scientific reports

OPEN

Comparison of artificial intelligence and multidisciplinary team recommendations in the management of colorectal cancer liver metastases

Mustafa Yılmaz[1✉], Najmaddın Abbaslı[2], Simge Tuna[2], Uğfe Kuyucuoğlu[2], Cumhur Özcan[2], Hilmi Bozkurt[3] & Tahsin Çolak[1]

Multidisciplinary teams (MDTs) are central to treatment planning for colorectal cancer liver metastases (CRCLM) but require time and consistent access to expertise. Chat-based large language models (LLMs) such as ChatGPT can generate recommendations from written clinical summaries; however, their concordance with MDT decisions in CRCLM is not well characterized. We conducted a single-center retrospective concordance study of 30 consecutive CRCLM cases discussed at an MDT. ChatGPT was provided a standardized anonymized text synopsis (without direct imaging access) and asked for management recommendations under two a priori conditions: (1) baseline synopsis only, and (2) a conditional query in which resectability status was explicitly specified. Each case and condition was queried three independent times in separate sessions using identical prompts; outputs were mapped to predefined management categories. Agreement between the final LLM recommendation and MDT decisions was assessed using percent agreement and Cohen's kappa. Across repeated runs, the LLM assigned the same management category in all cases (within-model consistency 100%, 3/3) for both querying conditions. In the baseline condition, agreement with MDT decisions was 66.7% (20/30; Cohen's kappa = 0.606, moderate agreement). In the conditional resectability-specified condition, agreement was 93.3% (28/30; Cohen's kappa = 0.924, very good agreement). Baseline discordant cases were characterized by conservative model outputs, including recommendations for systemic therapy and/or additional diagnostic work-up; only two cases remained discordant after resectability was specified. A chat-based LLM showed moderate concordance with unanimous MDT recommendations from minimal case summaries and very good concordance when resectability status was explicitly specified. These findings support feasibility as a supervised decision-support adjunct, but do not establish clinical benefit; prospective outcome-based validation is required.

Keywords Artificial intelligence, Liver neoplasms, Colorectal neoplasms, Decision support systems, Patient care team

Colorectal cancer (CRC) is a major cause of cancer-related mortality globally, with approximately 50% of cases developing liver metastases during the course of the disease[1]. The clinical management of these metastatic lesions necessitates multidisciplinary team (MDT) assessments to optimize survival rates and tailor treatment strategies to the patient. MDT meetings bring together experts from different disciplines, including surgical oncology, medical oncology, radiation oncology, radiology, and pathology, enabling the development of evidence-based and patient-centered treatment algorithms. Literature indicates that this approach increases resectability rates, improves adjuvant/neoadjuvant treatment compliance, and ultimately has significant positive effects on progression-free survival and overall survival[2,3].

[1]Department of Surgical Oncology, Faculty of Medicine , Mersin University , Mersin, Turkey. [2]Department of General Surgery, Faculty of Medicine , Mersin University , Mersin, Turkey. [3]Department of Gastroenterology Surgery, Faculty of Medicine , Mersin University , Mersin, Turkey. ✉email: dktrmstfylmz@gmail.com

However, MDTs face several challenges, including inconsistencies in clinical assessments, inadequate meeting times, time constraints for expert staff, and a lack of standardization in decision-making processes[4,5]. These factors can limit the effectiveness of MDT meetings, making optimal patient management difficult.

In recent years, the use of artificial intelligence (AI) technologies in healthcare has become increasingly widespread and has been suggested to offer significant potential in overcoming the challenges encountered in MDT meetings. In this manuscript, we use "AI" as an umbrella term that includes supervised machine-learning (ML) and radiomics models trained for specific prediction tasks as well as generative large language models (LLMs). Importantly, evidence derived from ML/radiomics applications cannot be directly extrapolated to chat-based LLMs, which generate natural-language recommendations from text inputs and may be sensitive to prompt framing and information completeness. AI-supported decision support systems can reduce assessment inconsistencies by enabling rapid and standardized analysis of clinical data. They can also facilitate the integration of telemedicine as a solution to the problem of expert availability and alleviate the impact of time constraints by providing recommendations on the basis of previous case studies[6,7].

