© 2025 EDIZIONI MINERVA MEDICA
Online version at https://www.minervamedica.it

Minerva Urology and Nephrology 2025 August;77(4):528-37
DOI: 10.23736/S2724-6051.25.06133-6

ORIGINAL ARTICLE

Awareness and use of generative AI-powered tools: results of one-year follow-up prospective cross-sectional global survey

Francesco CEI [1, 2, 3], Conner GANJAVI [1, 2], Ethan LAYNE [1, 2], Tesniem HUSSARI [1, 2], Rafael R. GEVORKYAN [1, 2], Michael EPPLER [1, 2], Pietro PIAZZA [4], Severin RODLER [5], Enrico CHECCUCCI [6], Juan GOMEZ RIVAS [7], Karl F. KOWALEWSKI [8], Ines RIVERO BELENCHÓN [9], Stefano PULIATTI [10], Mark TARATKIN [11], Alessandro VECCIA [12], Loïc BAEKELANDT [13], Pieter DE BACKER [14], Jeremy TEOH [15], Bhaskar SOMANI [16], Marcelo WROCLAWSKI [17, 18, 19], Andre DE CASTRO ABREU [1, 2], Alberto BRIGANTI [3], Andrea SALONIA [3], Francesco MONTORSI [3], Francesco PORPIGLIA [6], Declan MURPHY [20], David CANES [21], Inderbir GILL [1, 2], Giovanni E. CACCIAMANI [1, 2] *

[1]USC Institute of Urology, Catherine and Joseph Aresty Department of Urology, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA; [2]Artificial Intelligence Center, USC Urology, USC Institute of Urology, University of Southern California, Los Angeles, CA, USA; [3]Department of Urology, IRCCS San Raffaele Scientific Institute - Vita-Salute San Raffaele University, Milan, Italy; [4]Division of Urology, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy; [5]Department of Urology, Klinikum der Universitaet Muenchen, Munich, Germany; [6]Department of Surgery, Candiolo Cancer Institute, FPO-IRCCS Candiolo, Turin, Italy; [7]Department of Urology, Clinico San Carlos University Hospital, Madrid, Spain; [8]Department of Urology, University Medical Center Mannheim, Heidelberg University, Heidelberg, Germany; [9]Urology and Nephrology Department, Virgen del Rocío University Hospital, Seville, Spain; [10]Urology Department, University of Modena and Reggio Emilia, Modena, Italy; [11]Institute for Urology and Reproductive Health, Sechenov University, Moscow, Russia; [12]Department of Urology, Azienda Ospedaliera Universitaria Integrata Verona, Verona, Italy; [13]Department of Urology, University Hospitals Leuven, Leuven, Belgium; [14]ORSI Academy, Melle, Belgium; [15]S.H. Ho Urology Center, Department of Surgery, The Chinese University of Hong Kong, Hong Kong, China; [16]University Hospital Southampton NHS Foundation Trust, Southampton, UK; [17]Hospital Israelita Albert Einstein, São Paulo, Brazil; [18]Beneficência Portuguesa de São Paulo, São Paulo, Brazil; [19]Faculty of Medicine do ABC, Santo Andre, Brazil; [20]Division of Cancer Surgery, Peter MacCallum Cancer Centre, University of Melbourne, Melbourne, Australia; [21]Division of Urology, Lahey Hospital and Medical Center, Burlington, MA, USA

*Corresponding author: Giovanni Cacciamani, Artificial Intelligence Center, USC Urology, USC Institute of Urology, Catherine and Joseph Aresty Department of Urology, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.
E-mail: giovanni.cacciamani@med.usc.edu

