan T. et al. in January 7, 2025 utilized a verbal reinforcement learning framework with AI agents to improve on an agentless Llama 3.1 score from 70% to 81.3%. ChatGPT 4 scored 79.3% regarding complex knowledge otolaryngology. Their ENTAgents were found to "improve accuracy by over 11.3% in handling short questions, with 2.78 folds in the length of the text and explaining clearly [7]." Of particular importance was the Noda R. et al. 2025 study using o1 pro vs. GPT 4 regarding self-assessment questions from 2014–2023 for nephrology board renewal. "o1 pro scored 81.3% (170/209), significantly higher than GPT-4's 51.2% (107/209; $p < 0.001$)", with o1 pro exceeding the 60% passing criterion every year [8].

A review by Liu S. et al. found that LLM-RAG workflows "showed a 1.35 odds ratio increase in performance compared to baseline LLMs", and developed guidelines for improving clinical applications [9]. Moreover, Antonie N. et al. conducted a review regarding the use of ChatGPT and other AI chatbots to optimize antibiotic therapy, with the main finding that AI has "the potential to complement antibiotic stewardship efforts"; however rigorous clinical trials, interdisciplinary collaboration, regulatory clarity, and tailored algorithmic improvements are needed. Also, Gallifant J. et al. in January 2025 proposed a reporting guideline called TRIPOD-LLM as an extension to a prior work regarding biomedical applications aimed towards assisting academia, journal editors and peer reviewers, policymakers, funders, regulators, patients and study participants [10]. Furthermore, Malgaroli M. et al. published revised LLM guidelines for mental health regarding the building of a global clinical repository, the design of ethical usage standards, and offering further refinements and integrations while promoting digital inclusivity [11].

Yuan Z. et al. in May 2024 stated of LLMs: "As a result of the vast size of these models, they present considerable inference challenges, not only for devices with limited computational capabilities, but also for the most advanced hardware. Because of their complexity and scale, as well as their energy and computational demands, these models are difficult to deploy in real-world situations." "Additionally, the resource-intensive nature of these models raises concerns about energy consumption, scalability, and accessibility [12]." Li R. et al. in November 2024 discussed the importance of using improved processing techniques such as GPU tensor parallelism and load balancing computational tasks across devices to prevent device overloading [13]. The OpenAI GPT-4o system card detailed emerging features using the Apollo Research assessment: achieving scores of self-knowledge (2/3), explicit theory of mind (3/3), self-reasoning (1/3), applied theory of mind (1/3), and combined self-reasoning and theory of mind (1/3) [14]. Subsequently, the December 2024 GPT-o1 system card further highlighted the ongoing progress of reasoning models: "It is likely that improved scaffolding, inference-time compute techniques, or finetuning can further improve the performance of o1 [15]."

Nazi Z. et al. in January 2025 evaluated open and closed sourced LLMs with zero-shot, few-shot, and chain-of-thought prompting Chain-of-thought prompting (CoT), which LLMs generate a structured, step-by-step reasoning process before arriving at the final answer. However, their CoT experiments revealed that longer prompt lengths in CoT prompting can sometimes lead to increased hallucination where the model generates irrelevant or incorrect outputs [16]. Xu F. et al. in 2025 detailed process reward models that can be utilized during test-time phase, such as OpenAI o1 series models. "The new test-time scaling laws suggest that inference capabilities can be effectively enhanced by increasing test-time compute, providing a clear direction for the future development of LLMs." This allows for o1's reasoning capabilities which can "be effectively enhanced through advanced computation allocation during inference, particularly in complex problemsolving scenarios." "The constraints on scaling this approach differ substantially from those of LLM pretraining, with performance consistently improving with more time spent thinking [17]."

Agarwal M. et al. in October 2023 highlighted best practices for LLM inference performance in which input length is not generally significant for performance but important for hardware requirements, and that slower large model speed ratios won't necessarily match the parameter count ratio [18]. Shenoy V. et al. detailed the importance of the attention algorithm used in models such as ChatGPT, in which loading values from memory, performing a computation, and storing the results of that computation to memory are all affected by the algorithm [19]. Ferraris A. et al. published an article in January 2025 regarding self-attention mechanism which requires significant memory due to the requirement of "comparing every element in the input sequence with every other element, which scales quadratically with the input length." Therefore, with "very large datasets or long input sequences, this can become prohibitively expensive. [20]"

Hua S. et al. in February 2024 detailed top-p decoding in the training and the inference process may cause models such as ChatGPT to hallucinate. In addition, "the exposure bias problem arising from the inconsistency between training target and inferred target is also a major cause of inaccurate text." In addition, parametric knowledge bias has been found to cause extrinsic hallucinations [21]. Zhou Z. et al. in July 2024 addressed problems regarding batch inference and KV caching by utilizing a skip decoder to "ensures a monotonic decrease in exit points to prevent recomputation of KV cache, facilitating efficiency gains during inference [22]." Moreover, Li B. et al. in July 2024 detailed the complexities of large inputs: "Serving long-context LLM applications is particularly challenging as the size of the KV cache scales with the number of tokens." In addition LLM inference by generating text token by token is primarily limited by memory bandwidth [23].

2 Methods

AI software used in this study were unmodified LLMs with generation times being obtained from a digital stop watch. 3.5 Sonnet utilized document retrieval of full article pdfs found only in Table 2 using the 'paper clip' option and text prompts, while ChatGPT o1 processed text prompts plus large inputs. Each of the 100 reports for Part Two were combined and formatted with the type of study, the article title, the authors' names, and the date as shown in Supplementary S100p. The total input size including

Prompt 2a provided to ChatGPT o1 was 49,041 words (estimated number of tokens = 63,753, at 1.3 tokens/word), and 340,654 characters. Article lengths were obtained by pasting pdf text into Google Docs and using the word counts feature.

ChatGPT o1 (128K token limit through "Pro" service, but not the ChatGPT o1 pro model) prompts were first optimized with several larger text experiments. Once the output quality was sufficient, a chain of prompts were run in a single conversation. Text that was copied into LLM input fields was also pasted as plain text into the 169 page Supplementary S100p, available to developers and clinicians as a single reproducibility experiment. Supplementary V500 contains the 500 Part One verifications, SREA contains the Part Two generations, and VP11 contains Part Two verifications. Generations in the manuscript received white space formatting for readability and fitting. The author conducted experiments, analysis, and wrote the manuscript, with ChatGPT 4o and ChatGPT o1 primarily being utilized as a research tool.

2.1 LLMs in Study

- 01) 3.5 Sonnet: Claude website chat interface was accessed through MacOS 14.5 (23F79) and Google Chrome browser Version 131.0.6778.205 (Official Build) (arm64) [24].
- 02) ChatGPT o1: ChatGPT website pro interface was accessed through MacOS 14.5 (23F79) and Google Chrome browser Version 131.0.6778.205 (Official Build) (arm64) [25]. The ChatGPT o1 pro model was not utilized in this study.
- 03) ChatGPT 4o: ChatGPT website pro interface was accessed through MacOS 14.5 (23F79) and Google Chrome browser Version 131.0.6778.205 (Official Build) (arm64) [26].

3 Part One: 100 Articles

A PRISMA literature search based on 17 Bevacizumab related Google Scholar [27] searches screening 1,225 entries was conducted to obtain the 100 articles, as shown in Appendix Figure 5. Records excluded were based on license type, therapy type, and relevancy to yield 23 prospective clinical trials, 26 retrospective clinical studies, 10 case reports, 9 preclinical studies, and 32 review articles. Of these works, there were 25 colorectal cancer articles, 19 non-small cell lung cancer, 15 ovarian cancer, 15 hepatocellular carcinoma, 9 glioblastoma tumor, 5 cervical cancer, 4 glioma tumor, 2 breast cancer, 1 neuroblastoma, 1 appendiceal adenocarcinoma, 1 renal cell carcinoma, 1 PASC, and 2 various cancer articles. The resulting 100 article table separated by number, authors, 2024 published date online, words, and journal is shown in Table 2. The sum of the 100 articles was 914,257 words, with an average article length of 9,142.57 words, and a standard deviation of 4,119.72.

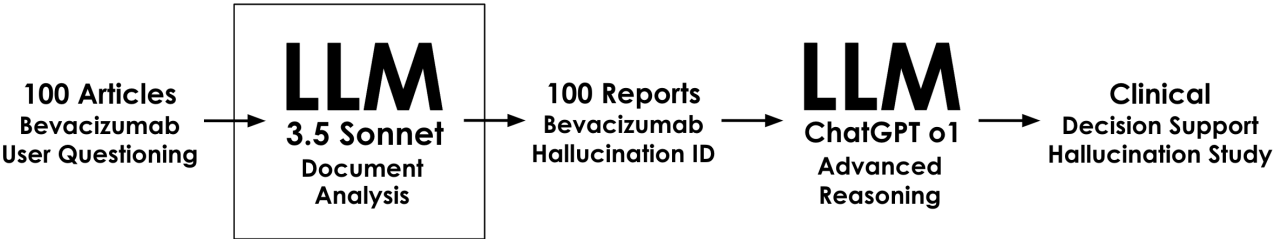


Figure 2: Part One LLM study

Standard 1: Bevacizumab Therapy Reports					
<p>Prompt 1: Write a two page report based on the attached document regarding the authors’ Bevacizumab therapy results, also abbreviated as BEV, BVZ, Bmab, BCZ, or BV using the following format: “Executive Summary”, “Technical Details”, “Key Insights”. Only include results related to Bevacizumab, deidentifying any patient information if present. Cite selections of numerical data enclosed in quotation marks corresponding to the authors’ names. Only use exact quotations derived from the attached document throughout the report. [Single Article Attached, Single Generation]</p>					
1 of 100 Articles	Bevacizumab Therapy	De-identify Patient IDs	Data Citations Author Names	Detailed Report Format	1 of 100 Reports

Table 1: Process diagram for Standard 1

Bevacizumab Therapy Articles 001-050			Bevacizumab Therapy Articles 051-100		
Article Authors	Date Words	Journal	Article Authors	Date Words	Journal
001 Boland P. et al. ^[28]	12/27/24 5488	<i>Invest New Drugs</i>	051 Boso D. et al. ^[29]	12/19/24 10587	<i>J Exp Clin Cancer Res</i>
002 Pan Q. et al. ^[30]	04/03/24 8476	<i>Sig Transduct Target Ther</i>	052 Alemzadeh E. et al. ^[31]	04/23/24 12995	<i>Oncol Res.</i>
003 Seufferlein T. et al. ^[32]	06/14/24 7666	<i>PLoS One</i>	053 Zak K. et al. ^[33]	03/05/24 14752	<i>Cancers</i>
004 Chen X. et al. ^[34]	08/12/24 9810	<i>Front. Oncol.</i>	054 He H. et al. ^[35]	11/21/24 11084	<i>Front. Pharmacol.</i>
005 Rietveld P. et al. ^[36]	05/28/24 4818	<i>Biomed. Pharmacother.</i>	055 Kim Y. et al. ^[37]	08/21/24 7283	<i>Pharmaceuticals</i>
006 Okawa M. et al. ^[38]	12/19/24 8875	<i>PLoS One</i>	056 Romero I. et al. ^[39]	01/29/24 7308	<i>Front. Oncol.</i>
007 Moisuc D. et al. ^[40]	01/16/24 10683	<i>Cancers</i>	057 Luvero D. et al. ^[41]	06/14/24 7536	<i>Pharmaceuticals</i>
008 Zhang P. et al. ^[42]	10/09/24 7735	<i>Front. Immunol.</i>	058 Nunes M. et al. ^[43]	05/04/24 24456	<i>Cells</i>
009 Li S. et al. ^[44]	12/11/24 5429	<i>Sig Transduct Target Ther</i>	059 Han Gr. et al. ^[45]	08/29/24 14971	<i>Clin. Transl. Med.</i>
010 Chautard R. et al. ^[46]	08/20/24 6643	<i>Dig Liver Dis</i>	060 Murata Y. et al. ^[47]	06/29/24 7391	<i>Cancers</i>
011 Zheng Z. et al. ^[48]	01/25/24 5511	<i>Front. Pharmacol.</i>	061 Nosaka T. et al. ^[49]	05/06/24 9397	<i>Cancers</i>
012 Furuya K. et al. ^[50]	02/02/24 8541	<i>BMC Cancer</i>	062 Shomura M. et al. ^[51]	10/25/24 8082	<i>Cancers</i>
013 Mazard T. et al. ^[52]	06/17/24 8202	<i>Cancer Imaging</i>	063 Hwang S. et al. ^[53]	02/19/24 10954	<i>Cancers</i>
014 He L. et al. ^[54]	03/28/24 5526	<i>Front. Immunol.</i>	064 Storandt M. et al. ^[55]	07/09/24 7588	<i>Oncol</i>
015 Yan L. et al. ^[56]	07/20/24 25901	<i>Discov Onc</i>	065 Hu Z. et al. ^[57]	11/19/24 6830	<i>Front. Immunol.</i>
016 Sun C. et al. ^[58]	12/23/24 10912	<i>PLoS One</i>	066 Khaled N. et al. ^[59]	04/08/24 10121	<i>JHEP Reports</i>
017 Song Y. et al. ^[60]	02/01/24 7951	<i>BMC Gastroenterol</i>	067 Ueno M. et al. ^[61]	09/18/24 6917	<i>J Gastroenterol</i>
018 André T. et al. ^[62]	09/26/24 7532	<i>Curr. Treat. Options in Oncol.</i>	068 Shirane Y. et al. ^[63]	03/28/24 9874	<i>Cancers</i>
019 Corrias G. et al. ^[64]	03/30/24 12471	<i>Cancers</i>	069 Egerer M. et al. ^[65]	10/29/24 7434	<i>Cancers</i>
020 Jacobsen A. et al. ^[66]	02/22/24 12801	<i>Cancers</i>	070 Abraham S. et al. ^[67]	01/04/24 6241	<i>Front. Oncol.</i>
021 Naz T. et al. ^[68]	06/06/24 7795	<i>J. Pharm. Policy Pract.</i>	071 Krupa K. et al. ^[69]	12/04/24 15001	<i>Cancers</i>
022 Rais T. et al. ^[70]	07/12/24 8863	<i>Front. Oncol.</i>	072 Oura K. et al. ^[71]	12/19/24 12732	<i>Int. J. Mol. Sci.</i>
023 De Felice F. et al. ^[72]	11/20/24 4047	<i>Clin. Oncol.</i>	073 Childs A. et al. ^[73]	06/08/24 12227	<i>JHEP Reports</i>
024 Su M. et al. ^[74]	07/12/24 4282	<i>Medicine</i>	074 Chen K. et al. ^[75]	05/17/24 10549	<i>Br J Cancer</i>
025 Xu Q. et al. ^[76]	12/03/24 14639	<i>Front. Immunol.</i>	075 Valerio J. et al. ^[77]	10/10/24 9398	<i>J. Pers. Med.</i>
026 Wu S. et al. ^[78]	12/23/24 10880	<i>Clin. Transl. Med.</i>	076 Guo G. et al. ^[79]	08/06/24 6756	<i>CII</i>
027 Huang Y. et al. ^[80]	01/29/24 6740	<i>Sig Transduct Target Ther</i>	077 Zhang J. et al. ^[81]	01/04/24 7916	<i>J Neurooncol</i>
028 Ge Y. et al. ^[82]	02/01/24 5688	<i>Cancer Med.</i>	078 Ezaki T. et al. ^[83]	04/15/24 5232	<i>Brain Tumor Pathol</i>
029 Rathbone M. et al. ^[84]	03/22/24 11756	<i>Cancers</i>	079 Zhong W. et al. ^[85]	09/04/24 4919	<i>Front. Pharmacol.</i>
030 Zhang X. et al. ^[86]	10/15/24 7061	<i>Front. Oncol.</i>	080 Laviv Y. et al. ^[87]	09/24/24 6796	<i>J Neurooncol</i>
031 Liao X. et al. ^[88]	10/15/24 5873	<i>CRJ</i>	081 Rahman M. et al. ^[89]	08/27/24 15882	<i>Cancers</i>
032 Kuo C. et al. ^[90]	03/25/24 6997	<i>Kaohsiung J Med Sci.</i>	082 Hoosemans L. et al. ^[91]	08/29/24 16624	<i>Cancers</i>
033 Xiong J. et al. ^[92]	06/24/24 4427	<i>Front. Oncol.</i>	083 Sadowski K. et al. ^[93]	05/26/24 19231	<i>Int. J. Mol. Sci.</i>
034 Hua C. et al. ^[94]	06/23/24 13889	<i>Biomolecules</i>	084 Yang H. et al. ^[95]	10/17/24 6761	<i>Biomol Biomed</i>
035 Lin X. et al. ^[96]	01/27/24 8087	<i>Cancer Immunol Immunother</i>	085 Ju M. et al. ^[97]	10/09/24 6660	<i>Cancer Med.</i>
036 Jarry U. et al. ^[98]	06/27/24 8754	<i>PLoS One</i>	086 Adachi T. et al. ^[99]	11/11/24 3302	<i>J Dermatol</i>
037 Huang S. et al. ^[100]	10/01/24 8485	<i>Exp Hematol Oncol</i>	087 Zhai C. et al. ^[101]	01/08/24 9862	<i>BMC Geriatr</i>
038 Kobayashi N. et al. ^[102]	12/19/24 8112	<i>Genes</i>	088 Chang C. et al. ^[103]	09/10/24 13680	<i>Taiwan J Obstet Gynecol.</i>
039 Tian J. et al. ^[104]	12/04/24 6177	<i>CRJ</i>	089 Zhao S. et al. ^[105]	01/02/24 7179	<i>BMC Cancer</i>
040 Han Gu. et al. ^[106]	01/22/24 8118	<i>Front. Pharmacol.</i>	090 Shen L. et al. ^[107]	08/19/24 11189	<i>Cell Prolif</i>
041 He Q. et al. ^[108]	12/16/24 10769	<i>Cancer Biol Ther</i>	091 Wei S. et al. ^[109]	04/02/24 5283	<i>Front. Med.</i>
042 Zhao Y. et al. ^[110]	02/22/24 7218	<i>Biomed. Pharmacother.</i>	092 Wang C. et al. ^[111]	11/14/24 6826	<i>World Neurosurg.</i>
043 Zhou Y. et al. ^[112]	03/13/24 5170	<i>Clin Transl Oncol</i>	093 Fleischer T. et al. ^[113]	04/26/24 12395	<i>Mol. Oncol.</i>
044 Li Q. et al. ^[114]	07/04/24 12263	<i>Front. Immunol.</i>	094 Janes P. et al. ^[115]	05/16/24 10305	<i>Cancers</i>
045 Freyer G. et al. ^[116]	03/05/24 8826	<i>Nat. Commun.</i>	095 Moreno L. et al. ^[117]	01/08/24 9479	<i>J Clin Oncol</i>
046 Roque D. et al. ^[118]	06/20/24 5475	<i>BJC Reports</i>	096 Hornstein N. et al. ^[119]	05/29/24 4197	<i>CRC</i>
047 Rosario S. et al. ^[120]	12/05/24 11995	<i>Nat. Commun.</i>	097 Eismann L. et al. ^[121]	10/02/24 3000	<i>Eur Urol</i>
048 Kokabu T. et al. ^[122]	07/25/24 6860	<i>Cancers</i>	098 Liu Q. et al. ^[123]	12/12/54 4325	<i>Front. Immunol.</i>
049 Kanno K. et al. ^[124]	08/25/24 6119	<i>Curr. Issues Mol. Biol</i>	099 Wang L. et al. ^[125]	01/02/24 9330	<i>J Pharm Health Care Sci</i>
050 Iida Y. et al. ^[126]	02/14/24 7314	<i>Front. Oncol.</i>	100 Wei Q. et al. ^[127]	01/04/24 21770	<i>J Hematol Oncol</i>

