Research Article

Artificial intelligence in urologic oncology: the actual clinical practice results of IBM Watson for Oncology in South Korea

Taeyoung Park, Philip Gu, Chang-Hee Kim, Kwang Taek Kim, Kyung Jin Chung,
Tea Beom Kim, Han Jung, Sang Jin Yoon, Jin Kyu Oh[*]

Department of Urology, Gachon University Gil Medical Center, Incheon, Korea

ABSTRACT

Background: Artificial intelligence (AI) is changing our life, including the medical field. Repeated machine learning using big data made various fields more predictable and accurate. In medicine, IBM Watson for Oncology (WFO), trained by Memorial Slone Kettering Cancer Center (MSKCC), was first introduced and applied in 14 countries worldwide.
Our study was designed to assess the feasibility of WFO in actual clinical practice. We aimed to investigate the concordance rate between WFO and multidisciplinary tumor board (MTB) in Urologic cancer patients.
Materials and methods: We reviewed retrospectively collected data for consecutive patients who underwent WFO and MTB simultaneously in the diagnosis of urologic malignancy before determining further treatment between August 2017 and September 2020. We compared the recommendation of the AI system, WFO (IBM Watson Health, Cambridge, MA), with the opinion of MTB for further managing all patients diagnosed such as prostate, bladder, and kidney cancer.
Results: A total of 55 patients were enrolled in our study. The number of patients with prostate cancer was 48. The number of bladder and kidney cancer patients was 5 and 2, respectively. The overall concordance rate between WFO and MTB was 92.7%. Three patients could not suggest proper treatment options using WFO, and the recommended choice of WFO was not feasible in the Korean Health Insurance Review and Assessment Service.
Conclusions: The decision of WFO showed a high concordance rate with a multidisciplinary tumor board for urologic oncology. However, some recommendations of WFO were not feasible in actual practice, and WFO still has some points to improve and modify. Interestingly, applying WFO is likely to facilitate a multidisciplinary team approach.
© 2023 The Asian Pacific Prostate Society. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Artificial intelligence (AI) is the term that includes machine learning and deep learning and has been developed rapidly according to the advancement of computational technology. Repeated machine learning using big data made various fields more predictable and accurate. In medicine, IBM Watson for Oncology (WFO), trained by Memorial Slone Kettering Cancer Center (MSKCC), was first introduced and applied in 14 countries worldwide.[1]

Accordingly, many clinical researchers began to identify the usefulness of WFO in clinical settings. Before the appearance of WFO, a multidisciplinary approach has been the mainstream of making decisions about treatment since its introduction.[2] The multidisciplinary team approach was regarded as a meaningful and successful tool to overcome many practical barriers, especially in the treatment decision for cancer.[3–5] Naturally, most studies on WFO focused on the relationship between WFO and the multidisciplinary team approach.

Many concordance studies exist between WFO and the actual clinical practice in some malignancies, including breast, colorectal, and gastric cancer. Somashekar et al reported that the treatment concordance rate between WFO and the multidisciplinary tumor board (MTB) was 93% of breast cancer cases. Researchers suggested that earlier stage (I) and more advanced stage (IV) showed less

* Corresponding author. Department of Urology, Gachon University Gil Medical Center, 1198, Guwol-Dong, Namdong-Gu, Incheon, 405-760, Korea.
  E-mail address: shaq21@empas.com (J.K. Oh).

concordant than stage II or III.[6] Zhao et al analyzed the concordance between WFO and MTB in patients with breast cancer and reported that the concordance rate was 77%.[7]

In the treatment of colon cancer, Kim et al showed the concordance rate was 87.0% between WFO and MTB. However, they suggested the concordance rate between MTB and NCCN guidelines was 97.1%. They concluded that the concordance rate increased to 88.4%, including the "For consideration".[8] In terms of the decision-making in gastric cancer using WFO, Tian et al reported that the overall concordance between WFO and MTB was 54.5%.[9]

Also, Jie et al reported that the total concordance rate of stage I−III was 86.0%, and that of stage IV was 80.8% between WFO and MTB in their meta-analysis.[10]

However, studies have yet to be reported on the concordance between WFO and actual treatments In the urologic field. Yu et al suggested that using clinical data, the overall prostate cancer treatment concordance rate was 73.6% between WFO and actual treatments.[11] However, their study has performed a comparison between WFO and actual clinical data instead of MTB.

