ORIGINAL ARTICLE

AI-ASSISTED RENAL TUMOUR MANAGEMENT: A COMPARATIVE EVALUATION OF CHATGPT AND MULTIDISCIPLINARY TEAM DECISIONS

Jawad Khan[1], Malik Furqan Mahmood[1], Muhammad Saleem[1], Mudassar Abbas[2✉], Saddam Hussain[1], Syed Muhammad Ali Hasnu[1], Abdul Basit Khan[3], Hifza Jadoon[3], Mezhgan Kiwan[4]

[1]Department of Urology, Ayub Teaching Hospital, Abbottabad-Pakistan
[2]Department of Acute Medicine, St. George's University Hospital, London-United Kingdom
[3]Department of Surgery, Ayub Teaching Hospital, Abbottabad-Pakistan, [4]Department Of Medicine, Gomal Medical College, D.I. Khan-Pakistan

Background: Renal tumours, including renal cell carcinoma (RCC), often require complex and personalized decision-making processes, typically achieved through multidisciplinary team (MDT) discussions. With advancements in artificial intelligence (AI), tools like ChatGPT have the potential to aid decision-making. This study evaluates ChatGPT's ability to align its recommendations with MDT decisions in renal tumour management. **Methods**: This retrospective study analyzed 13 renal tumour cases discussed by MDTs. Tumour classifications were based on the TNM staging system, including T1a, T1b, T2, T3/T4, and metastatic RCC. Treatment recommendations partial nephrectomy, radical nephrectomy, systemic therapy, and surveillance were generated by ChatGPT using the European Association of Urology (EAU) guidelines. Concordance with MDT decisions was assessed. **Results**: Localized RCC (T1-T2) accounted for 61.5% of cases, advanced RCC (T3/T4) for 15.4%, and metastatic RCC for 23.1%. Treatment recommendations included partial nephrectomy (23.1%), radical nephrectomy (46.2%), systemic therapy (23.1%), and surveillance (7.7%). ChatGPT demonstrated 100% concordance with MDT decisions across all cases. **Conclusion**: ChatGPT's ability to align with MDT decisions underscores its potential as a supplementary tool in renal tumour management. However, human oversight is essential to account for patient-specific and contextual factors. Larger studies are required to validate these findings and refine AI integration into clinical workflows.

Keywords: Artificial Intelligence; ChatGPT; Renal Tumours; Nephrectomy; Multidisciplinary Team; Decision-Making.

INTRODUCTION

Renal cell carcinoma (RCC) is the most common type of cancer arising in the kidney, accounting for 90% of renal cancers in adults. The occurrence of RCC has been on the rise at a steady rate and accounts for approximately 3% of adult malignancies.[1] With recent technological advancements AI has emerged as the central pillar of the fourth industrial revolution, transforming many fields including healthcare. ChatGPT is an AI language-based model developed by OpenAI and is based on the Generative Pre-Trained Transformer (GPT) architecture, specifically the GPT-4o version.[2] In medicine, ChatGPT's innovative applications are making significant contributions, ranging from streamlining record-keeping to assisting clinicians in decision-making, answering medical exam questions and writing medical reports.[3–8]

To our knowledge there is no study evaluating ChatGPT's ability and accuracy in delivering a management plan for patients with RCC and its comparison with real life multidisciplinary team decisions. We conducted a retrospective comparative study with aim to evaluate the accuracy of ChatGPT in decision making about management of renal tumours. To further assess the knowledge base of ChatGPT, we aim to compare it to the decision of Urology MDT meetings and test it as a potential supplementary decision-making tool in urology.

MATERIAL AND METHODS

This study was conducted following ethical approval from Ayub Teaching Hospital's ethical committee and