ABSTRACT

BACKGROUND: Generative AI (GenAI) frameworks, such as generative pre-trained transformer (GPTs) and large language models (LLMs), promise to transform clinical and research practices. Informed human opinion is key to guiding appropriate technological development and task refinement. Detailed data on how GPTs/LLMs powered-Chatbots usage, perceived risks and benefits among physicians has evolved over time and their impact on clinical and academic activities remain unclear. The aim of this study is to assess how the use of GPTs/LLMs chatbots by professionals working in urology has changed over time in the setting of academic and clinical activities.
METHODS: A one-year follow-up prospective cross-sectional survey was conducted among urology professionals. Results were reported highlighting the differences in baseline characteristics of participants and their responses between the

initial release (April 20th, 2023 and May 5th, 2023) and re-deployment of the survey 12 months after chi square and t-test were used to compare categorical and continuous variables.
RESULTS: A total of 129 participants completed the second survey. Eighty-six percent of participants reported having used any GPTs/LLMs chatbot for academic tasks, a significant increase from the previous survey (52.4%; P<0.001). When asked if they were using GPTs/LLMs chatbots more in academic settings compared to one year prior, 70.1% of participants answered affirmatively. Participants, when asked about the use of GPT/LLMs in particular clinical tasks after one year, reported less frequent use for deciding treatment options (18.6% *vs.* 31.0%; P=0.03) and patient follow-up care (10.1% *vs.* 21.4%; P=0.02). When participants were asked if they were using LLM chatbots more in clinical settings compared to one year before, 35.6% answered affirmatively.
CONCLUSIONS: GPTs/LLMs have a consolidated role in academic tasks, with increasing usage, while some resistance to their use in clinical practice remains. These results are relevant for driving the human-centered development of GenAI technology.

(*Cite this article as:* Cei F, Ganjavi C, Layne E, Hussari T, Gevorkyan RR, Eppler M, *et al*. Awareness and use of generative AI-powered tools: results of one-year follow-up prospective cross-sectional global survey. Minerva Urol Nephrol 2025;77:528-37. DOI: 10.23736/S2724-6051.25.06133-6)

KEY WORDS Generative artificial intelligence, Large language models; Surveys and questionnaire.

Generative AI (GenAI)-powered chatbots consist of two main components: a general-purpose AI system and a chat interface.[1] This interface enables users to input prompts freely and receive a natural-language response within seconds that is relevant to the input.[2] This exchange of prompts and responses facilitates a human-AI interaction resembling a conversation between humans.[3] The development of large language models (LLMs) and large multimodal models (LMMs) and generative pre-trained transformer (GPTs), such as GPT-4, Google's Bard, Microsoft Bing, Chatsonic, Palm, Med-PaLM 2, Clinical Camel, MedAlpaca and more, indicates that the implementation of highly complex multimodal reasoning across diverse medical datasets will soon be available.[4] As a result, GenAI is destined for broader adoption by patients[5] and physicians within clinical and academic practice.[6] Indeed, the enthusiasm surrounding GenAI has surged, exemplified by ChatGPT, which amassed over 100 million users and reached the milestone of 1 million users within just a few days – a benchmark that took Netflix three years to achieve.[7] This rapid adoption also raises concerns about the improper use of GenAI, necessitating the development of guidelines to regulate its use in clinical and academic practice.

A previous prospective cross-sectional study has already examined the use of GenAI chatbots.[8] However, detailed data on how GPT/LLMs-powered chatbots usage has evolved over time and their impact on clinical and academic activities remain unclear. Therefore, this study aims to assess how the use of GPTs/ LLMs chatbots by professionals working in urology has changed over time in the setting of academic and clinical activities.

Materials and methods

Survey evolution

As this study represents a one-year follow-up of the initial prospective, cross-sectional survey, the existing survey questions have not been modified.[8] This allows for appropriate response comparisons at multiple time points.

As previously discussed in the study protocol (https://osf.io/b73xf/), the survey is divided into four parts as follows:

• part 1 – baseline characteristics: collected baseline characteristics of respondents working in urology and assessed their experience with technology and artificial intelligence (AI) in healthcare;

• part 2 – ChatGPT for academic and clinical use: investigated the usage of ChatGPT in academic and clinical settings, including its effectiveness and limitations. GAI in academic setting include all those actions that can be employed in developing a scientific project like: employing GAI for literature reviews, generating ideas, drafting initial research proposals, summarizing findings, aiding in data analysis or interpretation, or even improoving grammar and english fluency for not native speakers;

• part 3 – ethical considerations: addressed ethical concerns related to the use of ChatGPT in urology, academic settings, and patient care;

• part 4 – patient perspective concerns: examined urologists' view of potential patient buy-in on the use of ChatGPT in healthcare, including the likelihood for harm and acceptance of AI-generated reports.