Our study was designed to assess the feasibility of WFO in actual clinical practice, and we aimed to investigate the concordance rate between WFO and MTB in Urologic cancer patients.

2. Materials and methods

2.1. Study population

Our study was approved by the Institutional Review Board of Gachon University Gil Medical Center (Incheon, South Korea). We reviewed retrospectively collected data for consecutive patients who underwent WFO and MTB simultaneously in the diagnosis of urologic malignancy before determining further treatment between August 2017 and September 2020. We compared the recommendation of the AI system, WFO (IBM Watson Health, Cambridge, MA), with the opinion of MTB for further managing all patients diagnosed with urologic malignancies such as prostate, bladder, and kidney cancer.

MTB consisted of all medical staff related to the urologic field, including urologist, radiologist, medical oncologist, pathologist, radiation oncologist, and nuclear medicine specialist.

2.2. Data collection and the criteria of concordance

The data set was collected retrospectively and analyzed according to the clinical stage. We also analyzed patients' age, Eastern Cooperative Oncology Group (ECOG) performance status, and the type of urologic cancer.

The suggested treatment options from WFO consist of three types of answers − Recommended, For consideration, and Not recommended. The concordance criteria were confined to the same suggestive choice between "Recommended" treatment from WFO and the opinion from MTB for the patients.

We calculated the concordance rate between the recommendation of WFO and the opinion of MTB for all patients.

2.3. Statistical analysis

A descriptive analysis of age, gender, cancer stage, and concordance rate was performed. Analyses of the concordance rate between the recommendation of WFO and the opinion of MTB were calculated for prostate, bladder, and kidney cancers. All statistical analyses were conducted using IBM SPSS Statistics for Windows, Version 28.0 (Armonk, NY: IBM Corp, USA).

Table 1
Various demographics of the study

Total patients' number (n)	55
Mean patients' age (y)	68.9
Sex (n)	
Male	55
Female	0
Sex ratio (M:F) (%)	100
Number of patients (n)	
Prostate cancer	48
Bladder cancer	5
Kidney cancer	2
History of surgical treatment (n)	39
Operation rate before WFO (%)	70.9

3. Results

A total of 55 patients were eligible for our study. The mean age of patients was 68.9 years old. All patients were male. The ECOG status of all patients in our study was 0. Of all patients, 48 patients had prostate cancer. Five patients and two patients were diagnosed with bladder cancer and kidney cancer each other (Table 1).

Thirty-nine patients had already gotten surgical treatment before visiting our center. The overall operation rate of all patients was 70.9%.

The opinion of MTB for all patients was the same as the National Comprehensive Cancer Network (NCCN) guidelines.[12]

Regarding prostate cancer, the clinical stages of 42 cases were T3 or higher, and only six patients were diagnosed with T2. The concordance rate between those of WFO and MTB was 91.7% in patients with prostate cancer (Table 2). The reasons for discordance were mainly due to no option in WFO and the feasibility of the suggested option in actual practice (Table 3).

The concordance of bladder and kidney cancer rates between WFO and MTB was 100%.

The overall concordance rate was 92.7% in urologic malignancy.

4. Discussion

Historically, AI was founded in 1955 and was newly focused by AlphaGo, which defeated a professional Go player in 2015. AI is changing almost everything, including the Internet of Things, self-driving cars, and robotic systems. In the health care system, there is no exception to this trend. Machine learning based on big data has become a valuable and increasingly necessary tool in concurrent healthcare systems.[13]

Table 2
Concordance rate according to types of cancer

Types of cancer	Number of patients	Concordance rate (%)
Prostate cancer		
Stage		
T2	6	100
T3	22	86.4
T4	20	95
Bladder cancer		
Stage		
T3	2	100
T4	3	100
Kidney cancer		
Stage		
T3	1	100
T4	1	100
Overall	55	92.7

Table 3
Reason of discordance between multi-disciplinary tumor board and Watson for Oncology

Case	Age (y)	T stage	N stage	M stage	ECOG	Prev. Op.	MDT	WFO	Reason of discordance
1	78	3	1	1	0	No	Docetaxel	Enzalutamide or abiraterone	Not feasible due to high medical cost of enzalutamide
2	72	3	0	0	0	Yes	Combined radiation & ADT	No option	No option for local recurrence in WFO
3	68	3	1	0	0	Yes	Combined radiation & ADT	No option	No option for local recurrence in WFO
4	62	3	1	0	0	Yes	Combined radiation & ADT	No option	No option for local recurrence in WFO

This retrospective study shows a significant difference between the realistic healthcare insurance system in South Korea and the global trends in cancer treatment. In our results, some patients' recommended treatment options from WFO were not feasible in actual practice due to the difference between the global oncologic treatment modalities and the Korean national healthcare system.