In the specific case of colorectal cancer liver metastasis (CRCLM), AI systems have the potential to improve disease staging, predict treatment response, and more accurately predict patient survival[8,9]. However, most prior work in this area focuses on ML- or radiomics-based prediction tasks, whereas the concordance of chat-based LLM recommendations with MDT decisions remains insufficiently characterized. Because treatment planning in CRCLM is largely gated by resectability assessment and treatment sequencing, clarifying this evidence gap is clinically relevant. Accordingly, AI can contribute to the process of optimizing patient care by helping MDTs make evidence-based, rapid, and consistent decisions.

This study aims to evaluate how a chat-based LLM (ChatGPT) can support traditional MDTs in the treatment of CRCLM by comparing its recommendations with MDT decisions under a standardized baseline clinical synopsis and a resectability-specified (conditional) information state, positioning the model as a decision-support adjunct rather than a replacement for MDT deliberation.

Methods

This retrospective study included 30 patients who were evaluated by the multidisciplinary oncology council of our hospital between January 2023 and January 2025 and who were diagnosed with CRCLM with histopathological confirmation and/or radiological findings. This study was conceived as a pilot feasibility/concordance analysis using a convenience sample of consecutive cases; no formal a priori sample size calculation was performed. Institutional ethics committee approval was obtained for the study (No: 2025/382). All methods were performed in accordance with relevant guidelines and regulations, including the Declaration of Helsinki. The demographic characteristics of the patients (age, sex), primary tumor parameters (localization, histological type), characteristics of the liver metastases (number, size, localization) and laboratory and radiological data were obtained from the hospital electronic records system. MDT decisions were compiled retrospectively from the meeting minutes. According to the meeting minutes, MDT recommendations were reached by unanimous consensus for all cases.

In this study, the GPT-4-turbo-based ChatGPT model (OpenAI, March 2025 version) was used, and the clinical, laboratory and radiological imaging data of the patients were anonymized in a standard format and submitted to the model. ChatGPT was provided a standardized anonymized text synopsis and had no direct access to the original imaging, radiology workstation review, or the full electronic medical record. A standard query was made for all patients as follows: "What is the most appropriate treatment approach for this patient?" To evaluate sensitivity to explicit resectability information, the resectability-specified conditional query was applied to all cases as a second, pre-defined information condition. Accordingly, ChatGPT was additionally asked: "The patient's hepatic metastases appear to have resectability potential; would you recommend a change in the therapeutic approach on the basis of this information?" These two queries were treated as two a priori information conditions (baseline vs. resectability-specified conditional) to reflect a clinically relevant gating variable rather than post hoc "optimization." Each case and condition was queried three independent times in separate sessions using identical prompts. Outputs were mapped to predefined management categories, and the final LLM recommendation for concordance analysis was defined by majority vote (noting that in this cohort all runs yielded 3/3 identical category assignments). A detailed example of the anonymized input format and full ChatGPT responses is provided in the supplementary appendix (see [link]).

The agreement between the recommendations generated by ChatGPT and the MDT decisions was assessed via the percentage of agreement and Cohen's kappa coefficient ($\kappa < 0.20$ poor, 0.21–0.40 low, 0.41–0.60 moderate, 0.61–0.80 good, 0.81–1.00 very good agreement). Kappa values were interpreted by magnitude (e.g., "moderate" for $\kappa \approx 0.60$) and the term "significant" was avoided unless referring to statistical testing.

Statistical analysis

The TIBCO Statistica 13.5.0.17 package program was used for statistical analysis. The numbers and percentages of categorical variables are reported as descriptive statistics. Given the pilot design and limited sample size, no subgroup analyses were performed to assess concordance across clinical strata, as such estimates would be statistically unstable.