Sum = 914,257 words; Avg = 9,142.57 words; St Dev = 4,119.72

Table 2: 100 Articles used for generating 100 LLM reports

100 LLM Generated Reports: Word Counts by Slice, and Response Times

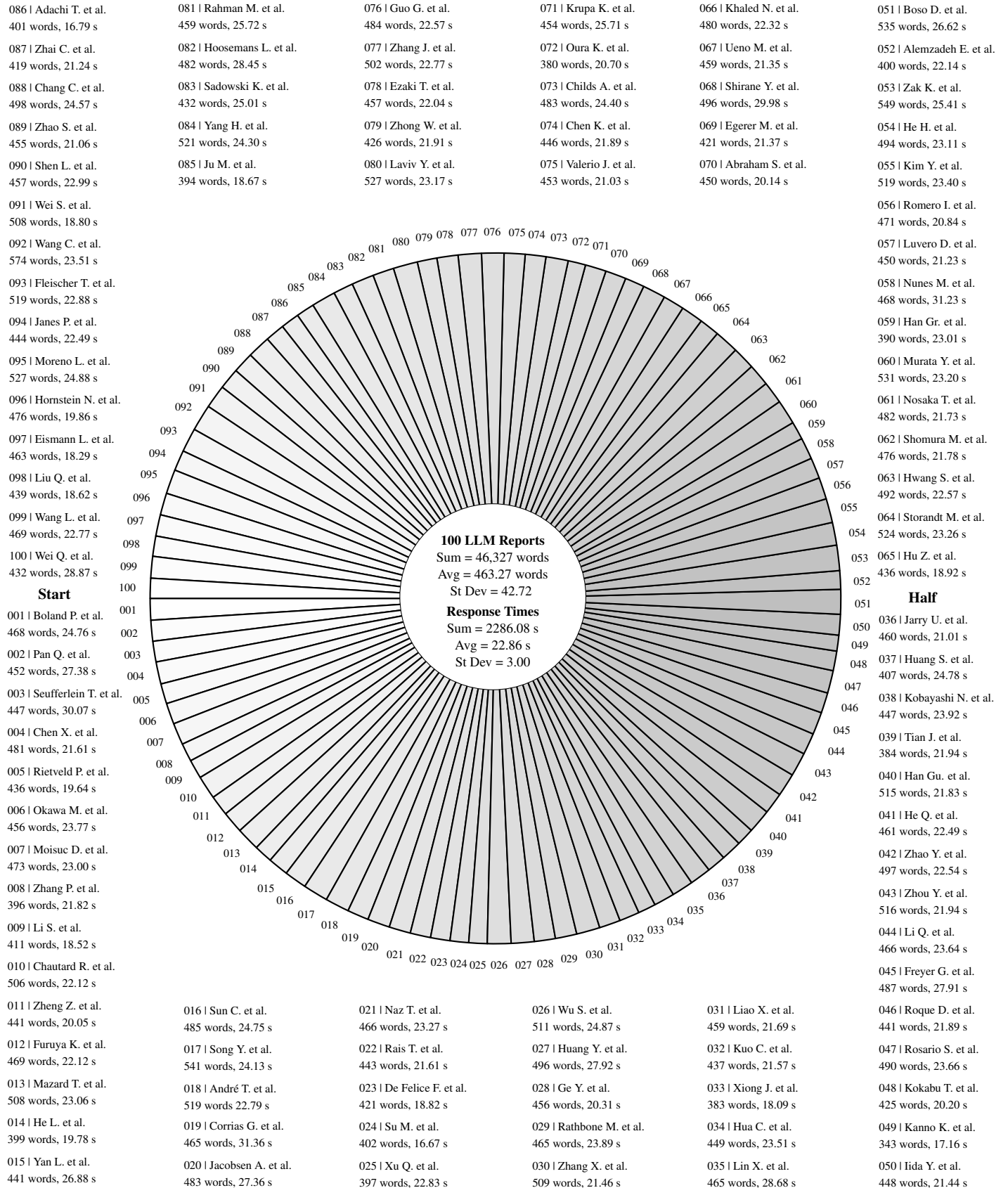


Figure 3: LLM 3.5 Sonnet summaries. Formatted size = 49,041 words, 169 pages in Supplementary S100p

LLM Detailed Report 023 of 100: De Felice F. et al. 2024. 421 words, 18.82 seconds

Here is a two-page report focused on the bevacizumab therapy results from the study:

Bevacizumab in Total Neoadjuvant Therapy for Locally Advanced Rectal Cancer: Analysis Report

Executive Summary

This report analyzes the bevacizumab (BEV) outcomes from a phase II clinical trial examining intensified total neoadjuvant therapy (TNT) for locally advanced rectal cancer (LARC). The study by De Felice et al. evaluated TNT consisting of targeted therapy (including bevacizumab for RAS-BRAF mutated patients) combined with FOLFOXIRI chemotherapy followed by intensified chemoradiotherapy. While specific bevacizumab-only outcomes were not separated in the results, the overall TNT regimen including bevacizumab demonstrated promising complete response rates and survival outcomes over a median 6.3 year follow-up period.

Technical Details

The study design included bevacizumab as follows:

- Patient Selection: Bevacizumab was administered specifically to patients with "mutated RAS-BRAF" status, as noted by De Felice et al.
- Treatment Protocol: The regimen consisted of "four cycles of FOLFOXIRI plus bevacizumab (mutated RAS-BRAF)" as described in the methods section.
- Timing: Bevacizumab was given during the induction chemotherapy phase, which occurred before chemoradiotherapy.

Key Outcome Metrics:

For the overall TNT regimen including bevacizumab for eligible patients:

- Complete Response Rate: De Felice et al. report "11 (39.3%) patients had a complete response (CR)"
- Survival Outcomes: The authors note "5-year overall survival (OS) and DFS were 74.6% and 57.1%, respectively"
- Follow-up Duration: As stated by De Felice et al., there was a "6.3 years (median follow-up)"

Key Insights

1. Integration with Standard Therapy:

The study demonstrated that bevacizumab could be successfully incorporated into an intensified TNT regimen for LARC patients with specific molecular profiles (RAS-BRAF mutations).

2. Safety Profile:

While bevacizumab-specific adverse events were not separately reported, the overall regimen showed manageable toxicity. As noted by De Felice et al., "There was no evidence of severe long-term toxicities."

3. Clinical Implications:

The authors conclude that "The addition of a targeted agent to induction FOLFOXIRI and oxaliplatin to 5-fluorouracil-based CRT, with the doses and intensities used in this study, resulted in high CR rates."

4. Research Limitations:

The authors acknowledge that "the results of our prospective phase II study should be interpreted with caution and do not unequivocally support the hypothesis that adding bevacizumab or panitumumab/cetuximab to the induction FOLFOXIRI and OXP to neoadjuvant 5-FU-based CRT substantially improves survival outcomes."

5. Future Directions:

As stated by De Felice et al., "Further research on TNT strategies in LARC is encouraged," suggesting the need for larger studies to validate these findings and better define the role of bevacizumab in TNT regimens.

LLM Detailed Report 086 of 100: Adachi T. et al. 2024. 401 words, 16.79 seconds

Here is a two-page report focused on the Bevacizumab therapy findings from the document:

Report on Bevacizumab Therapy Results from Adachi et al. Study

Executive Summary

This report analyzes the Bevacizumab (BEV) therapy outcomes from a case series examining cutaneous adverse events during pembrolizumab combination chemotherapy for cervical cancer. The study, conducted by Adachi et al. between October 2022 and May 2023, included patients receiving BEV as part of their treatment regimen. The findings provide insights into the safety profile and management of BEV when used in combination therapy for cervical cancer patients.

Technical Details

The treatment protocol included Bevacizumab as part of a combination therapy regimen. As described in Figure 1a of the paper, BEV was administered following paclitaxel, cisplatin/carboplatin, and pembrolizumab.

Out of the total 19 patients studied, two cases specifically involved Bevacizumab in their combination therapy:

Case 1 and Case 4 received what Adachi et al. describe as "paclitaxel + cisplatin + bevacizumab + PEM" combination therapy.

The study timeline covered a period quoted as "between October 2022 and May 2023" during which these treatments were administered.

The authors note that this combination represents their standard approach, describing it as "the conventional paclitaxel + cisplatin (or carboplatin) ± bevacizumab regimen."

Key Insights

1. Safety Profile:

- When Bevacizumab was included in the combination therapy, it was generally well-tolerated as part of the overall treatment protocol.
- No adverse events were specifically attributed to Bevacizumab alone in the reported cases.

2. Treatment Outcomes:

- For Case 1, which included Bevacizumab, the authors report the "effect of the treatment on cervical cancer was a partial response (PR)."
- Similarly for Case 4, which also included Bevacizumab, they note that "The effect of treatment on cervical cancer was PR."

3. Clinical Implementation:

- The study demonstrates that Bevacizumab can be successfully incorporated into combination therapy regimens with pembrolizumab and standard chemotherapy agents.
- The authors indicate this approach has shown "significant improvements in overall and progression-free survival for patients with Stage IVB metastatic or locally uncontrolled recurrent cervical cancer."

This analysis suggests that Bevacizumab can be effectively integrated into combination therapy protocols for cervical cancer treatment, with manageable adverse events and positive response rates. However, the limited sample size and specific focus of the study should be considered when interpreting these results.

Note: All quoted material is taken directly from the source document by Adachi et al. The report focuses exclusively on Bevacizumab-related findings while maintaining patient confidentiality and privacy.

4 Part One: Results

The 100 reports were obtained by individually running each article pdf with the Standard 1 prompt according to the process diagram in Figure 2 and Table 1. Generated report word counts are represented as slices in pie chart Figure 3, and include metrics for both word counts and run times. Notable values include a total of 46,326 words for the 100 summaries, with the formatted input to run in Part Two accounting for 49,041 words. The sum of all run times was 2,286.08 seconds, or 38.10 minutes at a standard deviation of 3.00. Each of the 100 generated summaries followed the prompt format task structuring request for “Executive Summary”, “Technical Details”, “Key Insights” in Table 1.

In addition, each generation correctly either identified the article’s authors or multiple authors if it was a review article. Secondly, each generation matched two shorter key terms for content anchoring, along with two exact quotes serving as explicit evidence citations found in the original articles, for a total of 500 verifications across 100 reports. There was no indication that any of these verifications were out of context from the original works. Several errors such as requotations, removal of references at ends of sentences, alternative hyphens, spacing, percentages added, caret insertions, and phrase combinations also maintained context with the original work; with many likely being a model’s preference for formatting or word reductions identified in Table 4. Both reports 023 and 086 shown in Table 3 are examples of reproducible report generations reflected in Bevacizumab S100p Supplementary.

The first example below illustrates the ability of the LLM to adhere to the article authors’ names, two exact terms, and two exact quotations shown as a full generation in Table 3. These verifications further support the quality benefit of using LLMs to analyze large amounts of complex data at high speeds. In addition, an example error type is shown below, reflected by the error summary Table 4. Given the proficiency of each of the 100 summaries meeting a total of 5 correct conditions per report, with some word reduction and formatting modifications made by the LLM, context was maintained for each individual report.

Report 023/100

5 Conditions Met

Author: "De Felice et al."

Exact Term 1: "TNT regimen"

Exact Term 2: "LARC patients"

Exact Quote 1: "mutated RAS-BRAF"

Exact Quote 2: "The addition of a targeted agent to induction FOLFOXIRI and oxaliplatin to 5-fluorouracil-based CRT, with the doses and intensities used in this study, resulted in high CR rates."

Context: = In-Context with original article.

Report 073/100

Error: Truncation, Phrase combinations.

Actual Quote: "atezolizumab plus bevacizumab-treated patients in IMbrave 150 showed no" Next paragraph "difference in response based on immune cell or tumour cell PD-L1 expression unless expression was $\geq 10\%$."

LLM Quote: "atezolizumab plus bevacizumab-treated patients in IMbrave 150 showed no difference in response based on immune cell or tumour cell PD-L1 expression unless expression was $\geq 10\%$."

5 Part One: Discussion

The PRISMA method to identify the 100 articles is believed to be effective for this LLM study across clinical trials, preclinical reports, case studies and reviews in regards to the scoped query of Bevacizumab therapy outcomes due to the 500 verifications of reports. A maximum number of prospective clinical trials was preferred, however the retrospective clinical studies and case reports provided additional useful patient outcome insights. The preclinical studies provided alternative perspectives, while the reviews ensured that leading studies and prevailing trends were included. The designations of colorectal cancer, NSCLC, ovarian cancer, and other cancers in the formatted input played a key role in structuring the Part Two input.

The 100 3.5 Sonnet reports shown in Figure 3 totaled over 46K words, and run times took less than 38 minutes. For frame of reference, when the author read and responded to a prompt on a prior bioprocess engineering article, the resulting 3.5 Sonnet speedup was over 600x faster than the human when adjusting for equivalent word counts (12.92 seconds vs. 9,188.06 seconds) [128]. Therefore, the amount of time for a human to perform this 100 article summaries task could reasonably be over 381 hours vs. 38 minutes for AI. The 3.5 Sonnet standard deviation for time was over tenfold smaller than for words at SD 3.00 vs. SD 42.72. Therefore, in addition to common inference-time compute restrictions, 3.5 Sonnet may also have more pronounced inference time restrictions.

The overall performance of the 100 3.5 Sonnet generations was very good, likely due to the relatively straightforward task of document analysis on a single pdf, combined with model improvements by the manufacturer in recent months [24]. The model was chosen mainly to reduce manufacturer bias between the two parts of the study, combined with recent ChatGPT 4o model modifications which has had difficulty in returning any text quotations followed by the authors’ names, and has opted for the use of hyperlinks that open pdfs. With each of the 100 articles returning the correct article format and five conditions shown in

Table 4, the remaining hallucinations appeared minor - in efforts to reduce word length through requotations, removal of end of sentence references for formatting purposes, and other small modifications; while maintaining the original context of the articles. The most significant error appeared to be the combination of two sentences from paragraphs close in proximity to form a new quotation regarding Article 073 in the error report below.

5.1 Context and Errors

Reports 001-100 Satisfied the Following Conditions					
Author Cited	Exact Term 1	Exact Term 2	Exact Quote 1	Exact Quote 2	Context
100%	100%	100%	100%	100%	In-context
Error Report for Select Bevacizumab Therapy Reports					
Article	Error Type	Instance	Probable Cause	Justification	Context
003	Truncation	Inside parenthesis	Word reduction	Conciseness	In-context
003	Reference removal	End of sentence	Formatting	Conciseness	In-context
003	Requotation	Sentence	Word reduction	Conciseness	In-context
004	Requotation	Sentence	Word reduction	Conciseness	In-context
004	Requotation	Sentence	Word reduction	Conciseness	In-context
009	Hyphen substitution	Mid sentence	Formatting	Consistency	In-context
011	Reference removal	End of sentence	Formatting	Conciseness	In-context
012	Requotation	Sentence	Word reduction	Conciseness	In-context
012	Requotation	Sentence	Word reduction	Conciseness	In-context
013	Additional space	Mid sentence	Formatting	Consistency	In-context
032	Percent added	Inside table	Formatting	Consistency	In-context
048	Table abbr.	Percents added	Formatting	Consistency	In-context
069	Caret insertion	Exponent	Formatting	Consistency	In-context
073	Phrase combinations	Two paragraphs	Formatting	Conciseness	In-context

Table 4: Top: 5 Conditions met for outputs. Bottom: Error types maintained context (not all errors are shown)

6 Part Two: Clinical Decision Support

The combined summaries in Supplementary S100p can be reproduced by copying the 169 page input containing Prompt 2a into ChatGPT o1 or other emerging reasoning models. ChatGPT o1 includes a "details" section returned in generations as shown in Supplementary SREA, but is a simple summary of events, not a full chain-of-thought process. Preliminary tests of other reasoning models either lacked the required context length or did not include sufficient generation details.