WFO, developed by IBM, is an artificial intelligence technology primarily utilized in medicine. However, it may have various limitations.

First of all, there needs to be more retrospective learning and creativity. WFO predominantly operates based on retrospectively learned data. This refers to utilizing existing data to analyze patterns and draw conclusions, primarily employing machine learning techniques. Consequently, Watson relies on existing knowledge and may encounter limitations in discovering or developing innovative treatment methods.

Regarding the early cases of locally recurrent prostate cancer after surgery, there was no recommended option from WFO for the patients (Table 3). After that, WFO and MTB were not tried to decide the treatment for the same-stage patients.

Furthermore, recommended options from WFO have some limitations in prescribing new medications. WFO can analyze existing research articles and medical data to suggest new drugs or treatment methods. Nevertheless, not all countries adopt such information and guidelines universally. The applicability of the information provided by WFO varies according to each country's healthcare systems, regulations, and unique characteristics.

In our study, the first patient with metastatic prostate cancer (T3N1M1) was applied for WFO and MTB in the early period of WFO (Table 3). However, the recommended option from WFO was not feasible because the Korean Healthcare insurance system did not cover medical costs at that time, unlike the US. Since then, WFO has not been applied for the same-stage patients.

Due to the above reasons, the number of cases in our study was only confined to some stage and gradually decreased.

Interestingly, regarding NCCN guidelines for the above patient, the opinion of MTB for metastatic prostate cancer was the same as that of NCCN in 2017.[12] However, recent NCCN guidelines for castration-sensitive metastatic prostate cancer suggest that abiraterone or enzalutamide be the initially recommended treatment.[14]

Due to constraints in cancer treatment and limited training data, WFO might have limited training data regarding urological treatments. Remarkably, the quantity and quality of data related to urological treatment might be insufficient compared to other cancer types like breast or colon cancer. Consequently, WFO might need help generating accurate and reliable conclusions in urological treatment.

These limitations represent challenges that various artificial intelligence systems, including Watson, could encounter. As new knowledge and technologies continue to emerge, efforts to overcome these limitations are anticipated to persist.

Despite the above clinical limitations of WFO in the urologic field, it is clear that WFO was very helpful in facilitating a multidisciplinary tumor board for urologic oncology in our center. Regarding the multidisciplinary tumor board in our center, there have been held only for a few cases in urologic oncology before applying for WFO.

However, after promoting the tumor board using WFO, the multi-disciplinary tumor board multiplied and invigorated. Paradoxically, this is one of the examples in which advanced technology may affect increasing face-to-face contact in the real world.

Our study has some limitations in that the sample size is tiny. In our center, WFO was applied for only intermediate or advanced oncologic cases due to the limited capacity for MTB. An MTB meeting was needed to call many specialists, including patients and their caregivers, simultaneously in the same place. Therefore, activating WFO and MTB in the low-stage cancer was impossible. Furthermore, the lack of updated data in WFO for some situations, such as castration-sensitive recurrent prostate cancer, prevented further MTB.

However, to the best of our knowledge, our study is the first report to assess the concordance of AI and a multidisciplinary tumor board for urologic oncology in actual practice in South Korea. Few studies showed the comparative results between WFO and the opinion of MTB based on clinical data only in urologic malignancy. However, there was no study based on actual clinical practice. In this aspect, our study is historically meaningful and valuable in the urologic field despite the small sample size. Recently, the Dr. Answer AI project, a newly designed software using a machine learning algorithm in South Korea, was introduced for the choice of further treatment of prostate cancer and is expected to broaden the area of AI in the urologic field.[15] However, Dr. Answer has some limitations to apply for clinical practice very soon.

Our future work will be the update and long-term follow-up of the patients to identify and assess the results of survival outcomes and the accuracy of WFO. Furthermore, a newly updated WFO based on a deep learning system will be introduced soon.