Results

In the analysis of tumor location distribution in 30 patients, rectal and sigmoid tumors constituted 40% ($n = 12$) and 30% ($n = 9$) of the cases, respectively, whereas the incidence of tumors in the proximal colon was lower (right colon: 13.33%, $n = 4$; transverse colon: 6.67%, $n = 2$; left colon: 10%, $n = 3$). In terms of sex distribution, the percentage of female patients was 53.33% ($n = 16$), and the percentage of male patients was 46.67% ($n = 14$). Age

Tumor location	n	%	Female patients	Male patients	Average age (years)
Right	4	13.33%	2	2	75.00
Transverse	2	6.67%	2	0	65.00
Left	3	10.00%	1	2	62.33
Sigmoid	9	30.00%	4	5	62.67
Rectum	12	40.00%	7	5	57.00
Total	30	100%	16 (53.33%)	14 (46.67%)	62.17 (avg)

Table 1. Tumor localization and demographics.

Age	Sex	Diagnosis	Metastasis location	MDT decision	ChatGPT decision
81	Male	Right colon adenocarcinoma + liver metastasis	Location: Posterolateral to the right lobe Size: 63 × 27 mm	Right hemicolectomy/metastasectomy	Surgical evaluation after systemic chemotherapy Palliative surgery or stenting if necessary
68	Female	Postoperative liver metastasis (Adenocarcinoma of the rectum)	Location: Right lobe posterior Size: 18 mm	Metastasectomy	Surgical evaluation after systemic chemotherapy
65	Female	Transverse colon adenocarcinoma + liver metastasis	Location: Multiple Size: Largest 27 mm	Colectomy/metastasectomy and adjuvant chemotherapy	Surgical evaluation after systemic chemotherapy Palliative surgery or stenting if necessary
73	Male	Sigmoid colon adenocarcinoma + liver metastasis	Location: Right lobe posterior Size: 78 × 54 mm	Anterior resection and hepatectomy	Surgical evaluation after systemic chemotherapy Palliative surgery or stenting if necessary
80	Female	Right colon adenocarcinoma + liver metastasis	Location: Right lobe posterior Size: 27 × 22 mm	Right hemicolectomy/metastasectomy	Surgical evaluation after systemic chemotherapy Palliative surgery or stenting if necessary
69	Female	Postoperative liver metastasis (Sigmoid colon adenocarcinoma)	Location: Segment 5(suspicious lesion) Size: unspecified	Metastasectomy	Biopsy, advanced imaging
65	Female	Transverse colon adenocarcinoma + liver metastasis	Location: Multiple Size: Largest 27 mm	Colectomy, partial metastasectomy and local ablation	Surgical evaluation after systemic chemotherapy Palliative surgery or stenting if necessary
85	Male	Postoperative liver metastasis (Adenocarcinoma of the rectum)	Location: Multiple Size: Largest 25 mm	Metastasectomy	Surgical evaluation after systemic chemotherapy
68	Female	Postoperative liver metastasis (Right colon adenocarcinoma)	Location: Left lobe Size: 1 cm (suspicious lesion)	Metastasectomy	Biopsy, advanced imaging
65	Female	Postoperative liver and spleen metastasis (Rectum adenocarcinoma)	Location: Right lobe inferior Size:18 mm	Metastasectomy/Splenectomy	Biopsy, advanced imaging, and systemic chemotherapy if metastasis is proven

Table 2. Cases of discordance between the ChatGPT recommendation and MDT recommendation in colorectal cancer patients with liver metastases.

analysis revealed that the mean age decreased from the right colon (75.00 years) to the rectum (57.00 years), and there was a gradient was observed proximal–distal age gradient in this direction. When all localizations were considered, the overall mean age was calculated as 62.17 years. When the histopathological examination results of all the patients were evaluated, the tumor type was defined as adenocarcinoma in all the patients. (Table 1)