employed a retrospective observational study with concordance analysis study design. A total of 13 cases of renal cell carcinoma (RCC) diagnosed between 1st December 2024 to 15th March, 2025 were analyzed. Multidisciplinary team (MDT) meetings typically evaluate treatment recommendations for RCC patients, considering various management strategies, including radiotherapy, chemotherapy, surgery, immunotherapy, surveillance or a combination thereof. The MDT decisions served as the reference standard for comparison with treatment recommendations generated by ChatGPT (paid version 4o). To maintain patient confidentiality, anonymized clinical data were input into ChatGPT in a manner analogous to patient case presentations during MDT discussions. This study included thirteen patients diagnosed with renal cell carcinoma (RCC) based on the European Association of Urology (EAU) guidelines.[9,10] Tumours were classified according to their size, stage, and metastatic status using the TNM staging system in accordance to EAU guidelines.[10] Demographic data and clinical history were obtained from MDT meeting records. To ensure patient confidentiality and eliminate investigator bias, all data were fully anonymized before analysis. ChatGPT was provided with structured prompts simulating patient case presentations. Following data entry into ChatGPT, the treatment responses produced were compared to treatment recommendations given by the MDT. The degree of concordance between the two was evaluated by two independent urologists, who were not part of the MDT discussions. Agreement was evaluated across different tumour stages (TNM classification) to determine whether concordance varied based on disease severity.

RESULTS

All cases of RCC were classified into different sub-stages according to TNM classification for RCC in accordance with EAU guidelines.[10] Amongst these 13 patients, both stage T1a and T1b comprised of 3 cases each, whereas, stage T2b and T3a had 2 patients each, while 3 cases were classified as metastatic RCC. (Table 1) (Figure 1)

All of the 13 patients that were recommended treatment options by the ChatGPT according to EAU guidelines was concordant with the urology MDT decisions. Out of those 13 patients, 3 underwent partial nephrectomy, 6 underwent radical nephrectomy and 1 was put on surveillance. Amongst the 3 patients who were recommended systemic therapy, 2 were lost to follow-up. (Table 2) (Figure 2)

Table-1: Descriptive Statistics

Tumour Type	Number of cases	Percentage
T1a	3	23.1
T1b	3	23.1
T2a	0	0
T2b	2	15.4
T3a	2	15.4
T3b	0	0
T3c	0	0
T4	0	0
Metastatic RCC	3	23.1

Table-2: Treatment Recommendations

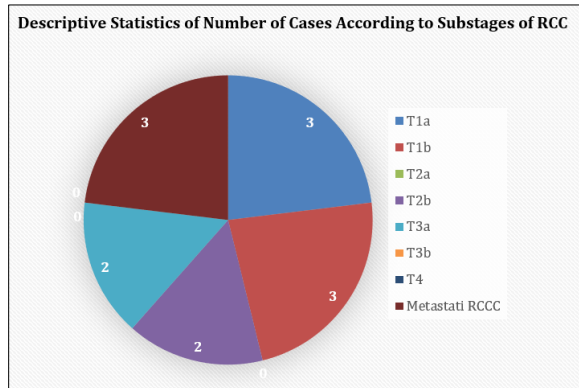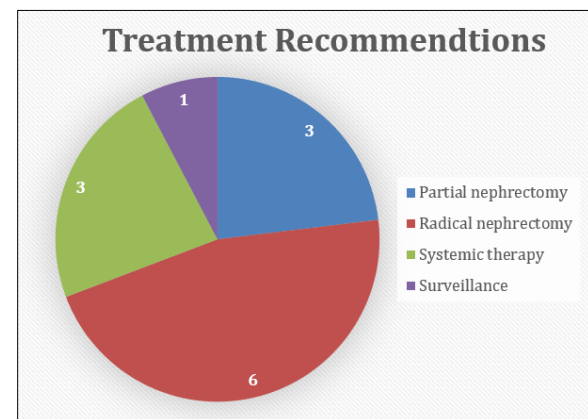| Treatment	Cases	Percentage
Partial nephrectomy	3	23.1
Radical nephrectomy	6	46.2
Systemic therapy	3	23.1
Surveillance	1	7.7

Figure-1: Descriptive Statistics

Figure-2: Treatment Recommendations

DISCUSSION

This study highlights ChatGPT's capacity to generate precise, guideline-concordant treatment recommendations for Renal Cell Carcinoma (RCC) that fully align with multidisciplinary team (MDT) decisions, suggesting its potential to provide, guideline-driven treatment suggestions on the basis of established clinical protocols.