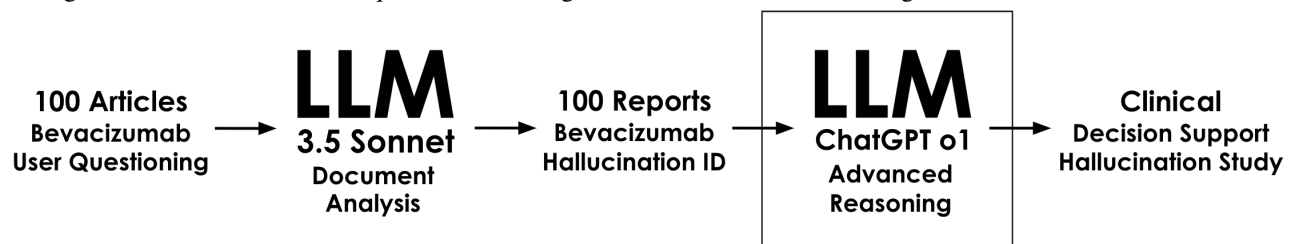


Figure 4: Part Two LLM study

Standard 2a: Bevacizumab Therapy LLM Clinical Decision Support

Prompt 2a: Write a four page report based on the 100 attached summaries regarding the authors' Bevacizumab therapy results, also abbreviated as BEV, BVZ, Bmab, BCZ, or BV using the following format: "Executive Summary", "Technical Details", "Key Insights". Only include results related to Bevacizumab or Bevacizumab combination therapies, deidentifying any patient information if present. Cite selections of numerical data enclosed in quotation marks corresponding to the authors' names according to this example format, Boland P. et al. Only use quotations derived from the attached summaries throughout the report without adding additional quotations. [First Prompt in Conversation] [100 LLM Reports Attached]

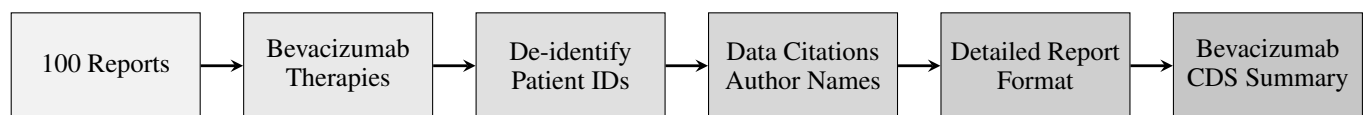


Table 5: Process diagram for Standard 2a

LLM Clinical Decision Report 2a: Various authors. 2024. 1249 words, 146.31 seconds

Executive Summary

Bevacizumab (often referred to as BEV, BVZ, Bmab, BCZ, or BV) is an anti-vascular endothelial growth factor (VEGF) monoclonal antibody used across multiple solid tumors, including colorectal cancer (CRC), non-small cell lung cancer (NSCLC), ovarian cancer, hepatocellular carcinoma (HCC), glioblastoma (GBM), and cervical cancer. Studies have demonstrated that bevacizumab's anti-angiogenic effects help stabilize or slow disease progression, although improvements in overall survival (OS) can be variable depending on tumor biology and treatment line. Across numerous clinical trials and real-world studies, bevacizumab-based combination regimens have shown meaningful benefits in progression-free survival (PFS) and, in certain settings, enhanced tumor response rates.

In metastatic colorectal cancer (mCRC), *"the addition of bevacizumab to chemotherapy provides meaningful survival advantages"* (Pan Q. et al.). However, the magnitude of benefit may be modest in heavily pretreated populations. Similarly, *"bevacizumab therapy remains limited in heavily pre-treated patients, not expected with bevacizumab alone"* (Boland P. et al.). In advanced NSCLC, bevacizumab has a well-established role combined with platinum-based chemotherapy, with some real-world studies highlighting *"bleeding complications were not significantly different. . . though attention is needed for variceal hemorrhage or other risk factors"* (Khaled N. et al.). Immunotherapy combinations with bevacizumab, such as atezolizumab plus bevacizumab, also demonstrate efficacy in patients with specific tumor profiles.

Ovarian cancer studies frequently underscore bevacizumab's benefit in prolonging PFS when used alongside chemotherapy and/or poly (ADP-ribose) polymerase (PARP) inhibitors. For example, *"The median OS was 56.5 months in the olaparib plus bevacizumab group. . . and 51.6 months in the placebo group"* (Ray-Coquard et al., reported by Luvero D. et al.). Meanwhile, in cervical cancer, *"the combination of chemotherapy and bevacizumab significantly improved OS from 13.3 months to 17.0 months"* (Tewari et al., cited by Chang C. et al.). Hepatocellular carcinoma (HCC) findings emphasize the combination of atezolizumab plus bevacizumab, with the pivotal IMbrave150 trial demonstrating *"a median overall survival of 19.2 months. . . compared to 13.4 months in the sorafenib arm"* (Finn et al., reported by Childs A. et al.). Glioblastoma (GBM) studies reveal bevacizumab's effectiveness in alleviating edema and improving PFS but note *"overall survival advantages remain modest"* (Hoosemans L. et al.).

Taken together, bevacizumab is widely recognized for its role in controlling advanced disease progression by inhibiting neoangiogenesis. Common adverse events include hypertension, proteinuria, bleeding risk, and, in some patients, arterial or venous thromboembolism. Real-world evidence suggests vigilance is required for *"rare but significant adverse events, with an overall acceptable safety profile"* (Wang L. et al.).

Technical Details

Mechanism of Action

Bevacizumab binds to VEGF-A, preventing it from engaging VEGF receptors (VEGFR-1, VEGFR-2) on endothelial cells. This blockade *"destroys existing vessels, disturbs neovascularization, and releases intratumoral pressure, ultimately inhibiting angiogenesis"* (Song Y. et al.). Such anti-angiogenic action can improve chemotherapy delivery ("vascular normalization") or relieve peritumoral edema.

Clinical Efficacy Across Indications

1. Colorectal Cancer (CRC):

"XELOX plus bevacizumab regimens is one of the optional first-line treatments of mCRC" (Pan Q. et al.). A heavily pretreated study population revealed *"the median PFS was 1.9 months"* in a phase 1b setting (Boland P. et al.). Predictive biomarkers include circulating angiogenic factors (CAFs), with machine learning efforts showing *"80.8% accuracy in predicting treatment resistance"* (Seufferlein T. et al.).

2. Non-Small Cell Lung Cancer (NSCLC):

Many trials underscore bevacizumab combined with platinum-based doublets. *"A meta-analysis involving advanced lung cancer indicated significantly improved PFS in bevacizumab arms compared to chemotherapy alone"* (Naz T. et al.). Immunotherapy combinations such as atezolizumab plus bevacizumab show *"a median OS improvement from 13.4 to 19.2 months"* (Childs A. et al.). Frequent adverse events include hypertension and proteinuria, while *"epistaxis occurred in up to 20%"* of certain bevacizumab-treated cohorts (Huang Y. et al.).

3. Ovarian Cancer:

Bevacizumab integrated into front-line regimens or used as maintenance therapy. *"ICON7 and GOG-218 data confirm that bevacizumab extended PFS in newly diagnosed advanced cases"* (Chang C. et al.). Combining bevacizumab with PARP inhibitors further enhances PFS in homologous recombination-deficient tumors, although *"the OS advantage in some subgroups remains under study"* (He H. et al.).

4. Hepatocellular Carcinoma (HCC):

"Atezolizumab plus bevacizumab represents an established first-line standard", with improvements in PFS and OS over sorafenib (Childs A. et al.). *"Grade ≥ 3 bleeding rates"* remain an area of caution, though *"in real-world cohorts, no grade 4/5 bleeding events were observed beyond typical expectations"* (Khaled N. et al.).

5. Glioblastoma (GBM) and High-Grade Glioma:

Bevacizumab alleviates vasogenic edema, leading to improved radiographic appearances (pseudoresponse). *"BEV combination therapy improved PFS but rarely translated into prolonged OS"* (Hoosemans L. et al.). Resistance often arises through alternative angiogenic pathways (Rahman M. et al.).

6. Cervical Cancer:

"The GOG 240 study found that adding bevacizumab to chemotherapy increased OS to 17.0 months from 13.3 months" (Tewari et al., referenced by Chang C. et al.). Toxicities, including fistula formation, hemorrhage, and hypertension, require vigilant monitoring.

Safety Profile

Common toxicities across indications include:

Hypertension: *"The incidence can be as high as 30% in certain patient populations"* (Zheng Z. et al.). **Proteinuria:** *"Grade 3 or higher proteinuria was observed in about 1.4% of patients"* (Wang L. et al.). **Bleeding & Thromboembolism:** While rates vary, *"Incidents of GI perforation, hemorrhage, and arterial or venous thrombosis have been documented"* (Jacobsen A. et al.). **Rare Events:** *"A few cases of nasal septal perforation, necrotizing fasciitis, and hypertensive encephalopathy are seen in post-marketing data"* (Wang L. et al.). In real-world pharmacovigilance, *"Five years of safety profile data highlight that bevacizumab's main AEs remain manageable, but active monitoring is advised"* (Wang L. et al.).

Key Insights

1. Combination Therapy Enhances Benefit

The majority of evidence indicates that bevacizumab performs best when paired with chemotherapy, immunotherapy, or targeted agents. In multiple tumor types (e.g., mCRC, NSCLC, ovarian cancer), it *"significantly improves progression-free survival, albeit OS advantages can be more modest"* (Song Y. et al.; Kim Y. et al.).

2. Patient Subset Identification

Studies highlight the importance of biomarkers to refine patient selection. *"Genetic and epigenetic profiles, VEGF polymorphisms, and circulating angiogenic factors"* may predict resistance or response (Seufferlein T. et al.; Fleischer T. et al.). In HCC, *"baseline tumor vasculature characteristics"* can correlate with bevacizumab efficacy (Mazard T. et al.).

3. Resistance and Alternative Pathways

Despite initial responses, *"tumors can develop VEGF-independent angiogenic routes, rendering bevacizumab less effective"* (Rahman M. et al.). Additional antiangiogenic or immunomodulatory strategies may be necessary to overcome resistance.

4. Significance of Safety Monitoring

Although well-established, bevacizumab requires careful surveillance for hypertension, proteinuria, bleeding events, and less common but serious complications. *"Bevacizumab discontinuation is recommended for severe grade 4 hypertension, arterial thrombotic events, or GI perforation"* (Chang C. et al.).

5. Emerging Applications

Newer trials explore bevacizumab with tyrosine kinase inhibitors, PARP inhibitors, immune checkpoint inhibitors, and locoregional therapies (e.g., TACE in HCC). Preliminary data show synergy in certain advanced or refractory settings, demanding further randomized validation.

Concluding Remarks

Bevacizumab remains a pivotal anti-angiogenic agent with documented improvements in PFS and symptom control across multiple solid tumors. While overall survival gains can vary, strategic combination regimens, biomarker-driven patient selection, and close safety monitoring may optimize bevacizumab's clinical value. The assembled evidence indicates a generally favorable risk-benefit ratio, especially when used alongside standard chemotherapy or immunotherapies, though ongoing research continues to refine its optimal role in advanced cancer care.

Table 6: ChatGPT o1 Prompt 2a generation, See Supplementary SREA for full response

LLM Dosing Recommendations Excerpt 2b. Various authors. 2024. 1685 words, 141.77 seconds

Proposed Dosing Recommendations

Building on the data synthesized above, the following recommendations aim to optimize bevacizumab dosing under specific clinical conditions:

1. Heavily Pretreated CRC

Recommendation: Increase bevacizumab from 5 mg/kg to **7.5–10 mg/kg** every 14 days in patients who have tolerated prior regimens and have no significant hypertension or renal impairment.

Rationale: “It is conceivable that a higher bevacizumab dose may have achieved greater synergy with ME-344 leading to a higher response rate” in refractory disease (Boland P. et al.).

2. Ovarian or Cervical Cancer in Frail or Older Patients

Recommendation: Initiate bevacizumab at **7.5 mg/kg** every 3 weeks alongside chemotherapy and/or immunotherapy, especially in patients with baseline controlled BP.

Rationale: “Bevacizumab at 7.5 mg/kg... shows similar benefits with reduced toxicity compared to 15 mg/kg doses” (Chen X. et al.; Li S. et al.).

3. Neuro-Oncology (GBM) for Symptomatic Edema Control

Recommendation: Employ short-course **5 mg/kg** bevacizumab every 2 weeks or prn for peritumoral edema alleviation.

Rationale: “Low-dose bevacizumab significantly improves tumor vasogenic edema with fewer grade 3–4 adverse events” (Guo G. et al.).

4. Standard High-Dose Regimens for Front-Line HCC or Advanced Ovarian Cancer

Recommendation: Use **15 mg/kg** bevacizumab Q3W for advanced HCC in combination with atezolizumab, or for front-line advanced ovarian cancer with carboplatin-based chemotherapy.

Rationale: “Atezolizumab plus bevacizumab 15 mg/kg... in advanced HCC demonstrated remarkable efficacy” (Childs A. et al.); “Maintenance therapy with bevacizumab 15 mg/kg every 3 weeks... in GOG-218 improved PFS” (Chang C. et al.).

5. Safety Monitoring and Dose Adjustments

Recommendation: In all scenarios, monitor for grade ≥ 3 hypertension, proteinuria (>2 g/24 h), or severe bleeding. Temporarily suspend bevacizumab and resume at a lower dose once toxicity resolves to grade ≤ 1 .

Rationale: “The combination therapy was associated with manageable toxicity, but timely intervention on arterial hypertension and proteinuria is required” (Zheng Z. et al.; Moiscu D. et al.).

LLM Combination Therapies Excerpt 2c. Various authors. 2024. 1280 words, 99.15 seconds

Technical Details

1. Metastatic Colorectal Cancer (mCRC)

Proposed New Combination:

Bevacizumab + OxPhos Inhibitor (e.g., ME-344) + Chemotherapy

Rationale: In a phase 1b trial, “the bevacizumab dose used (5 mg/kg) was lower than doses used in some other studies... conceivable that a higher bevacizumab dose may have achieved greater synergy” (Boland P. et al.).

Recommended Doses:

Bevacizumab: 7.5 mg/kg every 2 weeks (moderate increase from the prior 5 mg/kg), or escalate to 10 mg/kg in heavily pretreated patients with stable blood pressure and adequate renal function.

ME-344: Adhere to the protocol studied (depending on ongoing research), plus a standard chemotherapy backbone (FOLFOX or FOLFIRI).

Clinical Benefits: Potential for deeper tumor regression in refractory mCRC by blocking angiogenesis (via bevacizumab) and mitochondrial metabolism (via ME-344), while standard chemo sustains cytotoxic pressure.

Alternate Doublet:

Bevacizumab + Trifluridine/Tipiracil (TAS-102)

Rationale: “Third-line or later therapy with TAS-102 plus bevacizumab significantly prolonged OS... particularly in patients with prior BEV exposure” (André T. et al.).

Recommended Doses:

Bevacizumab: 5 mg/kg or 7.5 mg/kg every 2 weeks, depending on tolerability.

TAS-102: 35 mg/m² orally twice daily on days 1–5, repeated every 28 days.

2. Non-Small Cell Lung Cancer (NSCLC)

Proposed New Combination:

Bevacizumab + PD-1/PD-L1 Inhibitor + Platinum Doublet

Rationale: “Immune checkpoint inhibitors combined with bevacizumab improved PFS and OS in advanced NSCLC, with controllable toxicity profiles” (Huang Y. et al.; Wu S. et al.).

Recommended Doses:

Bevacizumab: 15 mg/kg every 3 weeks (the standard dose in many trials such as IMpower150).

Platinum Doublet (e.g., Carboplatin/Pemetrexed): Follow standard dosing intervals for 4–6 cycles.

PD-1/PD-L1 Inhibitor (e.g., Atezolizumab): 1200 mg every 3 weeks.

Clinical Benefits: Enhanced immunomodulation (via PD-1/PD-L1 blockade) plus angiogenesis inhibition has shown “superior survival outcomes in EGFR wild-type or post-TKI failure patients” (Nosaka T. et al.).

Low-Dose Alternative: Bevacizumab 7.5 mg/kg + Single-Agent Immunotherapy for frail patients or those with borderline performance status. Some studies indicate reduced severe hypertension or bleeding risk with lower bevacizumab doses (Chen X. et al.), though prospective data are limited.

LLM Biomarker-Driven Selections Excerpt 2d. Various authors. 2024. 1118 words, 79.78 seconds

Page 2: Biomarker-Driven Patient Selection and Response Evaluation

1. VEGF-Related Biomarkers and Genetic Screening

Multiple reports highlight the utility of identifying patients who harbor specific VEGF polymorphisms or other angiogenesis-related genetic variants. As one study notes, “the CC genotype of rs3025039 polymorphism of VEGF-A... was significantly related to time-to-treatment failure” (Sibertin-Blanc et al.). Incorporating such genetic insights may guide whether bevacizumab is likely to confer robust benefits or lead to early resistance.

In ovarian cancer, “HRD profiling could predict patients’ prolonged benefit from combined bevacizumab-PARP inhibitors” (Freyer G. et al.).

2. Imaging-Based Biomarkers

Advanced radiological tools—such as dynamic contrast-enhanced MRI, perfusion CT, or radiomics—might predict bevacizumab sensitivity. “Radiomics models predicted response similarly to expert radiologists (AUCs of 0.69–0.79)” (Van Griethuysen et al.). Routine integration of functional imaging could identify early changes in tumor vascularity, suggesting the need for dose escalation or therapy alteration.

3. Early Circulating Biomarkers

Several summaries referenced measuring “serum VEGF-A121 isoform” or other circulating cytokines to assess tumor angiogenesis adaptation (Okawa M. et al.). Monitoring these factors early in treatment might reveal subclinical resistance and enable timely therapy modifications.

4. Integration into Treatment Algorithms

Implementing biomarker assessments before therapy initiation and at defined intervals (e.g., every 6 weeks or after 2–3 cycles) may refine clinical decision-making. For instance, “biomarker-based early switch to another anti-angiogenic treatment could improve outcomes in mCRC patients with early disease progression” (Seufferlein T. et al.).

Clinically, such an approach helps avoid prolonged exposure to ineffective regimens, reducing toxicity and cost.

Table 7: ChatGPT o1 excerpts 2b, 2c, 2d. See Supplementary SREA for full responses

Protocol 2b-d: LLM Dosing, Drug Combinations, and Biomarkers

Prompt 2b: In a four page report, what are your new dosing recommendations to improve specific Bevacizumab or Bevacizumab combination therapy outcomes based on the 100 attached summaries? [Second Prompt in Conversation] [Inherited Partial Guardrails from Prompt 2a]

Prompt 2c: In a four page report, what are your new combination therapy with dosing recommendations to improve specific Bevacizumab or Bevacizumab combination therapy outcomes based on the 100 attached summaries? [Third Prompt in Conversation] [Inherited Partial Guardrails from Prompt 2a]

Prompt 2d: In a four page report, what else would be clinically useful for physicians to implement regarding improving specific Bevacizumab or Bevacizumab combination therapy outcomes based on the 100 attached summaries? [Fourth Prompt in Conversation] [Inherited Partial Guardrails from Prompt 2a]

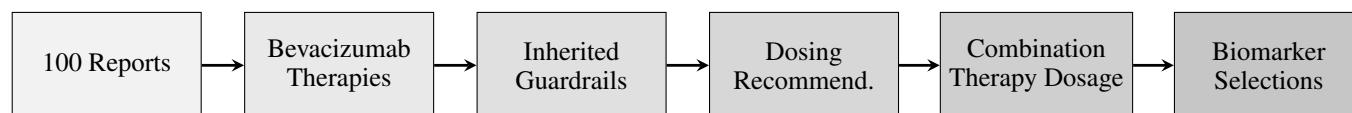


Table 8: Process diagram for Protocol 2b-d

7 Part Two: Results

7.1 Clinical Decision Support

Part Two used the 100 reports previously generated by 3.5 Sonnet and the prompt from Standard 2a shown in Figure 4 and Table 5. ChatGPT o1's response followed the "Executive Summary", "Technical Details", "Key Insights" format in 1249 words. Generation 2a: 5 lines of introductory and exiting statements are available in Supplementary SREA. The reasoning model took 146.31 seconds to produce a convincing summary of the 100 reports, with an emphasis placed on the therapy associated with the type of cancer (mCRC, NSCLC, Ovarian cancer, etc.). All of the ChatGPT o1 author quotation pairs appeared to be in context with the original 100 summaries for generation 2a, however several requotations and approximations were present, as depicted in Table 9. Of these errors, 34.5% were attributed to incorrect citations of the both the author and the sentence being quoted. The formatted input length of 49,041 words, estimated to be 64,000 tokens is less than the 128K token limit, but still posed challenges due to input complexity.