In 20 out of 30 cases (66.67%), the same recommendations were made between both decision makers. Cohen's kappa coefficient was calculated as 0.6063, indicating moderate agreement (Table 2). Across three independent runs per case and condition using identical prompts in separate sessions, the model assigned the same management category for all cases (3/3), indicating 100% within-model consistency under fixed prompts and inputs. When the primary sources of disagreement were analysed, the model recommended "surgical evaluation after systemic chemotherapy" and "palliative surgery or stent placement if necessary" for synchronous tumors in 7 patients, whereas MDT preferred curative resection in patients where "surgical evaluation after systemic chemotherapy" was recommended for metachronous tumors, and in 3 patients, MDT gave a direct surgical indication despite the model's recommendation for additional diagnostic procedures. (Table 2).

In the resectability assessment, ChatGPT preferred systemic therapy over surgical resection in the decision-making process. Because resectability is a key clinical gating variable in CRCLM treatment planning, a conditional (resectability-specified) query was applied as a pre-defined second information condition. Therefore, ChatGPT was asked, "If metastasectomy is a viable option, would you prefer surgical resection or would you prefer to continue with your current treatment plan?" In the second analysis performed after this specific questioning, it was observed that the agreement between the two decisions increased. (Fig. 1)

According to the findings, a high level of concordance of 93.33% (full match in 28 out of 30 cases) was determined between the MDT and ChatGPT decisions, and Cohen's kappa value of 0.924 was calculated, indicating very good agreement. In the two cases that remained discordant after resectability was specified, ChatGPT continued to recommend systemic therapy rather than metastasectomy. (Table 3)

Age/Gender: 68 years old, female patient
The patient underwent right hemicolectomy, and pathology revealed well-differentiated adenocarcinoma (pT3N0) with MSS (Microsatellite Stable) characteristics, and adjuvant treatment was initiated. PET-CT and dynamic liver MRI performed during follow-up revealed widespread cystic lesions in the liver parenchyma, a suspicious focus in the left lobe with SUVmax=7.7 and 1 cm in size, lesions compatible with necrotic hypermetabolic infection on the anterior abdominal wall, and FDG uptake compatible with inflammatory process in the colon loops. In addition, multiple large cystic lesions were observed in both kidneys, the largest of which was measured as 115 mm in the right kidney.

What is the most appropriate treatment approach for this patient?

In this 68-year-old female patient with a history of right hemicolectomy for well-differentiated pT3N0 MSS colorectal adenocarcinoma and current imaging findings revealing a small, hypermetabolic liver lesion, widespread hepatic and renal cysts, abdominal wall necrotic lesions suggestive of infection, and inflammatory changes in the colon, the most appropriate treatment approach is a comprehensive multidisciplinary assessment to distinguish between infectious/inflammatory versus metastatic processes. Given the solitary, small (1 cm) liver lesion with moderate FDG uptake, a liver biopsy may be warranted to confirm its nature, particularly in the context of widespread cystic changes that could confound imaging interpretation. Simultaneously, infectious etiologies—particularly for the abdominal wall and colonic findings—should be thoroughly evaluated and managed with appropriate antimicrobial therapy. The renal cysts, while likely benign (e.g., consistent with polycystic kidney disease), may require nephrology input if symptomatic or compromising renal function. Overall, oncologic management should be individualized based on biopsy results, infection control, and systemic assessment, with surveillance or systemic therapy considered accordingly.

The patient's hepatic metastases appear to have the potential for resectability; would you recommend a change in the therapeutic approach based on this information?

Yes, the identification of potentially resectable hepatic metastases in this patient significantly alters the therapeutic approach, shifting the focus toward a curative-intent strategy. In this context, a detailed multidisciplinary evaluation—including input from surgical oncology, hepatobiliary surgery, medical oncology, and radiology—is essential to confirm resectability and determine the optimal timing of intervention. If the solitary liver lesion is confirmed as metastatic and technically resectable, hepatic metastasectomy should be strongly considered, either upfront or following a short course of neoadjuvant chemotherapy to assess tumor biology and potentially downsize the lesion. Postoperative systemic therapy may be warranted based on the final pathology and response. Concurrently, infectious foci and cystic lesions should continue to be managed, but unless they contraindicate surgery or suggest disseminated disease, they should not delay curative treatment of hepatic metastasis.