To better elucidate the current use of GAI models by participants, questions were added specifically to assess the current prevalence of ChatGPT use in academic use and clinical use respectively.

Survey administration

Survey Monkey Plus (www.surveymonkey.com) was utilized in the re-deployment of the survey to administer the survey and collect the responses. The survey was distributed 12 months after the initial deployment, which took place from April 20th – May 5th, 2023 *via* email-list and social media. A total of 3 reminders were sent. Participants who met inclusion criteria for this survey were individuals who provided the information to be recontacted at 1 year, after taking the first survey,[8] working in the urological field who had completed the initial survey and had provided their contact information along with consent to participate in a follow-up survey. Those who met these criteria were invited to complete the survey between March 11th and May 20th, 2024. The survey was voluntary, and no incentives were granted upon completion. Participants were able to modify their answers at any point.

Data analysis and reporting

Reported results highlight the change in responses between the initial release (time 1) and re-deployment of the survey (time 2). To identify changes in study composition, the baseline demographic characteristics were compared for those who provided contact information for follow-up and those who did not provide contact information in the initial survey. The first page of the survey provided basic study information, a link to the survey protocol (https://osf.io/b73xf/), and an approximate survey length (16 minutes). Details about all questions, including free text responses, are reported in Supplementary Digital Material 1 (Supplementary Table I). Further, the baseline characteristics for those who completed the one-year follow-up survey were compared to those who were invited but did not complete the one-year follow-up. Data are reported as means and standard deviations for continuous variables, while numbers and percentages for dichotomous variables. Chi square and t-test were used to compare categorical and continuous variables. Additional analyses were performed on questions using a Likert scale, where the responses were converted to dichotomous outcomes: "strongly disagree," "disagree," and "neutral" were categorized as "no," while "agree," and "strongly agree" were categorized as "yes." Charts and tables are used when appropriate, with descriptions of the analysis type and measures used provided. Data analysis was completed using Rstudio 2022.02.2 statistical software. Data reporting follows the Checklist For Reporting Results of Internet E-Surveys (CHERRIES).[9]

Results

Baseline characteristics

Out of the 456 participants who responded to the initial survey, 280 (61.4%) provided their email addresses for future contact. After three rounds of reminders, 129 (46.0%) of these participants completed the second survey. The median time to complete the survey was 12 minutes.

When evaluating the baseline characteristics of the population that provided their email addresses after the first survey *versus* those who did not, some significant differences emerged (Supplementary Digital Material 2: Supplementary Table II). Similarly, differences in baseline characteristics were observed when comparing those who provided their email addresses but did not respond to the second survey with those who provided their email addresses and completed the second survey (Supplementary Digital Material 3: Supplementary Table III).

The majority of respondents were male (86.6%), with the largest age group being 30 to 40 years old (44.9%). Most respondents were urologists (60.6%), followed by urology residents (30.7%), urology research scholars (3.9%), and physician assistants (1.6%). Geographically, respondents were predominantly based in Europe (59.8%), followed by Asia (18.1%), Latin America (11.0%), North America (7.9%), Africa

(2.4%), and Oceania (0.8%). The demographic characteristics of the survey participants are detailed in Table I.

Academic use of LLMs

A comparison of the responses to the same survey questions taken at different times revealed a distinct pattern (Figure 1B). Eighty-six percent (86.0%) of participants reported having used a large language model (LLM) chatbot for academic tasks, a significant increase from the previous survey (52.4%; P<0.001). Additionally, 82.9% agreed that LLMs could play an important role in academic practice, compared to 77.0% in the previous year (P=0.02). Specifically, 58.9% of second-time survey takers reported using LLMs for academic brainstorming, up from 46.0% the previous year (P=0.01), while 72.1% *vs.* 81.7% agreed that LLMs can play a relevant role in this activity (P=0.4).