The following example for generation 2a includes notable AI clinical decision support insights followed by one hallucination example. Generation 2a provided an updated safety profile for Bevacizumab based on the 100 article summaries across multiple known side effects. Together, these examples illustrate the capacity for ChatGPT o1 to incorporate meaningful clinical decision support from a larger number of recent Bevacizumab studies.

LLM Clinical Decision Support 2a

Safety Profile: Common toxicities across indications include:

Hypertension: "The incidence can be as high as 30% in certain patient populations" (Zheng Z. et al.).

Proteinuria: "Grade 3 or higher proteinuria was observed in about 1.4% of patients" (Wang L. et al.).

Bleeding & Thromboembolism: While rates vary, "Incidents of GI perforation, hemorrhage, and arterial or venous thrombosis have been documented" (Jacobsen A. et al.).

Rare Events: "A few cases of nasal septal perforation, necrotizing fasciitis, and hypertensive encephalopathy are seen in post-marketing data" (Wang L. et al.). In real-world pharmacovigilance, "Five years of safety profile data highlight that bevacizumab's main AEs remain manageable, but active monitoring is advised" (Wang L. et al.).

Hallucination Example

Error Type: Author, sentence (Cheng et al., requotation).

Actual Quote: In the pivotal IMbrave150 trial, according to Cheng et al., the combination of atezolizumab and bevacizumab showed "a median OS of 19.2 months (95% CI 17-23.7) in the atezolizumab/bevacizumab arm compared to 13.4 months (95% CI 11.4-16.9; stratified HR 0.66, 95% CI 0.52-0.85, $p < 0.001$) in the sorafenib arm."

LLM Quote: Hepatocellular carcinoma (HCC) findings emphasize the combination of atezolizumab plus bevacizumab, with the pivotal IMbrave150 trial demonstrating "a median overall survival of 19.2 months. . . compared to 13.4 months in the sorafenib arm" (Finn et al., reported by Childs A. et al.). from Table 9

7.2 Dosage, Combination Therapy, and Biomarkers

The three remaining generations provided additional CDS findings, and were based on more brief and informal prompts as exemplified in Protocol 2b-d Table 8. Excerpt 2b provided recommended dosages based on cancer types, with details shown in

Table 7. Excerpt 2c illustrated combined therapies along with dosage recommendations, while Excerpt 2d offered biomarker-driven patient selection and response evaluations, also seen in Table 7. Generations c-d had a lower number of errors due to a lower number of author quotation pairs. All three of the generations are available in their entirety in Supplementary SREA.

In terms of hallucinations, each of the three responses struggled with both author and sentence quotations both in maintaining context of the summaries, and losing context in some cases. Generation 2b errors affecting both the author and quote had the highest error type rate of 59.3%, while generation 2c was 38.5%, and 2d was 46.2%. The loss of context ranged from citing the incorrect form of cancer to harder to difficulties in locating information across several articles. In a rare case, generation 2d referenced enhancing patient education, but this concept was not explicit in summaries. Generation 2c included a new NSLC proposed drug combination, which highlights recommended doses for each drug, clinical benefits, as well as a low-dose alternative shown below.

Combination Therapies Recommendations 2c

Non-Small Cell Lung Cancer (NSCLC))

Proposed New Combination:

Bevacizumab + PD-1/PD-L1 Inhibitor + Platinum Doublet

Rationale: “Immune checkpoint inhibitors combined with bevacizumab improved PFS and OS in advanced NSCLC, with controllable toxicity profiles” (Huang Y. et al.; Wu S. et al.).

Recommended Doses: Bevacizumab: 15 mg/kg every 3 weeks (the standard dose in many trials such as IMpower150).

Platinum Doublet (e.g., Carboplatin/Pemetrexed): Follow standard dosing intervals for 4–6 cycles.

PD-1/PD-L1 Inhibitor (e.g., Atezolizumab): 1200 mg every 3 weeks.

Clinical Benefits: Enhanced immunomodulation (via PD-1/PD-L1 blockade) plus angiogenesis inhibition has shown “superior survival outcomes in EGFR wild-type or post-TKI failure patients” (Nosaka T. et al.).

Low-Dose Alternative: Bevacizumab 7.5 mg/kg + Single-Agent Immunotherapy for frail patients or those with borderline performance status. Some studies indicate reduced severe hypertension or bleeding risk with lower bevacizumab doses (Chen X. et al.), though prospective data are limited.

Hallucination Example

Error Type: Sentence (requotation). “the bevacizumab dose used (5 mg/kg) was lower than doses used in some other studies. ... conceivable that a higher bevacizumab dose may have achieved greater synergy” from Table 9

LLM Quote: The authors note that the bevacizumab dose used (5 mg/kg) was lower than doses used in some other studies. As Boland et al. state, “it is conceivable that a higher bevacizumab dose may have achieved greater synergy with ME-344 leading to a higher response rate.”

8 Part Two: Discussion

ChatGPT o1’s response to Prompt 2a appears to be high quality given the size, detail, and complexity of the input provided to the model. Despite its longer generation time of 146.31 seconds required to solve the problem; a coherence priority and inference-time compute restrictions were the likely cause of poorer exact recall of quotations and author names. The problem was considerably more challenging than Part I articles as an input due to length: 49K words vs. 9K average words, type of study: 6 study types vs. 1, number of diseases: 12 cancer types vs. single publications, and disparate information: 100 summaries vs. a single cohesive article. The remaining three Part Two prompts could likely have been improved by explicitly defining guardrails similar to Prompt 2a for each additional prompt in the same conversation.

It is not believed that context length was a limitation to response quality, as the original input and prompt were less than the 128K token limit. Although the generations primarily had issues with auxiliary terms, they did not appear to provide misinformation from a clinical decision support system perspective. A main limitation was likely the explicit use of guardrails in the prompts, as verifications of quotations became more difficult for the manuscript author to identify, and generation 2d did not follow the correct format defined in Prompt 2a. The other limitation with not explicitly stating guardrails on each prompt within the same conversation was that although the output lengths were similar, the time to process requests decreased (146.31, 141.77, 99.15, 79.78 seconds), as this may indicate a lower reliance on analyzing the input data and a greater reliance on ChatGPT o1’s own parameters to solve problems.

9 Limitations

As LLMs have greatly improved since the release of ChatGPT 4o in May 2024, extrinsic hallucinations attributing models with inventing facts or details without grounding to the reference material have become less common. However intrinsic hallucination limitations were present where outputs were internally inconsistent or contradicted information provided within the source/context. This was most evident in the logical contradictions of cancer types, especially for generations 2b-d. Contextual hallucinations in the form of misattributing author names were present in generations 2a-d. Meta-level conflicts were likely also present due to issues or challenges that arose when the LLM operated across multiple, high-level contexts, domains, or frameworks.

Error Report for Bevacizumab Clinical Decision Support 2a					
Author	Error Type	Instance	Probable Cause	Justification	Context
Pan Q. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Boland P. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Khaled N. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Ray-Coquard et al.	Requotation	Author, sentence	Word reduction	Conciseness	In-context
Tewari et al.	Requotation	Author, sentence	Inference limit	Pattern approx	In-context
Finn et al.	Requotation	Author, sentence	Inference limit	Pattern approx	In-context
Hoosemans L. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Wang L. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Song Y. et al. 2024	Requotation	Author, sentence	Inference limit	Pattern approx	In-context
Boland P. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Seufferlein T. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Naz T. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Childs A. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Huang Y. et al.	Approximation	Sentence	Formatting	Conciseness	In-context
Chang C. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	In-context
He H. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	In-context
Childs A. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Khaled N. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	In-context
Hoosemans L. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	In-context
Tewari et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Zheng Z. et al.	Approximation	Sentence	Word reduction	Conciseness	In-context
Wang L. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Jacobsen A. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Wang L. et al.	Approximation	Sentence	Word reduction	Conciseness	In-context
Wang L. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Song Y. et al.	Approximation	Author, sentence	Inference limit	Pattern approx	In-context
Seufferlein T. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Rahman M. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Chang C. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	In-context
Error Report for Bevacizumab Dosing 2b, Partial Guardrails					
Author	Error Type	Instance	Probable Cause	Justification	Context
Boland P. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Li S. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Childs A. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	In-context
Boland P. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Li S. et al.	Requotation	Sentence	Inference limit	Pattern approx	Less context
Chen X. et al.	Requotation	Sentence	Inference limit	Pattern approx	Less context
Guo G. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Zhang J. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	In-context
Boland P. et al.	Requotation	Sentence	Inference limit	Pattern approx	Less context
Childs A. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	In-context
Chang C. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	In-context
Kuo C. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	In-context
Zheng Z. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	In-context
Moisuc D. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	In-context
Jacobsen A. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	In-context
Boland P. et al.	Approximation	Sentence	Inference limit	Pattern approx	Less context
Chen X. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	Less context
Zhang P. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	In-context
Pan Q. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	In-context
Guo G. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Childs A. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	In-context
Moisuc D. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	In-context
Chen X. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Guo G. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Childs A. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	Less context
Chang C. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	In-context
Zheng Z. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	In-context

Table 9: Report 2a: Less context errors = 0%, Report 2b: Less context errors = 22.2%. Lower is better

Error Report for Bevacizumab Combination Dosing 2c, Partial Guardrails					
Author	Error Type	Instance	Probable Cause	Justification	Context
Jacobsen A. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Chen X. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Boland P. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
André T. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Huang Y. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	Less context
Nosaka T. et al.,	Requotation	Sentence	Inference limit	Pattern approx	Less context
Freyer G. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Chen X. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Guo G. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Oura K. et al.	Approximation	Author, sentence	Inference limit	Pattern approx	In-context
Qin et al.	Approximation	Author, sentence	Inference limit	Pattern approx	In-context
Zheng Z. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	In-context
Hu Z. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	In-context
Error Report for Bevacizumab Biomarkers 2d, Partial Guardrails					
Author	Error Type	Instance	Probable Cause	Justification	Context
Sibertin-Blanc et al.	Approximation	Author, sentence	Inference limit	Pattern approx	In-context
Freyer G. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	In-context
Van Griethuysen et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Okawa M. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Seufferlein T. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Zheng Z. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Chen X. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	Less context
Jacobsen A. et al.	Requotation	Sentence	Inference limit	Pattern approx	Less context
Huang Y. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	Less context
Chang C. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	Less context
Zheng Z. et al.	Requotation	Author, sentence	Inference limit	Pattern approx	Less context
Wang L. et al.	Requotation	Sentence	Word reduction	Conciseness	In-context
Mazard T. et al.	Requotation	Sentence	Inference limit	Pattern approx	Less context

Table 10: Report 2c: Less context errors = 15.4%, Report 2d: Less context errors = 46.2%. Lower is better

10 Limitations (continued)

Citation and reference hallucinations were associated primarily with misreferenced sources. Syntactic hallucinations where the output was nonsensical or failed to parse were not present. ChatGPT 4o was utilized to assist in the understanding of semantic hallucinations - where a content's meaning was incorrect or completely made up were not apparent, except for a limited number of output from generations 2b-d. This included incorporating the importance of patient education, a topic not explicitly found in the 100 summaries. Embellishment or unwarranted speculation have also not been identified due to high proficiency in context able to be verified, especially for Prompts 1 and 2a. No self contradictions of earlier statements in the same response were identified. The LLM hallucination limitations covered here were believed to stem primarily from inference-time compute, inference-time budget, or output-length constraints which limit the number of floating-point operations or attention steps detailed in the following paragraphs.

ChatGPT o1 pro mode model trial responses were approximately 1.7x slower than ChatGPT o1, had shorter responses, had less convincing solutions, did not address hallucination issues, as was experienced in testing of a prior cancer research study [129]. As mentioned earlier, the ChatGPT o1 128K tokens context length limitation should not have been approached regarding the input size and subsequent prompts in the same conversation. However, context limitations may have contributed to paraphrasings due to the model's primary purpose to produce natural-sounding text. Instruction layer conflicts of the model's built-in compliance and style-generation mechanisms may have also contributed to its preference for coherent text over more strict citations. The LLM's architecture consists of self attention layers with fixed attention budgets which may have contributed to not giving equal weight to each token. As LLMs were trained to predict the next word in a sentence, there will likely continue to be a bias towards coherence and fluency vs. facts that are mentioned a limited number of times.

Essentially, LLMs are based on internal patterns to transform text to obtain a larger understanding under a coherence bias, instead of memorizing text, which can limit its ability for specific recall. An additional RAG system may have assisted in processing precise chunks of text on demand instead of ChatGPT o1's methods of pattern completion. A more structured format for prompts, especially for Prompts 2b-d could have also increased the ability to process exact details. When ChatGPT o1 was asked a question regarding LLM hardware limitations, its response was "Most Hosted Model Infrastructures Are Well-Resourced", and that commercial LLMs "typically have enough "horsepower" to handle their advertised maximum context window", where "cost, time, or token constraints" are the primary concern.

11 Future work

The immediate remedy to improve future performance across Prompts 2a-d would be to decrease the input size. An example of a smaller input for more effective processing was a 20 summary 8.8K word cancer research input using a similar format for ChatGPT o1, which effectively yielded over 1 citation per reference (1621 words, 57.03 seconds). The same was true for citations per reference in the identical Kawchak K. study, processing 20 summaries at 8.5K words regarding cancer mechanisms (1355 words, 68.85 seconds) [129]. The smaller input sizes represented by faster responses likely contributed to better representations of the source data, than with longer more extreme sized outputs experienced in this study vs. additional prior biosynthesis and bioprocess engineering works [130, 131].

The desire to increase both the scope and complexity in future drug discovery studies is a continual goal. With other reasoning models emerging [132, 133] that will likely compete with OpenAI's o1 model, a new OpenAI reasoning model termed ChatGPT o3 is in development [134], and plans to increase national AI infrastructure with a \$500B investment with operations led by OpenAI [135], it is reasonable to forecast improvements in inference-time compute, inference-time budget, and output-length constraints to more effectively process the high dimensional and complex details in this study. The number of critical steps in correctly preparing the study's dataset, processing with one LLM, formatting the data, and running experiments on a second LLM will likely not be realized by AI agents in the near-term, but will be evaluated on an on-going basis. In addition, this study currently has an advantage over agents regarding explainability and reproducibility, as the number of ways this complex problem can be solved likely increases significantly with more complex systems. Therefore, larger and more comprehensive cancer projects will be pursued using best available AI methods, learning from constraints in the study.

12 Conclusions

2025 LLM clinical decision support systems have improved in performance over the previous year's translational research applications, primarily due to larger and more capable AI models. These increasing LLM incorporation efforts using medical data are primarily being made by researchers to increase the dissemination of recent results to medical professionals who don't have time to read every new study and making correlations across dissimilar findings. The increased software size has led to a number of challenges to continue scaling inferencing quality due to manufacturer cost, time, or token considerations. Other challenges that affect key reasoning models such as ChatGPT o1 are known to exist, such as maintaining output accuracy for longer chain-of-thought prompt lengths [16], and the demand for more RAM dedicated to the attention mechanism [20]. However, most evidence supports the notion that there is not a lack of computational power, but rather financial considerations in hosting models that must maintain acceptable output quality to many users.

This study represented a comprehensive use case of real-world data with Bevacizumab therapy results, a humanized monoclonal IgG1 antibody used for treating a wide range of cancers. An increase of nearly 3x larger input size across twelve cancer types and five study types vs. a prior study were implemented to better understand how Bevacizumab combination therapies could be improved using LLM CDS. 3.5 Sonnet provided the same report format with privacy filtering across all 100 articles, with five human verifications across author name, key terms, and quotations per report. Hallucinations using document analysis were minor, appeared to maintain context with the original works, and may reflect the software manufacturer's efforts to manage costs by requesting larger phrases and following predetermined formatting rules.

The 169 page 49K word summary was included with a structured prompt for ChatGPT o1 to extract pertinent information regarding CDS over the complexities of Bevacizumab related cancer and study therapies. In general, very good evidence-based insight maintaining context with the 100 detailed reports was apparent in the form of clinical efficacy across multiple indications, as well as a detailed safety profile with common toxicities collected across the cancer studies. Additional less-structured prompts were run in the same conversation, with several convincing examples regarding proposed dosing recommendations, combination therapies, and biomarker-driven selections. However generated sentences experienced a higher prevalence of less context compared to the input data such as incorrect cancer types, and the model appearing to rely on its own parameters regarding CDS patient education.

Large language models (LLMs) are increasingly evaluated in clinical decision support, but their performance in highly specialized fields, such as Bevacizumab clinical trials, was not well established. The outcomes of the study from 3.5 Sonnet and ChatGPT o1 illustrate a larger capacity for models to incorporate meaningful clinical decision support from an extensive number of recent Bevacizumab studies. As many LLM cancer research studies have lacked a substantially large problem and sufficient rationale regarding the limitation of their work, this study helped to pave the way for developers and software manufacturers alike to confront issues regarding practical use of LLMs using real world data beyond small-scale problems and benchmarks. As LLMs continue to improve and inference-time compute is increased due to improved U.S. AI architecture, so should the quality of generations through the progress of reducing errors.