Fig. 1. Sample conversation screen of clinical decision-support interactions with ChatGPT.

Discussion

A high level of agreement (93.33%, Cohen's kappa 0.924) was observed between the ChatGPT and MDT decisions in our study. Complete agreement was achieved in 28 out of 30 patients. This finding is similar to the 91% agreement rate between IBM Watson for Oncology (WFO) and MDT decisions in a study of 250 CRC patients by Aikemu et al.[10]. Similarly, Kim et al. reported 87% agreement between WFO and MDT recommendations in the management of CRC[11]. Gabriel et al. reported 100% agreement between ChatGPT recommendations based on the European Association of Urology guidelines and MDT decisions in the management of prostate cancer[12].

Age	Gender	Diagnosis	Metastasis location	MDT decision	ChatGPT decision
81	Male	Right colon adenocarcinoma + liver metastasis	Location: posterolateral to the right lobe Size: 63×27 mm	Right hemicolectomy/metastasectomy	Surgical evaluation after systemic chemotherapy Palliative surgery or stenting if necessary
65	Female	Transverse colon adenocarcinoma + liver metastasis	Location: multiple size: largest 27 mm	Colectomy/metastasectomy and adjuvant chemotherapy	Surgical evaluation after systemic chemotherapy Palliative surgery or stenting if necessary

Table 3. Discrepancies detected between the ChatGPT and MDT recommendation after the resectability stage.

Choo et al. reported an 86.7% agreement rate in complicated CRC cases, which is in accordance with our results[13]. These high concordance rates suggest that such systems may be potentially useful as supervised decision-support adjuncts in oncological decision-making for CRC, particularly in guideline-concordant scenarios. The limited number of cases showing noncompliance emphasizes the importance of human expertise in complex cases and highlights the need to develop methodological standards and conduct prospective validation studies for the integration of these systems into clinical practice.

There are studies with high concordance rates in the literature. In a retrospective study by Lee et al.[14] including 656 CRC cases, the absolute concordance rate between WFO and MDT was reported to be 48.9% (increasing to 65.8% when the "Recommended" and "Considered" categories were evaluated together), revealing the variations in the performance of different AI systems. This difference in concordance in the literature can be explained by factors such as the difference in the natural language processing capabilities of the ChatGPT model used, the size of the study population and selection criteria, and the comparison methodology and treatment categorization systems. These findings indicate that methodological heterogeneity should be taken into account in the performance evaluations of AI-assisted decision systems and that caution should be exercised in directly comparing the clinical concordances of different systems.

In our study, the compliance rate, which was 66.67% in the first stage, increased to 93.33% in the second stage. This increase was observed when resectability status—a key clinical gating variable in CRCLM planning—was explicitly specified via a conditional query representing a second, pre-defined information condition, rather than post hoc "optimization." This methodological change reduced the number of noncompliant cases from 10 to 2. Similarly, Aikemu et al. (2021) reported that compliance rates increased after updates to the WFO database[10]. This finding shows that AI systems can generate different recommendations when additional clinically decisive information is provided and when clinical scenarios are more clearly defined. Importantly, this does not demonstrate that the model can independently replicate MDT deliberation; rather, it highlights sensitivity to the explicit availability of resectability information. This demonstrates the continuously improvable nature of AI systems and their adaptability to clinical applications.