LLM chatbots were more frequently utilized for medical education (50.4% *vs.* 34.9%; P=0.01) and academic writing (66.7% *vs.* 53.2%; P=0.03) compared to the previous survey. In response to the question, "Have you used ChatGPT (or any other LLM) in your academic or research work?" only 8.5% of participants answered "never," a significant decrease from 46.0% in the previous survey (P<0.001). Regarding the specific use of LLMs in academic activities, there was a significant increase in frequent use for grammar checking (60.5% *vs.* 33.3%; P<0.001), translating text into different languages (50.4% *vs.* 23.8%; P<0.001), and summarizing text (45.7% *vs.* 30.2%; P=0.01).

More participants agreed that LLMs are effective for scientific and academic writing (76.0% *vs.* 62.7%; P<0.001), and 70.6% agreed that LLMs positively impacted their academic work (70.6% *vs.* 46.0%; P<0.001). Finally, when asked if they were using LLM chatbots more in academic settings compared to one year prior, 70.1% of participants answered affirmatively (Figure 1A).

Clinical use of LLMs chatbots

The same comparison was conducted for responses to questions regarding the clinical use of LLMs taken one year apart (Figure 2B). When asked about the use of LLMs in clinical practice, a significant increase was observed (41.1% *vs.* 19.0%; P<0.001). Conversely, no significant difference was noted in the use of LLMs for clinical decision-making (50.4% *vs.* 51.6%; P=0.8).

Specifically, when asked about the use of LLMs in particular clinical tasks after one year,

TABLE I.—*Baseline characteristics of respondents.*

Parameters	Respondents
Gender	
Male	110 (85.3%)
Female	16 (12.4%)
Prefer not to report	1 (0.8%)
Age	
Under 30 years old	25 (19.4%)
30-40 years old	57 (44.2%)
41-50 years old	31 (24.0%)
51-60 years old	8 (6.2%)
Over 60 years old	6 (4.7%)
Practice location	
Europe	76 (58.9%)
Asia	23 (17.8%)
Latin America	14 (10.9%)
North America	10 (7.8%)
Africa	3 (2.3%)
Oceania	1 (0.8%)
Years of practice	
0-5 years	48 (37.2%)
6-10 years	34 (26.4%)
11-15 years	15 (11.6%)
16-20 years	15 (11.6%)
Over 20 years	15 (11.6%)
Respondent type	
Urologist	77 (59.7%)
Urology resident/fellow	39 (30.2%)
Urology research scholar	5 (3.9%)
Urology physician assistant	2 (1.6%)
Work environment	
Teaching hospital/academic institution	103 (79.8%)
Mixture of public and private practices	13 (10.1%)
Private practice	6 (4.7%)
Non-academic public hospital	5 (3.9%)
Current status	
Currently active in clinical practice	116 (89.9%)
Currently active in research practice	83 (64.3%)
Retired from clinical practice	2 (1.6%)
On a temporary break from clinical practice	1 (0.8%)
Primary area of interest	
Pediatric urology	10 (7.8%)
Female urology	8 (6.2%)
Andrology/male fertility	17 (13.2%)
Neuro-urology	5 (3.9%)
Urologic oncology	94 (72.9%)
Reconstructive urology	17 (13.2%)
Renal transplantation	12 (9.3%)
Endourology	60 (46.5%)
General urology	47 (36.4%)