5. Conclusions

The decision of WFO showed a high concordance rate with a multidisciplinary tumor board for urologic oncology. However, some recommendations of WFO were not feasible in actual practice, and WFO still has some points to improve and modify. Interestingly, applying WFO will likely facilitate a multidisciplinary team approach.

Ethical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Authors' contributions

T Park: data analysis, manuscript writing and editing, P Gu: data collection and analysis, CH Kim: data analysis, KT Kim: data analysis, KJ Chung: data collection and management, TB Kim: data collection and management, H Jung: data collection and management, SJ Yoon: data analysis, JK Oh: project development, data analysis, and supervision.

Conflicts of interest

The authors declare that there are no conflicts of interest.

Acknowledgments

Not applicable.

References

1. Somashekhar SP, Sepúlveda M-J, Norden AD, Rauthan A, Arun K, Patil P, et al. Early experience with IBM Watson for Oncology (WFO) cognitive computing system for lung and colorectal cancer treatment. J Clin Oncol 2017;35:8527.
2. Fleissig A, Jenkins V, Catt S, Fallowfield L. Multidisciplinary teams in cancer care: are they effective in the UK? Lancet Oncol 2006;7:935—43.
3. Villaflor VM, Allaix ME, Minsky B, Herbella FA, Patti MG. Multidisciplinary approach for patients with esophageal cancer. World J Gastroenterol 2012;18:6737—46.
4. Falzone L, Scandurra G, Lombardo V, Gattuso G, Lavoro A, Distefano AB, et al. A multidisciplinary approach remains the best strategy to improve and strengthen the management of ovarian cancer (Review) Int J Oncol 2021:59.
5. Cheng S, Wei X, Shi J, Guo W, Feng S, Zhai J, et al. A multidisciplinary team approach to the management of patients with hepatocellular carcinoma with portal vein tumor thrombus. Oncologist 2020;25:e998.
6. Somashekhar SP, Sepulveda MJ, Puglielli S, Norden AD, Shortliffe EH, Rohit Kumar C, et al. Watson for Oncology and breast cancer treatment rec-ommendations: agreement with an expert multidisciplinary tumor board. Ann Oncol 2018;29:418—23.
7. Zhao X, Zhang Y, Ma X, Chen Y, Xi J, Yin X, et al. Concordance between treat-ment recommendations provided by IBM Watson for Oncology and a multi-disciplinary tumor board for breast cancer in China. Jpn J Clin Oncol 2020;50: 852—8.
8. Kim EJ, Woo HS, Cho JH, Sym SJ, Baek JH, Lee WS, et al. Early experience with Watson for oncology in Korean patients with colorectal cancer. PLoS One 2019;14e0213640.
9. Tian Y, Liu X, Wang Z, Cao S, Liu Z, Ji Q, et al. Concordance between Watson for Oncology and a multidisciplinary clinical decision-making team for gastric cancer and the prognostic implications: retrospective study. J Med Internet Res 2020;22e14122.
10. Jie Z, Zhiying Z, Li L. A meta-analysis of Watson for Oncology in clinical application. Sci Rep 2021;11:5792.
11. Yu SH, Kim MS, Chung HS, Hwang EC, Jung SI, Kang TW, et al. Early experience with Watson for Oncology: a clinical decision-support system for prostate cancer treatment recommendations. World J Urol 2021;39:407—13.
12. Mohler JL, Armstrong AJ, Bahnson R, D'Amico AV, Davis B, Eastham J, et al. Prostate Cancer, Version 1.2016. J Natl Compr Canc Netw 2016 Jan;14:19—30. https://doi.org/10.6004/jnccn.2016.0004.
13. Beam AL, Kohane IS. Big data and machine learning in health care. JAMA 2018;319:1317—8.
14. Schaeffer EM, Srinivas S, Adra N, An Y, Barocas D, Bitting R, et al. NCCN Guidelines® Insights: Prostate Cancer, Version 1.2023. J Natl Compr Canc Netw 2022;20:1288—98. https://doi.org/10.6004/jnccn.2022.0063.
15. Rho MJ, Park J, Moon HW, Kim CS, Jeon SS, Kang M, et al. Dr. Answer AI for prostate cancer: intention to use, expected effects, performance, and concerns of urologists. Prostate Int 2022;10:38—44.