In both cases observed in our study, the AI system recommended more conservative treatment approaches. This pattern is plausibly consistent with safety-seeking behavior under uncertainty, particularly because the model was provided standardized text summaries and had no direct access to original imaging review or the full electronic medical record. These findings suggest that AI may lead to more conservative recommendations in some cases, whereas experienced clinicians may prefer more aggressive surgical approaches in selected cases. Similarly, Lee et al. reported cases where the AI system WFO recommended surveillance for liver metastases after surgical resection in CRC patients, whereas clinicians preferred chemotherapy[14]. Kim et al. reported that the agreement between the AI and MDT was more pronounced (88% agreement) in stage IV CRC patients[11]. These differences suggest that AI systems cannot yet completely replace human experts in personalized patient assessment and complex clinical decisions. If used clinically, such conservative outputs could help prompt completion of staging or clarify missing data, but may also risk undertreatment or delays in curative-intent local therapy if over-relied upon; therefore, any use should remain supervised decision support.

Despite the potential benefits of using AI systems in MDTs, several limitations and challenges exist. Lee et al. highlighted that WFO makes more conservative recommendations for elderly patients and that differences in local practices and reimbursement policies regarding the use of bioagents lead to noncompliance[14]. Tjhin et al. discussed medicolegal concerns regarding the use of AI in MDTs[15]. Issues such as patient privacy, data security, informed consent, and division of responsibility are issues that need to be carefully addressed when AI systems are integrated into clinical practice. Furthermore, AI systems cannot replace meaningful human interactions. MDT meetings serve as forums for discussions of patients' clinical, pathological, and radiological data, as well as patient preferences, values, and quality of life expectations. AI systems may not be able to perform such subjective assessments fully. From an operational perspective, chat-based systems may still be useful for supervised pre-MDT preparation (e.g., structuring summaries and prompting missing data), but our study did not quantify time savings, costs, or cost-effectiveness, and these potential advantages should be regarded as hypotheses rather than demonstrated outcomes.

The limitations of our study were that it was single-center and retrospective. As a pilot feasibility/concordance study using a convenience sample ($n = 30$), no formal a priori sample size calculation was performed, and estimates may be imprecise. In addition, our study only evaluated compliance with treatment decisions, and other important parameters, such as clinical outcomes or survival, were not evaluated; therefore, concordance does not establish clinical benefit or correctness. Given the modest cohort size, we did not perform subgroup analyses (e.g., by metastasis burden or age), as such stratum-specific concordance estimates would be statistically unstable. Moreover, the model was provided standardized text synopses without direct imaging review, which may have contributed to discordance in resectability-sensitive cases. Finally, we did not perform time–motion

or cost-effectiveness analyses; thus, operational advantages are discussed as plausible use cases rather than demonstrated outcomes.

Conclusions

In this study, we evaluated the concordance of ChatGPT recommendations with MDT decisions in CRCLM cases. Agreement between ChatGPT and MDT decisions increased from 66.7% in the baseline condition to 93.3% when resectability status was explicitly specified as a conditional information state. These results indicate that a chat-based LLM can show moderate-to-very good concordance with unanimous MDT recommendations when provided standardized text-based case summaries. Importantly, concordance with MDT decisions does not establish clinical correctness or outcome benefit; therefore, prospective outcome-based validation is required before clinical implementation.

Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

References

1. Milana, F. et al. Multidisciplinary tumor board in the management of patients with colorectal liver metastases: A Single-Center review of 847 patients. *Cancers* **14** (16), 3952. https://doi.org/10.3390/cancers14163952 (2022).
2. De Greef, K. et al. Multidisciplinary management of patients with liver metastasis from colorectal cancer. *World J. Gastroenterol.* **22** (32), 7215–7225. https://doi.org/10.3748/wjg.v22.i32.7215 (2016).
3. Li, X. et al. Effects of multidisciplinary team on the outcomes of colorectal cancer patients with liver metastases. *Ann. Palliat. Med.* **9** (5), 2741–2748. https://doi.org/10.21037/apm-20-193 (2020).
4. Jalil, R., Ahmed, M., Green, J. S. A. & Sevdalis, N. Factors that can make an impact on decision-making and decision implementation in cancer multidisciplinary teams: an interview study of the provider perspective. *Int. J. Surg.* **11** (5), 389–394. https://doi.org/10.1016/j.ijsu.2013.02.026 (2013).
5. Lamb, B. W. et al. Quality of care management decisions by multidisciplinary cancer teams: a systematic review. *Ann. Surg. Oncol.* **18** (8), 2116–2125. https://doi.org/10.1245/s10434-011-1675-6 (2011).
6. Jiang, F. et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol.* **2** (4), 230–243. https://doi.org/10.1136/svn-2017-000101 (2017).
7. Topol, E. J. High-performance medicine: the convergence of human and artificial intelligence. *Nat. Med.* **25** (1), 44–56. https://doi.org/10.1038/s41591-018-0300-7 (2019).
8. Mansur, A., Saleem, Z., Elhakim, T. & Daye, D. Role of artificial intelligence in risk prediction, prognostication, and therapy response assessment in colorectal cancer: current state and future directions. *Front. Oncol.* **13**, 1065402. https://doi.org/10.3389/fonc.2023.1065402 (2023).
9. Rompianesi, G., Pegoraro, F., Ceresa, C. D., Montalti, R. & Troisi, R. I. Artificial intelligence in the diagnosis and management of colorectal cancer liver metastases. *World J. Gastroenterol.* **28** (1), 108–122. https://doi.org/10.3748/wjg.v28.i1.108 (2022).
10. Aikemu, B. et al. Artificial intelligence in Decision-Making for colorectal cancer treatment strategy: an observational study of implementing Watson for oncology in a 250-Case cohort. *Front. Oncol.* **10**, 594182. https://doi.org/10.3389/fonc.2020.594182 (2020).
11. Kim, E. J. et al. Early experience with Watson for oncology in Korean patients with colorectal cancer. *PLoS One.* **14** (3), e0213640. https://doi.org/10.1371/journal.pone.0213640 (2019).
12. Gabriel, J., Gabriel, A., Shafik, L., Alanbuki, A. & Larner, T. Artificial intelligence in the urology multidisciplinary team meeting: can ChatGPT suggest European association of urology guideline-recommended prostate cancer treatments? *BJU Int.* **133** (4), 407–409. https://doi.org/10.1111/bju.16240 (2024).
13. Choo, J. M. et al. Conversational artificial intelligence (ChatGPT) in the management of complex colorectal cancer patients: early experience. *ANZ J. Surg.* **94** (3), 356–361. https://doi.org/10.1111/ans.18749 (2024).
14. Lee, W. S. et al. Assessing concordance with Watson for Oncology, a cognitive computing decision support system for colon cancer treatment in Korea. *JCO Clin. Cancer Inf.* **2**, 1–8. https://doi.org/10.1200/CCI.17.00109 (2018).
15. Tjhin, Y., Kewlani, B., Singh, H. K. S. I. & Pawa, N. Artificial intelligence in colorectal multidisciplinary team meetings. What are the medicolegal implications? *Colorectal Dis.* **26** (9), 1749–1752. https://doi.org/10.1111/codi.17091 (2024).

Author contributions

Mustafa Yılmaz and Cumhur Özcan conceived and designed the study. Mustafa Yılmaz, Uğfe Kuyucuoğlu, Simge Tuna, and Najmaddın Abbaslı acquired and analyzed the data. Mustafa Yılmaz drafted the manuscript. Cumhur Özcan and Hilmi Bozkurt critically revised the manuscript for important intellectual content. Tahsin Çolak supervised the entire project and provided administrative support. All authors contributed to the interpretation of results, reviewed and edited the manuscript, and approved its final version.

Declarations

Competing interests

The authors declare no competing interests.

Ethical approval

Ethics committee approval was received from our institution for our study.

Research involving human participants and/or animals and informed consent

This retrospective study used anonymized patient data from hospital electronic records. The requirement for

informed consent was waived by the ethics committee due to the retrospective nature of the study.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-026-38449-z.

Correspondence and requests for materials should be addressed to M.Y.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.