This study shows a hundred percent concordance between AI and MDT on treatment options that were according to European Association of Urology (EAU) guidelines. Although, we could not find any comparative studies on RCC management, recent studies have evaluated the role of ChatGPT in MDT decision-making for other cancer care. Significant concordance between the recommendations from ChatGPT and actual MDT decisions has been reported when standardized oncological guidelines were used, suggesting that AI may act as an adjunct tool to support clinical workflows.[11]

This study shows that the key strength is its adherence to established EAU guidelines, ensuring standardization, and its ability to generate accurate recommendations for localized, advanced, and metastatic RCC cases. This could potentially enhance MDT efficiency by offering consistent, guideline-driven insights.

However, limitations exist, including the lack of personalized considerations, as the AI cannot factor in individual patient preferences, socio-economic conditions, or nuanced clinical variables. Ethical concerns surrounding the accountability of AI-driven decision-making also remain, and the small sample size necessitates more extensive research to assess ChatGPT's broader applicability.

Future research should focus on integrating AI tools like ChatGPT into hybrid decision-making models in a structured manner, ensuring that AI complements clinical expertise while addressing patient-specific variables and ethical considerations.

CONCLUSIONS

The results indicate that ChatGPT could function as a complementary decision-support tool in renal tumour management, reinforcing the value of AI in clinical workflows. However, further large-scale studies are required to confirm its reliability and applicability in diverse clinical settings. Moreover, while promising AI should remain a supplementary tool requiring human oversight.

AUTHORS' CONTRIBUTION

JK: Conceptualization of study design, data analysis, proof reading. MFM, MA, SH, SMAH: Literature search, data interpretation. MS, ABK, HJ, MK: Literature search, data analysis, write-up, proof reading.

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. CA Cancer J Clin 2020;70(1):7–30.
2. Ray PP. ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. Internet Things Cyber Phys Syst 2023;6:100075.
3. Strunga M, Urban R, Surovková J, Thurzo A. Artificial Intelligence Systems Assisting in the Assessment of the Course and Retention of Orthodontic Treatment. Healthcare (Basel) 2023;11(5):683.
4. Yeo YH, Samaan JS, Ng WH, Ting PS, Trivedi H, Vipani A, *et al*. Assessing the performance of ChatGPT in answering questions regarding cirrhosis and hepatocellular carcinoma. Clin Mol Hepatol 2023;29(4):594–604.
5. Lecler A, Duron L, Soyer P, De Margerie-Mellon C, Ben Hassen W, Chassang M, *et al*. Revolutionizing radiology with GPT-based models: Current applications, future possibilities and limitations of ChatGPT. Diagn Interv Imaging 2023;104(4):211–9.
6. Liu J, Wang C, Liu S, Li P, Zhang Y, Chen H, *et al*. Utility of ChatGPT in clinical practice. J Med Internet Res 2023;25:e48568.
7. Gilson A, Safranek CW, Huang T, Socrates V, Chi L, Taylor RA, *et al*. How does ChatGPT perform on the United States Medical Licensing Examination (USMLE)? The implications of large language models for medical education and knowledge assessment. JMIR Med Educ 2023;9:e47860.
8. Jeblick K, Schachtner B, Dexl J, Mittermeier A, Stuber AT, Weber T, *et al*. ChatGPT makes medicine easy to swallow: An exploratory case study on simplified radiology reports. Eur Radiol 2023;33(5):2771–7.
9. Escudier B, Porta C, Schmidinger M, Rioux-Leclercq N, Bex A, Khoo V, *et al*. Renal cell carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment, and follow-up. Ann Oncol 2021;32(4):389–400.
10. European Association of Urology. Staging and Classification Systems. In: Renal Cell Carcinoma: EAU Guidelines. Arnhem (NL): European Association of Urology; 2023. [Internet]. [cited 2025 Jan]. Available from: https://uroweb.org/guidelines/renal-cell-carcinoma/chapter/staging-and-classification-systems
11. Dogan I, Bartin MK, Sonmez E, Seyran E, Bozkurt HA, Yuksek M, *et al*. Chat GPT Performance in Multi-Disciplinary Boards—Should AI Be a Member of Cancer Boards? Healthcare 2025;13(18):2254.

Address for Correspondence:
Mudassar Abbas, Department of Acute Medicine, St. George's University Hospital, London-United Kingdom
Email: mudassarabbas511@gmail.com