References

- [1] Christoph Raphael Buhr, Benjamin Philipp Ernst, Andrew Blaikie, Harry Smith, Tom Kelsey, Christoph Matthias, Maximilian Fleischmann, Florian Jungmann, Jürgen Alt, Christian Brandts, Peer W. Kämmerer, Sebastian Foersch, Sebastian Kuhn, and Jonas Eckrich. Assessment of decision-making with locally run and web-based large language models versus human board recommendations in otorhinolaryngology, head and neck surgery. *European Archives of Oto-Rhino-Laryngology*, January 2025. doi:10.1007/s00405-024-09153-3.
- [2] Brian Hyeonseok Kim and Chao Wang. Large language models for interpretable mental health diagnosis. *arXiv*, (arXiv:2501.07653), January 2025. arXiv:2501.07653. URL: <http://arxiv.org/abs/2501.07653>, doi:10.48550/arXiv.2501.07653.
- [3] Luis Felipe Ensina, Matheus Matos Machado, Joice B. Machado Marques, Monica Pugliese H. Dos Santos, Fábio Cerqueira Lario, Chayanne Andrade Araújo, Fabiana Andrade Nunes Oliveira, and Dilvan De Abreu Moreira. Artificial intelligence for detecting anaphylaxis in electronic medical records. *Asia Pacific Allergy*, January 2025. URL: <https://journals.lww.com/10.5415/apallergy.0000000000000179>, doi:10.5415/apallergy.0000000000000179.
- [4] Yanjun Gao, Skatje Myers, Shan Chen, Dmitriy Dligach, Timothy Miller, Danielle S Bitterman, Guanhua Chen, Anoop Mayampurath, Matthew M Churpek, and Majid Afshar. Uncertainty estimation in diagnosis generation from large language models: next-word probability is not pre-test probability. *JAMIA Open*, 8(1):ooae154, December 2024. URL: <https://academic.oup.com/jamiaopen/article/doi/10.1093/jamiaopen/ooae154/7951510>, doi:10.1093/jamiaopen/ooae154.
- [5] Maxime Griot, Coralie Hemptinne, Jean Vanderdonckt, and Demet Yuksel. Large language models lack essential metacognition for reliable medical reasoning. *Nature Communications*, 16(1):642, January 2025. URL: <https://www.nature.com/articles/s41467-024-55628-6>, doi:10.1038/s41467-024-55628-6.
- [6] David Steybe, Philipp Poxleitner, Suad Aljohani, Bente Brokstad Herlofson, Ourania Nicolatou-Galitis, Vinod Patel, Stefano Fedele, Tae-Geon Kwon, Vittorio Fusco, Sarina E. C. Pichardo, Katharina Theresa Obermeier, Sven Otto, Alexander Rau, and Maximilian Frederik Russe. Evaluation of a context-aware chatbot using retrieval-augmented generation for answering clinical questions on medication-related osteonecrosis of the jaw. *Journal of Cranio-Maxillofacial Surgery*, January 2025. URL: <https://www.sciencedirect.com/science/article/pii/S101051822400341X>, doi:10.1016/j.jcms.2024.12.009.
- [7] Tsz Kin Chan and Ngoc-Duy Dinh. Entagents: Ai agents for complex knowledge otolaryngology. *medRxiv*, January 2025. URL: <https://www.medrxiv.org/content/10.1101/2025.01.01.25319863v1>, doi:10.1101/2025.01.01.25319863.
- [8] Ryunosuke Noda, Chiaki Yuasa, Fumiya Kitano, Daisuke Ichikawa, and Yugo Shibagaki. Performance of o1 pro and gpt-4 in self-assessment questions for nephrology board renewal. *medRxiv*, January 2025. URL: <https://www.medrxiv.org/content/10.1101/2025.01.14.25320525v1>, doi:10.1101/2025.01.14.25320525.
- [9] Siru Liu, Allison B McCoy, and Adam Wright. Improving large language model applications in biomedicine with retrieval-augmented generation: a systematic review, meta-analysis, and clinical development guidelines. *Journal of the American Medical Informatics Association*, page ocaf008, January 2025. URL: <https://academic.oup.com/jamia/advance-article/doi/10.1093/jamia/ocaf008/7954485>, doi:10.1093/jamia/ocaf008.
- [10] Jack Gallifant, Majid Afshar, Saleem Ameen, Yindalon Aphinyanaphongs, Shan Chen, Giovanni Cacciamani, Dina Demner-Fushman, Dmitriy Dligach, Roxana Daneshjou, Chrystinne Fernandes, Lasse Hyldig Hansen, Adam Landman, Lisa Lehmann, Liam G. McCoy, Timothy Miller, Amy Moreno, Nikolaj Munch, David Restrepo, Guergana Savova, Renato Umeton, Judy Wawira Gichoya, Gary S. Collins, Karel G. M. Moons, Leo A. Celi, and Danielle S. Bitterman. The tripod-llm reporting guideline for studies using large language models. *Nature Medicine*, page 1–10, January 2025. URL: <https://www.nature.com/articles/s41591-024-03425-5>, doi:10.1038/s41591-024-03425-5.
- [11] Matteo Malgaroli, Katharina Schultebrasucks, Keris Jan Myrick, Alexandre Andrade Loch, Laura Ospina-Pinillos, Tanzeem Choudhury, Roman Kotov, Munmun De Choudhury, and John Torous. Large language models for the mental health community: framework for translating code to care. *The Lancet Digital Health*, page S2589750024002553, January 2025. URL: <https://linkinghub.elsevier.com/retrieve/pii/S2589750024002553>, doi:10.1016/S2589-7500(24)00255-3.
- [12] Zhihang Yuan, Yuzhang Shang, Yang Zhou, Zhen Dong, Zhe Zhou, Chenhao Xue, Bingzhe Wu, Zhikai Li, Qingyi Gu, Yong Jae Lee, Yan Yan, Beidi Chen, Guangyu Sun, and Kurt Keutzer. Llm inference unveiled: Survey and roofline model insights. *arXiv*, (arXiv:2402.16363), May 2024. arXiv:2402.16363. URL: <http://arxiv.org/abs/2402.16363>, doi:10.48550/arXiv.2402.16363.
- [13] Rui Li, Deji Fu, Chunyu Shi, Zhilan Huang, and Gang Lu. Efficient llms training and inference: An introduction. *IEEE Access*, page 1–1, 2024. URL: <https://ieeexplore.ieee.org/abstract/document/10756602>, doi:10.1109/ACCESS.2024.3501358.
- [14] OpenAI. Openai 40 system card. this report outlines the safety work carried out prior to releasing gpt-4o including external red teaming, frontier risk evaluations according to our preparedness framework, and an overview of the mitigations we built in to address key risk areas., 2024. URL: <https://openai.com/index/gpt-4o-system-card/>.

- [15] OpenAI. Openai o1 system card. this report outlines the safety work carried out prior to releasing openai o1, including external red teaming and frontier risk evaluations according to our preparedness framework., 2024. URL: <https://openai.com/index/openai-o1-system-card/>.
- [16] Zabir Al Nazi, Md. Rajib Hossain, and Faisal Al Mamun. Evaluation of open and closed-source llms for low-resource language with zero-shot, few-shot, and chain-of-thought prompting. *Natural Language Processing Journal*, 10:100124, March 2025. URL: <https://www.sciencedirect.com/science/article/pii/S2949719124000724>, doi:10.1016/j.nlp.2024.100124.
- [17] Fengli Xu, Qianye Hao, Zefang Zong, Jingwei Wang, Yunke Zhang, Jingyi Wang, Xiaochong Lan, Jiahui Gong, Tianjian Ouyang, Fanjin Meng, Chenyang Shao, Yuwei Yan, Qinglong Yang, Yiwen Song, Sijian Ren, Xinyuan Hu, Yu Li, Jie Feng, Chen Gao, and Yong Li. Towards large reasoning models: A survey on scaling llm reasoning capabilities. *arXiv*, (arXiv:2501.09686), January 2025. arXiv:2501.09686. URL: <http://arxiv.org/abs/2501.09686>, doi:10.48550/arXiv.2501.09686.
- [18] Databricks. Llm inference performance engineering: Best practices. large language models (llms) generate text in a two-step process: "prefill", where the tokens in the input prompt are processed in parallel, and "decoding", where text is generated one 'token' at a time in an autoregressive manner., October 2023. URL: <https://tinyurl.com/3e7596at>.
- [19] Baseten. URL: <https://www.baseten.co/blog/llm-transformer-inference-guide/>.
- [20] Andrea Filippo Ferraris, Davide Audrito, Luigi Di Caro, and Cristina Poncibò. The architecture of language: Understanding the mechanics behind llms. *Cambridge Forum on AI: Law and Governance*, 1:e11, jan 2025. URL: <https://www.cambridge.org/core/journals/cambridge-forum-on-ai-law-and-governance/article/architecture-of-language-understanding-the-mechanics-behind-llms/E3DDEFB9C04883733380E04331D6F782>, doi:10.1017/cfl.2024.16.
- [21] Shangying Hua, Shuangci Jin, and Shengyi Jiang. The limitations and ethical considerations of chatgpt. *Data Intelligence*, 6(1):201–239, February 2024. URL: <https://direct.mit.edu/dint/article/6/1/201/118839/The-Limitations-and-Ethical-Considerations-of>, doi:10.1162/dint_a_00243.
- [22] Zixuan Zhou, Xuefei Ning, Ke Hong, Tianyu Fu, Jiaming Xu, Shiyao Li, Yuming Lou, Luning Wang, Zhihang Yuan, Xiuhong Li, Shengen Yan, Guohao Dai, Xiao-Ping Zhang, Yuhang Dong, and Yu Wang. A survey on efficient inference for large language models. *arXiv*, (arXiv:2404.14294), July 2024. arXiv:2404.14294. URL: <http://arxiv.org/abs/2404.14294>, doi:10.48550/arXiv.2404.14294.
- [23] Baolin Li, Yankai Jiang, Vijay Gadepally, and Devesh Tiwari. Llm inference serving: Survey of recent advances and opportunities. *arXiv*, (arXiv:2407.12391), July 2024. arXiv:2407.12391. URL: <http://arxiv.org/abs/2407.12391>, doi:10.48550/arXiv.2407.12391.
- [24] Claude. Anthropic is an ai safety and research company that's working to build reliable, interpretable, and steerable ai systems. enhanced reasoning, state-of-the-art coding skills, computer use, and 200k context window. URL: <https://www.anthropic.com/claude/sonnet>.
- [25] OpenAI. Openai o1 pricing. get the best of openai with the highest level of access. everything in plus. unlimited access to o1, o1-mini, gpt-4o, and voice (audio only.) higher limits for video and screensharing in voice., 2024. URL: <https://openai.com/chatgpt/pricing/>.
- [26] OpenAI. Openai flagship models. our versatile, high-intelligence flagship model. text and image input, text output. 128k context length. smarter model, higher price per token, 2024. URL: <https://platform.openai.com/docs/models/gpt-4o>.
- [27] Google Scholar. Google scholar provides a simple way to broadly search for scholarly literature. search across a wide variety of disciplines and sources: articles, theses, books, abstracts and court opinions. URL: <https://scholar.google.com/>.
- [28] Patrick M. Boland, Heinz-Josef Lenz, Kristen K. Ciombor, Vaia Florou, Michael J. Pishvaian, Michael Cusnir, Deirdre Cohen, Jessie Y. Guo, Min Tang, Prabhu Rajagopalan, Sandra E. Wiley, Richard G. Ghalie, and Howard S. Hochster. A phase 1b study of the oxphos inhibitor me-344 with bevacizumab in refractory metastatic colorectal cancer. *Investigational New Drugs*, December 2024. URL: <https://link.springer.com/10.1007/s10637-024-01489-1>, doi:10.1007/s10637-024-01489-1.
- [29] Daniele Boso, Iliara Piga, Chiara Trento, Sonia Minuzzo, Eleonora Angi, Luisa Iommarini, Elisabetta Lazzarini, Leonardo Caporali, Claudio Fiorini, Luigi D'Angelo, Monica De Luise, Ivana Kurelac, Matteo Fassan, Anna Maria Porcelli, Filippo Navaglia, Iliara Billato, Giovanni Esposito, Giuseppe Gasparre, Chiara Romualdi, and Stefano Indraccolo. Pathogenic mitochondrial dna variants are associated with response to anti-vegf therapy in ovarian cancer pdx models. *Journal of Experimental Clinical Cancer Research*, 43(1):325, December 2024. URL: <https://jccr.biomedcentral.com/articles/10.1186/s13046-024-03239-w>, doi:10.1186/s13046-024-03239-w.
- [30] Qiu-Zhong Pan, Jing-Jing Zhao, Liang Liu, Dong-Sheng Zhang, Li-Ping Wang, Wen-Wei Hu, De-Sheng Weng, Xiang Xu, Yi-Zhuo Li, Yan Tang, Wei-Hong Zhang, Jie-Yao Li, Xiao Zheng, Qi-Jing Wang, Yong-Qiang Li, Tong Xiang, Li Zhou, Shuang-Ning Yang, Chen Wu, Rong-Xing Huang, Jia He, Wei-Jiao Du, Lu-Jun Chen, Yue-Na Wu, Bin Xu, Qiong Shen, Yi Zhang, Jing-Ting Jiang, Xiu-Bao Ren, and Jian-Chuan Xia. Xelox (capecitabine plus oxaliplatin) plus bevacizumab

- (anti-vegf-a antibody) with or without adoptive cell immunotherapy in the treatment of patients with previously untreated metastatic colorectal cancer: a multicenter, open-label, randomized, controlled, phase 3 trial. *Signal Transduction and Targeted Therapy*, 9(1):79, April 2024. URL: <https://www.nature.com/articles/s41392-024-01788-2>, doi:10.1038/s41392-024-01788-2.
- [31] Effat Alemzadeh, Leila Allahqoli, Afrooz Mazidimoradi, Esmat Alemzadeh, Fahimeh Ghasemi, Hamid Salehiniya, and Ibrahim Alkatout. Deciphering resistance mechanisms and novel strategies to overcome drug resistance in ovarian cancer: a comprehensive review. *Oncology Research*, 32(5):831–847, 2024. URL: <https://www.techscience.com/or/v32n5/56240>, doi:10.32604/or.2024.031006.
- [32] Thomas Seufferlein, Ludwig Lausser, Alexander Stein, Dirk Arnold, Gerald Prager, Stefan Kasper-Virchow, Michael Niedermeier, Lothar Müller, Stefan Kubicka, Alexander König, Petra Büchner-Steudel, Kai Wille, Andreas W. Berger, Angelika M. R. Kestler, Johann M. Kraus, Silke D. Werle, Lukas Perkhofer, Thomas J. Ettrich, and Hans A. Kestler. Prediction of resistance to bevacizumab plus folfox in metastatic colorectal cancer—results of the prospective multicenter permad trial. *PLOS ONE*, 19(6):e0304324, June 2024. URL: <https://dx.plos.org/10.1371/journal.pone.0304324>, doi:10.1371/journal.pone.0304324.
- [33] Klaudia Żak, Małgorzata Satora, Ilona Skrabalak, Rafał Tarkowski, Marta Ostrowska-Leśko, and Marcin Bobiński. The potential influence of residual or recurrent disease on bevacizumab treatment efficacy in ovarian cancer: Current evidence and future perspectives. *Cancers*, 16(5):1063, March 2024. URL: <https://www.mdpi.com/2072-6694/16/5/1063>, doi:10.3390/cancers16051063.
- [34] Xiaoqian Chen, Wenkui Li, Xiaogai Lei, Zhanhong Li, Qijing Guo, Xinfu Ma, Yushuang Luo, and Liang Wang. Efficacy of immune checkpoint inhibitors combined with bevacizumab in mss/pmmr advanced colorectal cancer after first-line treatment failure. *Frontiers in Oncology*, 14:1429095, August 2024. URL: <https://www.frontiersin.org/articles/10.3389/fonc.2024.1429095/full>, doi:10.3389/fonc.2024.1429095.
- [35] Haining He and Fei Zhou. Efficacy and safety of anti-angiogenic drugs combined with chemotherapy in the treatment of platinum-sensitive/resistant ovarian cancer: a meta-analysis with trial sequential analysis of randomized controlled trials. *Frontiers in Pharmacology*, 15:1446403, November 2024. URL: <https://www.frontiersin.org/articles/10.3389/fphar.2024.1446403/full>, doi:10.3389/fphar.2024.1446403.
- [36] Pascale C.S. Rietveld, Niels A.D. Guchelaar, Ruben A.G. Van Eerden, Nadine L. De Boer, Peter De Bruijn, Sebastiaan D.T. Sassen, Eva V.E. Madsen, Birgit C.P. Koch, Cornelis Verhoef, Jacobus W.A. Burger, Ron H.J. Mathijssen, and Stijn L.W. Koolen. Intraperitoneal pharmacokinetics of systemic oxaliplatin, 5-fluorouracil and bevacizumab in patients with colorectal peritoneal metastases. *Biomedicine Pharmacotherapy*, 176:116820, July 2024. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0753332224007042>, doi:10.1016/j.biopha.2024.116820.
- [37] Yu Jin Kim, Hee Min Lee, Ga Eun Lee, Jin Hui Yoo, Hwa Jeong Lee, and Sandy Jeong Rhie. Optimizing outcomes: Bevacizumab with carboplatin and paclitaxel in 5110 ovarian cancer patients—a systematic review and meta-analysis. *Pharmaceuticals*, 17(8):1095, August 2024. URL: <https://www.mdpi.com/1424-8247/17/8/1095>, doi:10.3390/ph17081095.
- [38] Masashi Okawa, Munekazu Yamakuchi, Aryal Bibek, Kazunori Takenouchi, Drew N. Maywar, Shingo Yamada, Keiichi Inoue, Kazuhiko Higurashi, Junichi Nakazawa, Masahiro Kawahira, Tomoko Kodama, Kiyonori Tanoue, Yoko Oyama, Sadayuki Higashi, Chieko Fujisaki, Hirohito Hashinokuchi, Akito Tabaru, Hideaki Kanda, Shuji Tachioka, Yutaka Imoto, Teruto Hashiguchi, and Yoshiharu Soga. Plasma and serum concentrations of vegf-a121, but not of vegf-a165, increase post-bevacizumab administration. *PLOS ONE*, 19(12):e0316035, December 2024. URL: <https://dx.plos.org/10.1371/journal.pone.0316035>, doi:10.1371/journal.pone.0316035.
- [39] Ignacio Romero, Eva Guerra, Ainhua Madariaga, and Luis Manso. Safety of bevacizumab and olaparib as frontline maintenance therapy in advanced ovarian cancer: expert review for clinical practice. *Frontiers in Oncology*, 13:1304303, January 2024. URL: <https://www.frontiersin.org/articles/10.3389/fonc.2023.1304303/full>, doi:10.3389/fonc.2023.1304303.
- [40] Diana Cornelia Moisuc, Daniela Constantinescu, Mihai Vasile Marinca, Bogdan Gafton, Mariana Pavel-Tanasa, and Petru Cianga. Cyclophilin a: An independent prognostic factor for survival in patients with metastatic colorectal cancer treated with bevacizumab and chemotherapy. *Cancers*, 16(2):385, January 2024. URL: <https://www.mdpi.com/2072-6694/16/2/385>, doi:10.3390/cancers16020385.
- [41] Daniela Luvero, Roberto Angioli, Federica Celoro, Francesco Plotti, Corrado Terranova, Federica Guzzo, Gianna Barbara Cundari, Federico Liparulo, Camilla Verdone, and Roberto Montera. Tailored treatment strategies in first line therapy for ovarian cancer patients: A critical review of the literature. *Pharmaceuticals*, 17(6):778, June 2024. URL: <https://www.mdpi.com/1424-8247/17/6/778>, doi:10.3390/ph17060778.
- [42] Pei Zhang, Xiaofen Li, Xin Wang, Yu Yang, Jianfei Wang, and Dan Cao. Shr-8068 combined with adabrelimab and bevacizumab in the treatment of refractory advanced colorectal cancer: study protocol for a single-arm, phase ib/ii study. *Frontiers in Immunology*, 15:1450533, October 2024. URL: <https://www.frontiersin.org/articles/10.3389/fimmu.2024.1450533/full>, doi:10.3389/fimmu.2024.1450533.