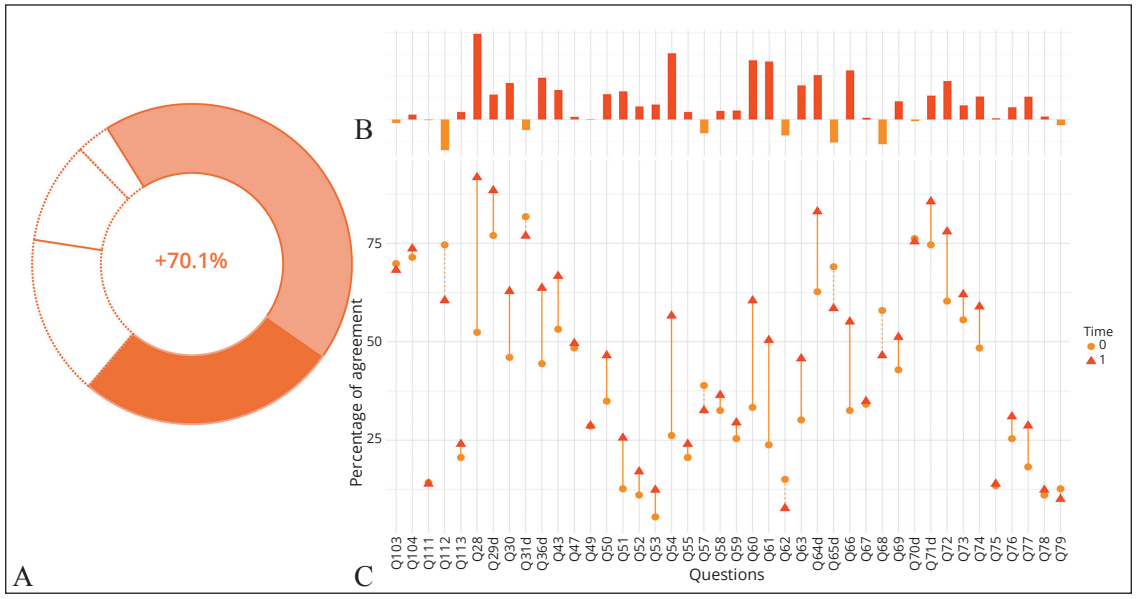

Figure 1.—A) Percentage of participants who increased their use of large language models (LLMs) in academic settings compared to the previous year; B) difference in agreement after one year regarding the use of LLMs in academic settings; C) change in agreement levels over one year regarding the use of LLMs in academic settings.

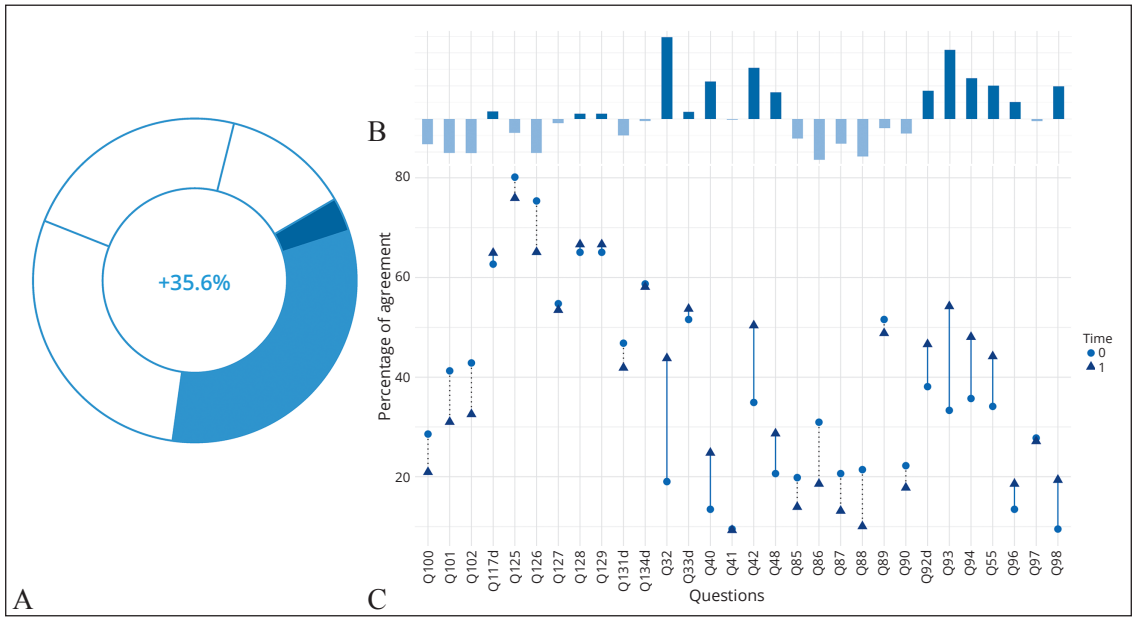

Figure 2.—A) Percentage of participants who increased their use of Large Language Models (LLMs) in clinical settings compared to the previous year; B) difference in agreement after one year regarding the use of LLMs in clinical practice settings; C) change in agreement levels over one year regarding the use of LLMs in clinical practice settings.