- [43] Mariana Nunes, Carla Bartosch, Miguel Henriques Abreu, Alan Richardson, Raquel Almeida, and Sara Ricardo. Deciphering the molecular mechanisms behind drug resistance in ovarian cancer to unlock efficient treatment options. *Cells*, 13(9):786, May 2024. URL: <https://www.mdpi.com/2073-4409/13/9/786>, doi:10.3390/cells13090786.
- [44] Sheng Li, Xiaoyou Li, Hanfeng Xu, Jiayuan Huang, Jingni Zhu, Ying Peng, Jun Bao, and Liangjun Zhu. Alternating modified capox/capiri plus bevacizumab in untreated unresectable metastatic colorectal cancer: a phase 2 trial. *Signal Transduction and Targeted Therapy*, 9(1):346, December 2024. URL: <https://www.nature.com/articles/s41392-024-02048-z>, doi:10.1038/s41392-024-02048-z.
- [45] Grace Y. Q. Han, Monica Alexander, Julia Gattozzi, Marilyn Day, Elayna Kirsch, Narges Tafreshi, Raafat Chalar, Soraya Rahni, Gabrielle Gossner, William Burke, and Mehdi Damaghi. Ecological and evolutionary dynamics to design and improve ovarian cancer treatment. *Clinical and Translational Medicine*, 14(9):e70012, September 2024. URL: <https://onlinelibrary.wiley.com/doi/10.1002/ctm2.70012>, doi:10.1002/ctm2.70012.
- [46] Romain Chautard, Morgane Caulet, Olivier Bouché, Christophe Borg, Sylvain Manfredi, Olivier Capitain, Jean-Philippe Spano, William Raoul, Maxime Guéguinou, Olivier Herault, Aurélie Ferru, Cédric Pobel, Olivier Sire, and Thierry Lecomte. Evaluation of serum mid-infrared spectroscopy as new prognostic marker for first-line bevacizumab-based chemotherapy in metastatic colorectal cancer. *Digestive and Liver Disease*, 57(1):141–148, January 2025. URL: <https://linkinghub.elsevier.com/retrieve/pii/S1590865824008867>, doi:10.1016/j.dld.2024.07.022.
- [47] Yosuke Murata, Takuto Nosaka, Yu Akazawa, Tomoko Tanaka, Kazuto Takahashi, Tatsushi Naito, Hidetaka Matsuda, Masahiro Ohtani, and Yasunari Nakamoto. Serial changes of circulating tumor cells in patients with hepatocellular carcinoma treated with atezolizumab plus bevacizumab. *Cancers*, 16(13):2410, June 2024. URL: <https://www.mdpi.com/2072-6694/16/13/2410>, doi:10.3390/cancers16132410.
- [48] Zhuoling Zheng, Yihong Zhao, Jingwen Xie, Min Gao, Yiting Wang, and Xiaoyan Li. Clinical risk factors of bevacizumab-related hypertension in patients with metastatic colorectal cancer: a retrospective study. *Frontiers in Pharmacology*, 15:1463026, October 2024. URL: <https://www.frontiersin.org/articles/10.3389/fphar.2024.1463026/full>, doi:10.3389/fphar.2024.1463026.
- [49] Takuto Nosaka, Yosuke Murata, Yu Akazawa, Tomoko Tanaka, Kazuto Takahashi, Tatsushi Naito, Hidetaka Matsuda, Masahiro Ohtani, Yoshiaki Imamura, and Yasunari Nakamoto. Programmed death ligand 1 expression in circulating tumor cells as a predictor and monitor of response to atezolizumab plus bevacizumab treatment in patients with hepatocellular carcinoma. *Cancers*, 16(9):1785, May 2024. URL: <https://www.mdpi.com/2072-6694/16/9/1785>, doi:10.3390/cancers16091785.
- [50] Kei Furuya, Masao Nakajima, Ryouichi Tsunedomi, Yuki Nakagami, Ming Xu, Hiroto Matsui, Yukio Tokumitsu, Yoshitaro Shindo, Yusaku Watanabe, Shinobu Tomochika, Noriko Maeda, Michihisa Iida, Nobuaki Suzuki, Shigeru Takeda, Shoichi Hazama, Tatsuya Ioka, Yoshinobu Hoshii, Tomio Ueno, and Hiroaki Nagano. High serum proteinase-3 levels predict poor progression-free survival and lower efficacy of bevacizumab in metastatic colorectal cancer. *BMC Cancer*, 24(1):165, February 2024. URL: <https://bmccancer.biomedcentral.com/articles/10.1186/s12885-024-11924-4>, doi:10.1186/s12885-024-11924-4.
- [51] Masako Shomura, Haruka Okabe, Maya Sakakibara, Emi Sato, Koichi Shiraishi, Yoshitaka Arase, Kota Tsuruya, Yusuke Mishima, Shunji Hirose, and Tatehiro Kagawa. Impact of atezolizumab + bevacizumab therapy on health-related quality of life in patients with advanced hepatocellular carcinoma. *Cancers*, 16(21):3610, October 2024. URL: <https://www.mdpi.com/2072-6694/16/21/3610>, doi:10.3390/cancers16213610.
- [52] Thibault Mazard, Caroline Mollevi, Evelyne M. Loyer, Julie Léger, Romain Chautard, Olivier Bouché, Christophe Borg, Paul Armand-Dujardin, Aurore Bleuzen, Eric Assenat, and Thierry Lecomte. Prognostic value of the tumor-to-liver density ratio in patients with metastatic colorectal cancer treated with bevacizumab-based chemotherapy: a post-hoc study of the stic-avastin trial. *Cancer Imaging*, 24(1):77, June 2024. URL: <https://cancerimagingjournal.biomedcentral.com/articles/10.1186/s40644-024-00722-7>, doi:10.1186/s40644-024-00722-7.
- [53] Sang Youn Hwang, Hyun Young Woo, Jeong Heo, Hyung Jun Kim, Young Joo Park, Ki Youn Yi, Yu Rim Lee, Soo Young Park, Woo Jin Chung, Byoung Kuk Jang, and Won Young Tak. Outcome of atezolizumab plus bevacizumab combination therapy in high-risk patients with advanced hepatocellular carcinoma. *Cancers*, 16(4):838, February 2024. URL: <https://www.mdpi.com/2072-6694/16/4/838>, doi:10.3390/cancers16040838.
- [54] Lijuan He, Haiyuan Li, Yunpeng Wang, Weidong Li, Lei Gao, Bo Xu, Jike Hu, Puyi He, Weigao Pu, Guodong Sun, Zhuanfang Wang, Qinying Han, Ben Liu, and Hao Chen. Complete remission in a pretreated, microsatellite-stable, kras-mutated colon cancer patient after treatment with sintilimab and bevacizumab and platinum-based chemotherapy: a case report and literature review. *Frontiers in Immunology*, 15:1354613, March 2024. URL: <https://www.frontiersin.org/articles/10.3389/fimmu.2024.1354613/full>, doi:10.3389/fimmu.2024.1354613.
- [55] Michael H Storandt, Tyler J Zemla, Kanchi Patell, Nikolas Naleid, Jennifer J Gile, Nguyen H Tran, Sakti Chakrabarti, Zhaohui Jin, Mitesh Borad, and Amit Mahipal. Atezolizumab plus bevacizumab as first-line systemic therapy for hepatocellular carcinoma: a multi-institutional cohort study. *The Oncologist*, 29(11):986–996, November 2024. URL: <https://academic.oup.com/oncolo/article/29/11/986/7709605>, doi:10.1093/oncolo/oyae142.

- [56] Lei Yan, Jia Shi, and Jiazuo Zhu. Cellular and molecular events in colorectal cancer: biological mechanisms, cell death pathways, drug resistance and signalling network interactions. *Discover Oncology*, 15(1):294, July 2024. URL: <https://link.springer.com/10.1007/s12672-024-01163-1>, doi:10.1007/s12672-024-01163-1.
- [57] Zili Hu, Zhoutian Yang, Zibo Fu, Yizhen Fu, Zeliang Ma, Zhongguo Zhou, Minshan Chen, and Yaojun Zhang. Efficacy and safety of atezolizumab–bevacizumab vs pembrolizumab–lenvatinib in unresectable hepatocellular carcinoma: a retrospective, cohort study. *Frontiers in Immunology*, 15:1472870, November 2024. URL: <https://www.frontiersin.org/articles/10.3389/fimmu.2024.1472870/full>, doi:10.3389/fimmu.2024.1472870.
- [58] Chengyu Sun, Enguo Fan, Luqiao Huang, and Zhengguo Zhang. Second-line systemic treatment for metastatic colorectal cancer: A systematic review and bayesian network meta-analysis based on rct. *PLOS ONE*, 19(12):e0313278, December 2024. URL: <https://dx.plos.org/10.1371/journal.pone.0313278>, doi:10.1371/journal.pone.0313278.
- [59] Najib Ben Khaled, Marie Möller, Leonie S. Jochheim, Catherine Leyh, Ursula Ehmer, Katrin Böttcher, Matthias Pinter, Lorenz Balcar, Bernhard Scheiner, Alexander Weich, Hans Benno Leicht, Valentina Zarka, Liangtao Ye, Julia Schneider, Ignazio Piseddu, Osman Öcal, Monika Rau, Friedrich Sinner, Marino Venerito, Simon Johannes Gairing, Friedrich Förster, Julia Mayerle, Enrico N. De Toni, Andreas Geier, and Florian P. Reiter. Atezolizumab/bevacizumab or lenvatinib in hepatocellular carcinoma: Multicenter real-world study with focus on bleeding and thromboembolic events. *JHEP Reports*, 6(6):101065, June 2024. URL: <https://linkinghub.elsevier.com/retrieve/pii/S2589555924000661>, doi:10.1016/j.jhepr.2024.101065.
- [60] Yu Song, Qianqian Mao, Manling Zhou, Cheng-Jiang Liu, Li Kong, and Ting Hu. Effectiveness of bevacizumab in the treatment of metastatic colorectal cancer: a systematic review and meta-analysis. *BMC Gastroenterology*, 24(1):58, February 2024. URL: <https://bmcgastroenterol.biomedcentral.com/articles/10.1186/s12876-024-03134-w>, doi:10.1186/s12876-024-03134-w.
- [61] Masayuki Ueno, Haruhiko Takeda, Atsushi Takai, Hiroki Morimura, Norihiro Nishijima, Satoru Iwamoto, Shunsuke Okuyama, Makoto Umeda, Takeshi Seta, Atsuyuki Ikeda, Tomoyuki Goto, Shin'ichi Miyamoto, Takahisa Kayahara, Yoshito Uenoyama, Kazuyoshi Matsumura, Shigeharu Nakano, Masako Mishima, Tadashi Inuzuka, Yuji Eso, Ken Takahashi, Hiroyuki Marusawa, Yukio Osaki, Etsuro Hatano, and Hiroshi Seno. Crafty score as a predictive marker for refractoriness to atezolizumab plus bevacizumab therapy in hepatocellular carcinoma: a multicenter retrospective study. *Journal of Gastroenterology*, 59(12):1107–1118, December 2024. URL: <https://link.springer.com/10.1007/s00535-024-02150-7>, doi:10.1007/s00535-024-02150-7.
- [62] Thierry André, Eric Van Cutsem, Julien Taieb, Marwan Fakih, Gerald W. Prager, Fortunato Ciardiello, Alfredo Falcone, Mark Saunders, Nadia Amellal, Lucas Roby, Josep Tabernero, and Per Pfeiffer. Clinical trial data review of the combination ftid/tpi + bevacizumab in the treatment landscape of unresectable metastatic colorectal cancer. *Current Treatment Options in Oncology*, 25(10):1312–1322, October 2024. URL: <https://link.springer.com/10.1007/s11864-024-01261-w>, doi:10.1007/s11864-024-01261-w.
- [63] Yuki Shirane, Yasutoshi Fujii, Atsushi Ono, Hikaru Nakahara, Clair Nelson Hayes, Ryoichi Miura, Serami Murakami, Naoya Sakamoto, Shinsuke Uchikawa, Hatsue Fujino, Takashi Nakahara, Eisuke Murakami, Masami Yamauchi, Daiki Miki, Tomokazu Kawaoka, Koji Arihiro, Masataka Tsuge, and Shiro Oka. Peripheral t cell subpopulations as a potential surrogate biomarker during atezolizumab plus bevacizumab treatment for hepatocellular carcinoma. *Cancers*, 16(7):1328, March 2024. URL: <https://www.mdpi.com/2072-6694/16/7/1328>, doi:10.3390/cancers16071328.
- [64] Giuseppe Corrias, Eleonora Lai, Pina Ziranu, Stefano Mariani, Clelia Donisi, Nicole Liscia, Giorgio Saba, Andrea Pretta, Mara Persano, Daniela Fanni, Dario Spanu, Francesca Balconi, Francesco Loi, Simona Deidda, Angelo Restivo, Valeria Pusceddu, Marco Puzzoni, Cinzia Solinas, Elena Massa, Clelia Madeddu, Clara Gerosa, Luigi Zorcolo, Gavino Faa, Luca Saba, and Mario Scartozzi. Prediction of response to anti-angiogenic treatment for advanced colorectal cancer patients: From biological factors to functional imaging. *Cancers*, 16(7):1364, March 2024. URL: <https://www.mdpi.com/2072-6694/16/7/1364>, doi:10.3390/cancers16071364.
- [65] Mara Egerer, Kathrin Schuch, David Schöler, Fabian Artusa, Tobias Püngel, Theresa Maria Holtman, Sven H. Loosen, Münevver Demir, Alexander Wree, Tom Luedde, Frank Tacke, Christoph Roderburg, and Raphael Mohr. Extracellular vesicles may predict response to atezolizumab plus bevacizumab in patients with advanced hepatocellular carcinoma. *Cancers*, 16(21):3651, October 2024. URL: <https://www.mdpi.com/2072-6694/16/21/3651>, doi:10.3390/cancers16213651.
- [66] Anne Jacobsen, Jürgen Siebler, Robert Grützmann, Michael Stürzl, and Elisabeth Naschberger. Blood vessel-targeted therapy in colorectal cancer: Current strategies and future perspectives. *Cancers*, 16(5):890, February 2024. URL: <https://www.mdpi.com/2072-6694/16/5/890>, doi:10.3390/cancers16050890.
- [67] Shalin Abraham and Adel Samson. Case report: Successful treatment of a patient undergoing haemodialysis with multifocal hepatocellular carcinoma using atezolizumab and bevacizumab. *Frontiers in Oncology*, 13:1279501, January 2024. URL: <https://www.frontiersin.org/articles/10.3389/fonc.2023.1279501/full>, doi:10.3389/fonc.2023.1279501.
- [68] Tehnia Naz, Anees Ur Rehman, Aleena Shahzad, Muhammad Fawad Rasool, Zikria Saleem, and Rabia Hussain. Impact of bevacizumab on clinical outcomes and its comparison with standard chemotherapy in metastatic colorectal cancer patients:

- a systematic review and meta-analysis. *Journal of Pharmaceutical Policy and Practice*, 17(1):2354300, December 2024. URL: <https://www.tandfonline.com/doi/full/10.1080/20523211.2024.2354300>, doi:10.1080/20523211.2024.2354300.
- [69] Kamila Krupa, Marta Fudalej, Anna Cencelewicz-Lesikow, Anna Badowska-Kozakiewicz, Aleksandra Czerw, and Andrzej Deptała. Current treatment methods in hepatocellular carcinoma. *Cancers*, 16(23):4059, December 2024. URL: <https://www.mdpi.com/2072-6694/16/23/4059>, doi:10.3390/cancers16234059.
- [70] Taruba Rais, Rumaisa Riaz, Tasmiyah Siddiqui, Amna Shakeel, Afsheen Khan, and Habiba Zafar. Innovations in colorectal cancer treatment: trifluridine and tipiracil with bevacizumab for improved outcomes – a review. *Frontiers in Oncology*, 14:1296765, July 2024. URL: <https://www.frontiersin.org/articles/10.3389/fonc.2024.1296765/full>, doi:10.3389/fonc.2024.1296765.
- [71] Kyoko Oura, Asahiro Morishita, Tomoko Tadokoro, Koji Fujita, Joji Tani, and Hideki Kobara. Immune microenvironment and the effect of vascular endothelial growth factor inhibition in hepatocellular carcinoma. *International Journal of Molecular Sciences*, 25(24):13590, December 2024. URL: <https://www.mdpi.com/1422-0067/25/24/13590>, doi:10.3390/ijms252413590.
- [72] F. De Felice, L. Archetti, G. D'Ambrosio, F. Iafrate, V. Picone, F.M. Magliocca, D. Musio, M. Roberto, G. Casella, I. Clementi, N. Bulzonetti, A. Picchetto, E. Vitti, E. Merenda, C. Gentili, M. Lanzilao, M. Miccini, G. Illuminati, A. Delle Donne, D. Crocetti, M. Minozzi, M. Mongardini, R. Caronna, E. Fiori, and E. Cortesi. Intensified total neoadjuvant therapy in patients with locally advanced rectal cancer: Long-term results of a prospective phase ii study. *Clinical Oncology*, 37:103698, January 2025. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0936655524004965>, doi:10.1016/j.clon.2024.103698.
- [73] Alexa Childs, Gloryanne Aidoo-Micah, Mala K. Maini, and Tim Meyer. Immunotherapy for hepatocellular carcinoma. *JHEP Reports*, 6(10):101130, October 2024. URL: <https://linkinghub.elsevier.com/retrieve/pii/S2589555924001344>, doi:10.1016/j.jhepr.2024.101130.
- [74] Mingming Su, Lili Zhao, Jing Zhou, Xuan Li, and Ning Ding. Celiac trunk aortic dissection induced by bevacizumab therapy for rectal cancer: A case report. *Medicine*, 103(28):e38882, July 2024. URL: <https://journals.lww.com/10.1097/MD.00000000000038882>, doi:10.1097/MD.00000000000038882.
- [75] Kaina Chen, Timothy W. H. Shuen, and Pierce K. H. Chow. The association between tumour heterogeneity and immune evasion mechanisms in hepatocellular carcinoma and its clinical implications. *British Journal of Cancer*, 131(3):420–429, August 2024. URL: <https://www.nature.com/articles/s41416-024-02684-w>, doi:10.1038/s41416-024-02684-w.
- [76] Qinlan Xu and Dong Shao. Leveraging the synergy between anti-angiogenic therapy and immune checkpoint inhibitors to treat digestive system cancers. *Frontiers in Immunology*, 15:1487610, December 2024. URL: <https://www.frontiersin.org/articles/10.3389/fimmu.2024.1487610/full>, doi:10.3389/fimmu.2024.1487610.
- [77] José E. Valerio, Aizik L. Wolf, Penelope Mantilla-Farfan, Guillermo De Jesús Aguirre Vera, María P. Fernández-Gómez, and Andrés M. Alvarez-Pinzon. Efficacy and cognitive outcomes of gamma knife radiosurgery in glioblastoma management for elderly patients. *Journal of Personalized Medicine*, 14(10):1049, October 2024. URL: <https://www.mdpi.com/2075-4426/14/10/1049>, doi:10.3390/jpm14101049.
- [78] Shang-Gin Wu, Chao-Chi Ho, James Chih-Hsin Yang, Shu-Han Yu, Yen-Feng Lin, Shu-Chin Lin, Bin-Chi Liao, Ching-Yao Yang, Yen-Ting Lin, Chong-Jen Yu, Ya-Ting Chuang, Wei-Yu Liao, Kah Yi Yap, Weng Si Kou, and Jin-Yuan Shih. Atezolizumab, bevacizumab, pemetrexed and platinum for egfr -mutant nsccl patients after egfr tki failure: A phase ii study with immune cell profile analysis. *Clinical and Translational Medicine*, 15(1):e70149, January 2025. URL: <https://onlinelibrary.wiley.com/doi/10.1002/ctm2.70149>, doi:10.1002/ctm2.70149.
- [79] Guangzhong Guo, Ziyue Zhang, Jiubing Zhang, Dayang Wang, Sensen Xu, Guanzheng Liu, Yushuai Gao, Jie Mei, Zhaoyue Yan, Ruijiao Zhao, Meiyun Wang, Tianxiao Li, and Xingyao Bu. Predicting recurrent glioblastoma clinical outcome to immune checkpoint inhibition and low-dose bevacizumab with tumor in situ fluid circulating tumor dna analysis. *Cancer Immunology, Immunotherapy*, 73(10):193, August 2024. URL: <https://link.springer.com/10.1007/s00262-024-03774-7>, doi:10.1007/s00262-024-03774-7.
- [80] Yan Huang, Yunpeng Yang, Yuanyuan Zhao, Hongyun Zhao, Ningning Zhou, Yaxiong Zhang, Likun Chen, Ting Zhou, Gang Chen, Ting Wu, Lu Lu, Shilin Xue, Xiaoyan Kang, Li Zhang, and Wenfeng Fang. Q11706 (anti-pd-1 igg4/ctla-4 antibody) plus chemotherapy with or without bevacizumab in advanced non-small cell lung cancer: a multi-cohort, phase ii study. *Signal Transduction and Targeted Therapy*, 9(1):23, January 2024. URL: <https://www.nature.com/articles/s41392-023-01731-x>, doi:10.1038/s41392-023-01731-x.
- [81] Jeff F. Zhang, Bernard Okai, Austin Iovoli, Victor Goulenko, Kristopher Attwood, Jaims Lim, Ryan M. Hess, Ajay P. Abad, Dheerendra Prasad, and Robert A. Fenstermaker. Bevacizumab and gamma knife radiosurgery for first-recurrence glioblastoma. *Journal of Neuro-Oncology*, 166(1):89–98, January 2024. URL: <https://link.springer.com/10.1007/s11060-023-04524-y>, doi:10.1007/s11060-023-04524-y.
- [82] Yihui Ge, Juan Li, Wenjing Gong, Jian Wang, Xiaojuan Wei, Jing Liu, Shuyun Wang, Leirong Wang, Haifeng Sun, Qinglei Cheng, Yanxin Sun, Qi Dang, Yuping Sun, and Aiqin Gao. Efficacy of first-line treatment options beyond ret-tkis in

- advanced ret-rearranged non-small cell lung cancer: A multi-center real-world study. *Cancer Medicine*, 13(2):e6960, January 2024. URL: <https://onlinelibrary.wiley.com/doi/10.1002/cam4.6960>, doi:10.1002/cam4.6960.
- [83] Taketo Ezaki, Toshihide Tanaka, Ryota Tamura, Kentaro Ohara, Yohei Yamamoto, Jun Takei, Yukina Morimoto, Ryotaro Imai, Yuki Kuranai, Yasuharu Akasaki, Masahiro Toda, Yuichi Murayama, Keisuke Miyake, and Hikaru Sasaki. Status of alternative angiogenic pathways in glioblastoma resected under and after bevacizumab treatment. *Brain Tumor Pathology*, 41(2):61–72, April 2024. URL: <https://link.springer.com/10.1007/s10014-024-00481-0>, doi:10.1007/s10014-024-00481-0.
- [84] Marcus Rathbone, Conor O'Hagan, Helen Wong, Adeel Khan, Timothy Cook, Sarah Rose, Jonathan Heseltine, and Carles Escriu. Intracranial efficacy of atezolizumab, bevacizumab, carboplatin, and paclitaxel in real-world patients with non-small-cell lung cancer and egfr or alk alterations. *Cancers*, 16(7):1249, March 2024. URL: <https://www.mdpi.com/2072-6694/16/7/1249>, doi:10.3390/cancers16071249.
- [85] Wangyan Zhong, Jiwei Mao, Dongping Wu, Jianghua Peng, and Wanli Ye. The efficacy of stereotactic radiotherapy followed by bevacizumab and temozolomide in the treatment of recurrent glioblastoma: a case report. *Frontiers in Pharmacology*, 15:1401000, September 2024. URL: <https://www.frontiersin.org/articles/10.3389/fphar.2024.1401000/full>, doi:10.3389/fphar.2024.1401000.
- [86] Xinyi Zhang, Xiaofei Chu, Jun Wang, Wenjing Gu, Xiaoyan Fu, Jing Zhang, Congying Wang, Qinghui Han, Jing Zhou, Yongjing Zhang, and Xiaoyan Liu. Retrospective analysis of survival and safety of bevacizumab biosimilar and original drug combination chemotherapy in non-small cell lung cancer. *Frontiers in Oncology*, 14:1437762, October 2024. URL: <https://www.frontiersin.org/articles/10.3389/fonc.2024.1437762/full>, doi:10.3389/fonc.2024.1437762.
- [87] Yosef Laviv, Ohad Regev, Andrew A. Kanner, Susana Fichman, Dror Limon, Tali Siegal, Shlomit Yust-Katz, and Alexandra Benouaich-Amiel. Stem the blood flow: beneficial impact of bevacizumab on survival of subventricular zone glioblastoma patients. *Journal of Neuro-Oncology*, 171(1):201–211, January 2025. URL: <https://link.springer.com/10.1007/s11060-024-04828-7>, doi:10.1007/s11060-024-04828-7.
- [88] Xin Liao, Tingting He, Xiong Wan, Pian Liu, Jing Li, Yong He, and Yubo Wang. Clinical management in nscl patients with egfr mutation after osimertinib progression with unknown resistance mechanisms. *The Clinical Respiratory Journal*, 18(10):e70025, October 2024. URL: <https://onlinelibrary.wiley.com/doi/10.1111/crj.70025>, doi:10.1111/crj.70025.
- [89] Md Ataur Rahman and Meser M. Ali. Recent treatment strategies and molecular pathways in resistance mechanisms of antiangiogenic therapies in glioblastoma. *Cancers*, 16(17):2975, August 2024. URL: <https://www.mdpi.com/2072-6694/16/17/2975>, doi:10.3390/cancers16172975.
- [90] Chia-Yu Kuo, Ming-Ju Tsai, Jen-Yu Hung, Mei-Hsuan Lee, Kuan-Li Wu, Yu-Chen Tsai, Cheng-Hao Chuang, Chung-Wen Huang, Chin-Ling Chen, Chih-Jen Yang, and Inn-Wen Chong. Clinical outcome of bevacizumab or ramucirumab combined with epidermal growth factor receptor (egfr) tyrosine kinase inhibitors as the first line therapy in susceptible egfr-mutated advanced non-small-cell lung. *The Kaohsiung Journal of Medical Sciences*, 40(5):467–476, May 2024. URL: <https://onlinelibrary.wiley.com/doi/10.1002/kjm2.12822>, doi:10.1002/kjm2.12822.
- [91] Linde Hoosmans, Marc Vooijs, and Ann Hoebe. Opportunities and challenges of small molecule inhibitors in glioblastoma treatment: Lessons learned from clinical trials. *Cancers*, 16(17):3021, August 2024. URL: <https://www.mdpi.com/2072-6694/16/17/3021>, doi:10.3390/cancers16173021.
- [92] Jin Xiong and Lei Xia. Case report: Precise ngs and combined bevacizumab promote durable response in alk-positive lung adenocarcinoma with multiple-line alk-tyki resistance. *Frontiers in Oncology*, 14:1419306, June 2024. URL: <https://www.frontiersin.org/articles/10.3389/fonc.2024.1419306/full>, doi:10.3389/fonc.2024.1419306.
- [93] Karol Sadowski, Adrianna Jazdzewska, Jan Kozłowski, Aleksandra Zacny, Tomasz Lorenc, and Wioletta Olejarz. Revolutionizing glioblastoma treatment: A comprehensive overview of modern therapeutic approaches. *International Journal of Molecular Sciences*, 25(11):5774, May 2024. URL: <https://www.mdpi.com/1422-0067/25/11/5774>, doi:10.3390/ijms25115774.
- [94] Chunyu Hua, Ziyue Guo, Meng Dai, Jie Zhou, Hanxiao Ge, Guoqing Xue, Fahui Xu, Liyuan Ru, Kuan Lv, Guohui Zhang, Lina Zheng, Meiyi Wang, Yun Teng, Wendan Yu, and Wei Guo. Lumbrokinase extracted from earthworms synergizes with bevacizumab and chemotherapeutics in treating non-small cell lung cancer by targeted inactivation of bptf/vegfr and nf-b/cox-2 signaling. *Biomolecules*, 14(7):741, June 2024. URL: <https://www.mdpi.com/2218-273X/14/7/741>, doi:10.3390/biom14070741.
- [95] Hua Yang, Shi Gao Huang, Mohan Dong, Xiaomeng Wang, JunHua He, Huyan Su, Changhao Liu, Yong Zhu, Lichun Wei, and Zi Liu. Efficacy and safety of bevacizumab in neoadjuvant and concurrent chemoradiotherapy for refractory cervical cancer patients. *Biomolecules and Biomedicine*, 24(6):1586–1594, October 2024. URL: <https://bjbms.org/ojs/index.php/bjbms/article/view/10528>, doi:10.17305/bb.2024.10528.
- [96] Xiuhua Lin, Shichuan Tang, Yutong Guo, Ruijing Tang, Zhenli Li, Xinting Pan, Geng Chen, Liman Qiu, Xiuqing Dong, Ling Zhang, Xiaolong Liu, Zhixiong Cai, and Baosong Xie. Personalized neoantigen vaccine enhances the therapeutic efficacy of bevacizumab and anti-pd-1 antibody in advanced non-small cell lung cancer. *Cancer Immunology*,

Immunotherapy, 73(2):26, January 2024. URL: <https://link.springer.com/10.1007/s00262-023-03598-x>, doi:10.1007/s00262-023-03598-x.

- [97] Mingxiu Ju, Baoyue Pan, Yongwen Huang, Yun Zhou, Jieping Chen, Huiling Xiang, Shije Xu, Siyu Chen, Chunyan Lan, Jundong Li, and Min Zheng. The efficacy of first and second immunotherapy exposure in patients with recurrent or metastatic cervical cancer. *Cancer Medicine*, 13(19):e70204, October 2024. URL: <https://onlinelibrary.wiley.com/doi/10.1002/cam4.70204>, doi:10.1002/cam4.70204.
- [98] Ulrich Jarry, Megane Bostoen, Jérôme Archambeau, Raphaël Pineau, Laura Chaillot, Florence Jouan, Hélène Solhi, Hugo Ferrari, Rémy Le Guevel, Valentine Mennessier, Hervé Lena, Romain Corre, Charles Ricordel, Thierry Guillaudeux, and Rémy Pedoux. Afatinib or bevacizumab in combination with osimertinib efficiently control tumor development in orthotopic murine models of non-small lung cancer. *PLOS ONE*, 19(6):e0304914, June 2024. URL: <https://dx.plos.org/10.1371/journal.pone.0304914>, doi:10.1371/journal.pone.0304914.
- [99] Takeya Adachi, Tomoya Matsui, Utako Okata-Karigane, Chiaki Takahashi, Umi Tahara, Mari Hyodo, Akihiro Miyagawa, Kenta Kobayashi, Yoshio Nakamura, Takeru Funakoshi, Hiroshi Nishio, Wataru Yamagami, and Hayato Takahashi. Delayed and immediate cutaneous adverse events during pembrolizumab combination chemotherapy against cervical cancer: Case series. *The Journal of Dermatology*, 52(1):132–137, January 2025. URL: <https://onlinelibrary.wiley.com/doi/10.1111/1346-8138.17521>, doi:10.1111/1346-8138.17521.
- [100] Shanshan Huang, Yaling Long, Yuan Gao, Wanling Lin, Lei Wang, Jizong Jiang, Xun Yuan, Yuan Chen, Peng Zhang, and Qian Chu. Combined inhibition of met and vegf enhances therapeutic efficacy of egfr tkis in egfr-mutant non-small cell lung cancer with concomitant aberrant met activation. *Experimental Hematology Oncology*, 13(1):97, October 2024. URL: <https://ehoonline.biomedcentral.com/articles/10.1186/s40164-024-00565-9>, doi:10.1186/s40164-024-00565-9.
- [101] Chongya Zhai, Xuanhong Jin, Liangkun You, Na Yan, Jie Dong, Sai Qiao, Yuhong Zhong, Yu Zheng, and Hongming Pan. Hemophagocytic lymphohistiocytosis following pembrolizumab and bevacizumab combination therapy for cervical cancer: a case report and systematic review. *BMC Geriatrics*, 24(1):32, January 2024. URL: <https://bmgeriatr.biomedcentral.com/articles/10.1186/s12877-023-04625-3>, doi:10.1186/s12877-023-04625-3.
- [102] Nobuaki Kobayashi, Seigo Katakura, Nobuhiko Fukuda, Kohei Somekawa, Ayami Kaneko, and Takeshi Kaneko. The impact of bevacizumab and mir200c on emt and egfr-tki resistance in egfr-mutant lung cancer organoids. *Genes*, 15(12):1624, December 2024. URL: <https://www.mdpi.com/2073-4425/15/12/1624>, doi:10.3390/genes15121624.
- [103] Che-Wei Chang, Szu-Ting Yang, Hung-Hsien Liu, Wen-Hsun Chang, Wen-Ling Lee, and Peng-Hui Wang. Surgery-based radiation-free multimodality treatment for locally advanced cervical cancer. *Taiwanese Journal of Obstetrics and Gynecology*, 63(5):651–664, September 2024. URL: <https://linkinghub.elsevier.com/retrieve/pii/S1028455924002092>, doi:10.1016/j.tjog.2024.07.014.
- [104] Jin Tian, Zhiqi Shi, Lili Zhao, Peng Liu, Xiaojun Sun, Lin Long, Jianhua Zang, and Jun Xiao. Revolutionizing nsclc treatment: Immunotherapy strategies for egfr-tkis resistance. *The Clinical Respiratory Journal*, 18(12):e70037, December 2024. URL: <https://onlinelibrary.wiley.com/doi/10.1111/crj.70037>, doi:10.1111/crj.70037.
- [105] Shuangshuang Zhao, Minmin Zhang, Qing Zhang, Jingjun Wu, and Hui Dai. Anlotinib alone or in combination with bevacizumab in the treatment of recurrent high-grade glioma: a prospective single-arm, open-label phase ii trial. *BMC Cancer*, 24(1):6, January 2024. URL: <https://bmccancer.biomedcentral.com/articles/10.1186/s12885-023-11776-4>, doi:10.1186/s12885-023-11776-4.
- [106] Guangsu Han, Chenlu Li, and Ping Yi. The efficacy of bevacizumab combined with platinum-containing chemotherapy in the treatment of advanced non-small cell lung cancer in china: a systematic review and meta-analysis of randomized clinical trials. *Frontiers in Pharmacology*, 15:1293039, January 2024. URL: <https://www.frontiersin.org/articles/10.3389/fphar.2024.1293039/full>, doi:10.3389/fphar.2024.1293039.
- [107] Lin Shen, Shuai Li, Yalin Wang, Yi Yin, Yiting Liu, Yunlei Zhang, and Xuesheng Zheng. Alternating electric fields transform the intricate network of tumour vasculature into orderly parallel capillaries and enhance the anti-angiogenesis effect of bevacizumab. *Cell Proliferation*, 58(1):e13734, January 2025. URL: <https://onlinelibrary.wiley.com/doi/10.1111/cpr.13734>, doi:10.1111/cpr.13734.
- [108] Qi He, Xiaoyan Liu, Liyan Jiang, Ping Liu, Weixia Xuan, Yudong Wang, Rui Meng, Huijing Feng, Shuang Lv, Qian Miao, Di Zheng, Yan Xu, and Mengzhao Wang. First-line treatments for kras-mutant non-small cell lung cancer: current state and future perspectives. *Cancer Biology Therapy*, 26(1):2441499, December 2025. URL: <https://www.tandfonline.com/doi/full/10.1080/15384047.2024.2441499>, doi:10.1080/15384047.2024.2441499.
- [109] SiYao Wei, LanYin Chang, and Yi Zhong. The efficacy and adverse events of bevacizumab combined with temozolomide in the treatment of glioma: a systemic review and meta-analysis of randomized controlled trials. *Frontiers in Medicine*, 11:1419038, July 2024. URL: <https://www.frontiersin.org/articles/10.3389/fmed.2024.1419038/full>, doi:10.3389/fmed.2024.1419038.
- [110] Yuanyuan Zhao, Ting Mei, Feifei Na, Xiaoman Tian, Rui Ao, Xiangyu Long, Qiang Luo, Ping Duan, Jiang Zhu, Yongsheng Wang, Meijuan Huang, Yongmei Liu, and Youling Gong. First-line treatment of driver gene-negative metastatic lung adenocarcinoma with malignant pleural effusion: Should chemotherapy be combined with an immune checkpoint inhibitor

- or bevacizumab? *Investigational New Drugs*, 42(2):196–206, April 2024. URL: <https://link.springer.com/10.1007/s10637-024-01424-4>, doi:10.1007/s10637-024-01424-4.
- [111] Chuheng Wang, Linan Duan, Yao Zhao, Yubo Wang, and Yunqian Li. Efficacy and safety of bevacizumab combined with temozolomide in the treatment of glioma: A systematic review and meta-analysis of clinical trials. *World Neurosurgery*, 193:447–460, January 2025. URL: <https://linkinghub.elsevier.com/retrieve/pii/S1878875024017753>, doi:10.1016/j.wneu.2024.10.071.
- [112] Yuanliang Zhou, Jingchao Li, Yankang Li, Guangchuan Deng, Qi Wang, Hongyue Qin, Jianbin Li, and Zhenxiang Li. Bevacizumab improved prognosis for advanced egfr-mutant lung adenocarcinoma with brain metastasis receiving cerebral radiotherapy. *Clinical and Translational Oncology*, 26(8):1968–1975, March 2024. URL: <https://link.springer.com/10.1007/s12094-024-03418-3>, doi:10.1007/s12094-024-03418-3.
- [113] Thomas Fleischer, Mads Haugland Haugen, Jørgen Ankill, Laxmi Silwal-Pandit, Anne-Lise Børresen-Dale, Ingrid Hedenfalk, Thomas Hatschek, Jörg Tost, Olav Engebraaten, and Vessela N. Kristensen. An integrated omics approach highlights how epigenetic events can explain and predict response to neoadjuvant chemotherapy and bevacizumab in breast cancer. *Molecular Oncology*, 18(8):2042–2059, August 2024. URL: <https://febs.onlinelibrary.wiley.com/doi/10.1002/1878-0261.13656>, doi:10.1002/1878-0261.13656.
- [114] Qiuling Li, Jieshan Lin, Guojun Hao, Aihua Xie, Shuangxin Liu, and Bin Tang. Nephrotoxicity of targeted therapy used to treat lung cancer. *Frontiers in Immunology*, 15:1369118, July 2024. URL: <https://www.frontiersin.org/articles/10.3389/fimmu.2024.1369118/full>, doi:10.3389/fimmu.2024.1369118.
- [115] Peter W. Janes, Adam C. Parslow, Diana Cao, Angela Rigopoulos, Fook-Thean Lee, Sylvia J. Gong, Glenn A. Cartwright, Ingrid J. G. Burvenich, Ulf Eriksson, Terrance G. Johns, Fiona E. Scott, and Andrew M. Scott. An anti-vegfb antibody reduces abnormal tumor vasculature and enhances the effects of chemotherapy. *Cancers*, 16(10):1902, May 2024. URL: <https://www.mdpi.com/2072-6694/16/10/1902>, doi:10.3390/cancers16101902.
- [116] Gilles Freyer, Anne Floquet, Olivier Tredan, Aurore Carrot, Carole Langlois-Jacques, Jonathan Lopez, Frédéric Selle, Cyril Abdeddaim, Alexandra Leary, Coraline Dubot-Poitelon, Michel Fabbro, Laurence Gladieff, and Michele Lamuraglia. Bevacizumab, olaparib, and durvalumab in patients with relapsed ovarian cancer: a phase ii clinical trial from the gineco group. *Nature Communications*, 15(1):1985, March 2024. URL: <https://www.nature.com/articles/s41467-024-45974-w>, doi:10.1038/s41467-024-45974-w.
- [117] Lucas Moreno, Rebekah Weston, Cormac Owens, Dominique Valteau-Couanet, Marion Gambart, Victoria Castel, C. Michel Zwaan, Karsten Nysom, Nicolas Gerber, Aurora Castellano, Genevieve Laureys, Ruth Ladenstein, Jochen Rössler, Guy Makin, Dermot Murphy, Bruce Morland, Sucheta Vaidya, Estelle Thebaud, Natasha Van Eijkelenburg, Deborah A. Tweddle, Giuseppe Barone, Julie Tandonnet, Nadege Corradini, Pascal Chastagner, Catherine Paillard, Francisco J. Bautista, Soledad Gallego Melcon, Bram De Wilde, Lynley Marshall, Juliet Gray, Susan A. Burchill, Gudrun Schleiermacher, Louis Chesler, Andrew Peet, Martin O. Leach, Kieran McHugh, Roisin Hayes, Neil Jerome, Hubert Caron, Jennifer Laidler, Nicola Fenwick, Grace Holt, Veronica Moroz, Pamela Kearns, Simon Gates, Andrew D.J. Pearson, Keith Wheatley, on behalf of Innovative Therapies for Children with Cancer (ITCC), European Association for Neuroblastoma Research (SIOPEN), Dominique Valteau-Couanet, Francisco José Bautista, Dermot Murphy, Guy Makin, Estelle Thebaud, Natasha Van Eijkelenburg, Dave Hobin, Sucheta Vaidya, Martin Elliott, Gudrun Schleiermacher, Cormac Owens, Giuseppe Barone, Deborah Tweddle, Victoria Castel, Karsten Nysom, Julie Tandonnet, Nadege Corradini, Pascal Chastagner, Marion Gambart, Sarah Jannier, Aurora Castellano, Soledad Gallego, Carole Coze, Anne Auvrignon, Anne Sophie Defachelles, Ricardo Lopez Almaraz, Nicolas Gerber, Ruth Ladenstein, Bram De Wilde, Benedicte Brichard, Roberto Luksch, Michel Zwaan, Beck Popovic Maja, Courtney Willis, Juliet Gray, Simone Hettmer, Lothar Schweigerer, Udo Kontny, Holger Christiansen, Alberto Garaventa, Anthony Ng, Lisa Howell, and Daniel Yeomanson. Bevacizumab, irinotecan, or topotecan added to temozolomide for children with relapsed and refractory neuroblastoma: Results of the itcc-siopen beacon-neuroblastoma trial. *Journal of Clinical Oncology*, 42(10):1135–1145, April 2024. URL: <https://ascopubs.org/doi/10.1200/JCO.23.00458>, doi:10.1200/JCO.23.00458.
- [118] Dana M. Roque, Eric R. Siegel, Natalia Buza, Stefania Bellone, Gloria S. Huang, Gary Altwerger, Vaagn Andikyan, Mitchell Clark, Masoud Azodi, Peter E. Schwartz, Gautam G. Rao, Elena Ratner, and Alessandro D. Santin. Randomized phase ii trial of weekly ixabepilone ± biweekly bevacizumab for platinum-resistant or refractory ovarian/fallopian tube/primary peritoneal cancer (nct03093155): Updated survival and subgroup analyses. *BJC Reports*, 2(1):43, June 2024. URL: <https://www.nature.com/articles/s44276-024-00067-5>, doi:10.1038/s44276-024-00067-5.
- [119] Nicholas J. Hornstein, Mohammad A. Zeineddine, Betul B. Gunes, Andrew J. Pellatt, Mark Knafl, Haifeng Zhu, Anneleis F. Willett, Abdelrahman Yousef, Suyu Liu, Ryan Sun, Andrew Futreal, Scott E. Woodman, Melissa W. Taggart, Michael J. Overman, Daniel M. Halperin, Kanwal P. Raghav, and John Paul Shen. Efficacy and safety of atezolizumab and bevacizumab in appendiceal adenocarcinoma. *Cancer Research Communications*, 4(5):1363–1368, May 2024. URL: <https://aacrjournals.org/cancerrescommun/article/4/5/1363/745526/Efficacy-and-Safety-of-Atezolizumab-and>, doi:10.1158/2767-9764.CRC-24-0019.
- [120] Spencer R. Rosario, Mark D. Long, Shanmuga Chilakapati, Eduardo Cortes Gomez, Sebastiano Battaglia, Prashant K. Singh, Jianmin Wang, Katy Wang, Kristopher Attwood, Suzanne M. Hess, Aj Robert McGray, Kunle Odunsi, Brahm H.

- Segal, Gyorgy Paragh, Song Liu, Jennifer A. Wargo, and Emese Zsiros. Integrative multi-omics analysis uncovers tumor-immune-gut axis influencing immunotherapy outcomes in ovarian cancer. *Nature Communications*, 15(1):10609, December 2024. URL: <https://www.nature.com/articles/s41467-024-54565-8>, doi:10.1038/s41467-024-54565-8.
- [121] Lennert Eismann, Amy X. Xie, Cerise Tang, Andrea Knezevic, Irina Ostrovnaya, Fengshen Kuo, A. Ari Hakimi, Ed Reznik, and Ritesh R. Kotecha. Sample site impacts rna biomarkers for renal cell carcinoma. *European Urology*, 87(1):79–83, January 2025. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0302283824025958>, doi:10.1016/j.eururo.2024.09.004.
- [122] Tetsuya Kokabu, Yosuke Tarumi, Kota Aoki, Ayaka Okamura, Kohei Aoyama, Hisashi Kataoka, Kaori Yoriki, and Taisuke Mori. Effects of parp inhibitors on subsequent platinum-based chemotherapy in patients with recurrent ovarian cancer. *Cancers*, 16(15):2651, July 2024. URL: <https://www.mdpi.com/2072-6694/16/15/2651>, doi:10.3390/cancers16152651.
- [123] Qinghua Liu, Ruoyun Li, Wei Zhu, and Pengfei Zheng. Case report: Microsatellite instability-high pancreas adenocarcinoma with postoperative liver metastasis recurrence treated with multimodality therapy achieving complete pathological response. *Frontiers in Immunology*, 15:1456343, December 2024. URL: <https://www.frontiersin.org/articles/10.3389/fimmu.2024.1456343/full>, doi:10.3389/fimmu.2024.1456343.
- [124] Kosuke Kanno, Kentaro Nakayama, Sultana Razia, Sohel Hasibul Islam, Zahan Umme Farzana, Shahataj Begum Sonia, Hiroki Sasamori, Hitomi Yamashita, Tomoka Ishibashi, Masako Ishikawa, Kayo Imamura, Noriyoshi Ishikawa, and Satoru Kyo. Molecular analysis of high-grade serous ovarian carcinoma exhibiting low-grade serous carcinoma and serous borderline tumor. *Current Issues in Molecular Biology*, 46(9):9376–9385, August 2024. URL: <https://www.mdpi.com/1467-3045/46/9/555>, doi:10.3390/cimb46090555.
- [125] Li Wang, Yibo Fei, Han Qu, Haiyang Zhang, Yuanyuan Wang, Zhenghua Wu, and Guorong Fan. Five years of safety profile of bevacizumab: an analysis of real-world pharmacovigilance and randomized clinical trials. *Journal of Pharmaceutical Health Care and Sciences*, 10(1):1, January 2024. URL: <https://jphcs.biomedcentral.com/articles/10.1186/s40780-023-00314-w>, doi:10.1186/s40780-023-00314-w.
- [126] Yasushi Iida, Nozomu Yanaiharu, Yuki Yoshino, Misato Saito, Ryosuke Saito, Junya Tabata, Ayako Kawabata, Masataka Takenaka, Natsuko Chiba, and Aikou Okamoto. Bevacizumab increases the sensitivity of olaparib to homologous recombination-proficient ovarian cancer by suppressing cryl via pi3k/akt pathway. *Frontiers in Oncology*, 14:1302850, February 2024. URL: <https://www.frontiersin.org/articles/10.3389/fonc.2024.1302850/full>, doi:10.3389/fonc.2024.1302850.
- [127] Qing Wei, Peijing Li, Teng Yang, Jiayu Zhu, Lu Sun, Ziwen Zhang, Lu Wang, Xuefei Tian, Jiahui Chen, Can Hu, Junli Xue, Letao Ma, Takaya Shimura, Jianmin Fang, Jieer Ying, Peng Guo, and Xiangdong Cheng. The promise and challenges of combination therapies with antibody-drug conjugates in solid tumors. *Journal of Hematology Oncology*, 17(1):1, January 2024. URL: <https://jhoonline.biomedcentral.com/articles/10.1186/s13045-023-01509-2>, doi:10.1186/s13045-023-01509-2.
- [128] Kevin Kawchak. mab bioprocess engineering in-context table forecasts using conversational ai literature insight generations. *ChemRxiv*, December 2024. URL: <https://chemrxiv.org/engage/chemrxiv/article-details/674804b05a82cea2fa3af213>, doi:10.26434/chemrxiv-2024-jzbj0.
- [129] Kevin Kawchak. Cancer vs. conversational artificial intelligence. *bioRxiv*, December 2024. URL: <https://www.biorxiv.org/content/10.1101/2024.12.28.630597v1>, doi:10.1101/2024.12.28.630597.
- [130] Kevin Kawchak. Paclitaxel biosynthesis ai breakthrough. *ChemRxiv*, October 2024. URL: <https://chemrxiv.org/engage/chemrxiv/article-details/66fe1dc412ff75c3a1412054>, doi:10.26434/chemrxiv-2024-pqjd3.
- [131] Kevin Kawchak. Monoclonal antibody bioprocess engineering advancements using conversational artificial intelligence. *ChemRxiv*, October 2024. URL: <https://chemrxiv.org/engage/chemrxiv/article-details/671dc78d1fb27ce124b81d22>, doi:10.26434/chemrxiv-2024-3m7m1.
- [132] Google DeepMind. Gemini 2.0 flash thinking experimental is our enhanced reasoning model, capable of showing its thoughts to improve performance and explainability., December 2024. URL: <https://deepmind.google/technologies/gemini/flash-thinking/>.
- [133] DeepSeek. We introduce our first-generation reasoning models, deepseek-r1-zero and deepseek-r1. deepseek-r1-zero, a model trained via large-scale reinforcement learning (rl) without supervised fine-tuning (sft) as a preliminary step, demonstrated remarkable performance on reasoning, January 2025. URL: <https://github.com/deepseek-ai/DeepSeek-R1>.
- [134] OpenAI. Openai 12 days: o3 preview call for safety researchers. deliberative alignment: reasoning enables safer language models introducing our new alignment strategy for o-series models, which are directly taught safety specifications and how to reason over them., 2024. URL: <https://openai.com/12-days/>.
- [135] OpenAI. Openai stargate project. the stargate project is a new company which intends to invest usd 500 billion over the next four years building new ai infrastructure for openai in the united states., 2024. URL: <https://openai.com/index/announcing-the-stargate-project/>.

13 Appendix

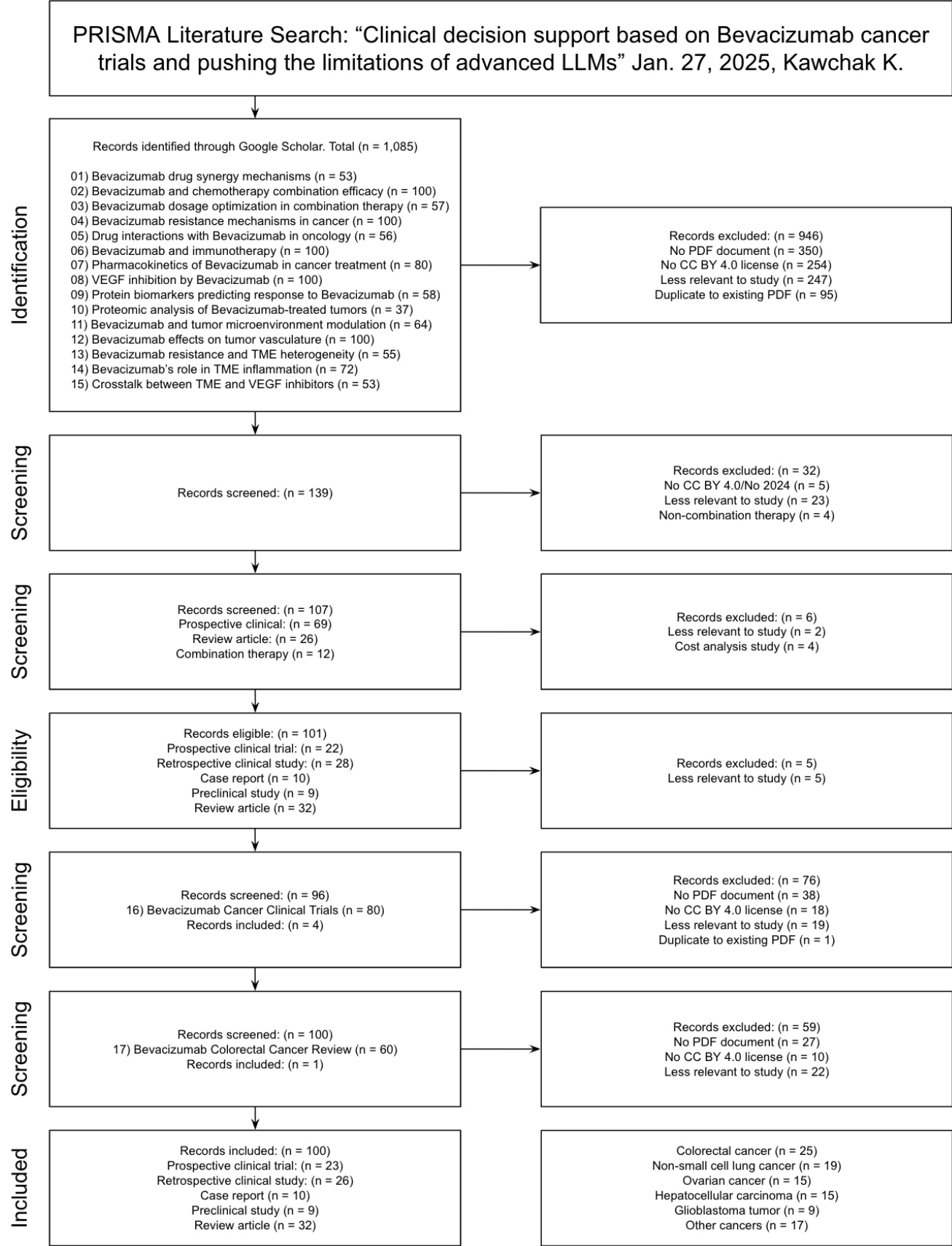


Figure 5: PRISMA study identifying the 100 articles