participants reported less frequent use for deciding treatment options (18.6% *vs.* 31%; P=0.03) and patient follow-up care (10.1% *vs.* 21.4%; P=0.02). Conversely, no significant differences were noted regarding diagnosis (14.0% *vs.* 19.8%; P=0.2) and medication management (13.2% *vs.* 20.6%; P=0.1). When asked about the frequency of LLM use and their opinions on the

usefulness of LLMs in clinical practice, no significant differences were found (P=0.2). Finally, when participants were asked if they were using LLM chatbots more in clinical settings compared to one year before, 35.6% answered affirmatively (Figure 2A).

Moreover, we noticed a significant difference in the percentage of people who would pay for LLMs if no free versions were available (54.3% *vs*. 45.2%; P=0.03). Finally, when asked in the second survey who was currently paying for premium LLMs, 24.0% responded positively.

Ethical considerations

Approximately half of the participants acknowledged the presence of ethical concerns in using LLM chatbots for scientific or academic writing. This represented a significant increase in the number of participants encountering ethical concerns compared to the previous survey (50.4% *vs*. 32.5%; P<0.001).

The most frequently encountered ethical concerns were artificial intelligence hallucinations (51.2%), defined as outputs that are false, illogic, misleading, or not grounded information, plagiarism (46.5%), and appropriately citing LLM-generated content (34.9%). These concerns were not significantly different from the prior survey (all P>0.05).

Regarding the appropriateness of citing LLM-generated content, only 14.0% of participants believed that LLM-generated text should be avoided, a percentage unchanged from the previous survey (14.3%). However, the number of participants suggesting that LLM-generated text should be clearly disclosed decreased by 14.1% (60.5% *vs*. 74.6%; P=0.02). When asked about possible ethical issues regarding manuscript originality, human creativity, and the potential for spreading misinformation, 48.8%, 41.9%, and 71.3% of participants, respectively, expressed concerns. No significant differences were noted compared to the previous year's survey.

Furthermore, 58.9% of participants agreed that LLM chatbot outputs should not be used in clinical care since they have not been validated. However, almost 30% of these respondents reported using LLMs more in clinical practice compared to the previous year. When asked about additional potential ethical concerns associated with the use of LLMs in patient care and research, 48.1% expressed concerns regarding the privacy and confidentiality of patients' data, 65.1% agreed on the presence of bias in AI algorithms, 68.2% noted the lack of human oversight in decision-making, and 67.4% were concerned about the risk of plagiarism and loss of academic integrity. No significant differences were noted compared to the previous survey (all P>0.05).

Regarding regulations, 76.0% of participants agreed on the need to develop guidelines for LLM chatbot use and oversight, 65.1% supported regular monitoring and auditing of AI systems, and 66.7% supported ethical and legal oversight. No significant differences were noted compared to the previous survey (all P>0.05). Concerning the responsibility to ensure the appropriate use of LLMs, 50.4% of participants agreed on shared responsibility between users and the AI system, 23.3% agreed on the exclusive responsibility of the user, and 17.1% agreed on the exclusive responsibility of the AI system. These results were consistent with the previous year's distribution of responses.

Patient perspective

We also asked participants to give their opinion concerning patients' perspectives on the topic. Thirty-eight percent (38.0%) agreed that patients are likely or very likely to accept reports generated by LLMs. Sixty-two percent (62.0%) believed that patients are likely to use LLMs for self-treatment and self-medication. In this context, more than half of the participants (52.7%) agreed that LLMs can potentially cause harm to patients. None of these responses was significantly different compared to the previous survey (P>0.05).

Discussion

The present prospective 1-year survey shows how people working in urology are using GPTs/LLMs. It appears that GPTs/LLMs have a consolidated role in academic tasks, with increasing usage, while some resistance to their use in clinical practice remains. Since GenAI promises to transform clinical and research practices,[10-13] informed human opinion is key to guiding ap-

propriate technological development and task refinement.[14] This survey serves as a follow-up to assess how professionals in urology are utilizing LLMs in academic and clinical activities and evaluate as well how their ethical concerns have changed at 1-year distance form the first survey. Participants' answers highlighted a significant increase in the use of GPTs/LLMs for academic activities and a higher level of trust in this context. Conversely, when evaluating the role of GPTs/LLMs in clinical settings, respondents were less likely to trust them, and the survey indicated that their use in clinical practice did not increase as much as in academic tasks.

Since the respondents were selected from those who answered the first survey, the baseline characteristics were similar, with a higher prevalence of males, urologists, aged between 30 to 40 years old, and based in Europe. Despite this, significant differences emerged when comparing the baseline characteristics of respondents who did not leave their email for recontacting for the second survey with those who did. However, those who left their email are likely more motivated to contribute compared to those who did not, and consequently, their answers may be more informed and reliable.

The second survey showed that over the past year, the use of GPTs/LLMs for academic activities has increased. Indeed, 82.9% of participants agreed that GPTs/LLMs could play an important role in academic practice, and only 8.5% have never used GPTs/LLMs in research and academics. This overwhelming utilization frequency among urologists surveyed compared to the prior survey shows the increasing popularity of GPTs/LLMs and allows respondents to discuss benefits and limitations of this technology as contextualized from personal experiences. Participants reported increased use of GPTs/LLMs for brainstorming, medical education, academic writing, grammar checking, translating text into different languages, and summarizing text. This trend is plausible considering that Europe, Asia, South America, and Africa represent almost 60% of urology literature production, and most urologists in these regions are not native English speakers. Therefore, many doctors and scientists in these continents are likely benefiting significantly from GPTs/LLMs chatbots.[15] This is in line with other disciplines as well. According to the National Science Board Science and Engineering Indicators, 49% of postdoctoral fellows trained in the United States were born overseas.[16] These results, taken together, suggest that non-native English speakers can significantly benefit in terms of academic productivity and language refinement from the use of GPTs/LLMs in research activities.

While there has been a significant increase in the usage of GPTs/LLMs for academic activities, the same cannot be said for clinical activities. Although 30% of respondents reported using GPTs/LLMs powered chatbots in clinical practice, the percentage of participants agreeing on the use of GPTs/LLMs for making diagnoses and managing medications remained stable over one year. Moreover, the use of GPTs/LLMs powered chatbots for deciding treatment options and patient follow-up care actually decreased after one year, suggesting that GPTs/LLMs powered chatbots have lost some users' trust for these activities.

This is likely related to the fact that GPTs/LLMs still have certain inaccuracies and are far from being perfect.[17-19] When applied to clinical practice without appropriate human oversight, this can lead to severe consequences. Additionally, the poor quality of hospital software in certain countries and the privacy issues related to patient care make the applicability of GPTs/LLMs in clinical practice challenging in many instances.[20] Even with a shift away from clinical decision-making support, those who are using GPTs/LLMs in clinical practice are using it more frequently than they were one year ago. This may indicate that as the technology is further integrated into clinical practice, its use-cases will be refined towards applications that are safe and effective. Some of these applications include improving patient-physician communication, streamlining patient encounters, and assisting with documentation and billing.[1, 21, 22]

Despite GPTs/LLMs' limitations, the potential for their use without human supervision raises some ethical concerns of which survey respondents seem to be aware. Indeed, an 18% increase in ethical issues related to using GPTs/LLMs chatbots for scientific or academic writ-

ing has been observed over one year. However, when specifically analyzing ethical issues such as hallucinations, plagiarism, manuscript originality, human creativity, the potential for spreading misinformation, privacy and confidentiality of patients' data, bias in AI algorithms, and lack of human oversight in decision-making, no significant differences emerged compared to the previous year's survey. This suggests that while awareness of ethical concerns has increased, the nature and frequency of specific issues like hallucinations or plagiarism have not changed significantly over the year. This could be related to the fact that the technology behind GPTs/LLMs' might have stabilized, meaning that the types of issues they produce (*e.g.*, hallucinations, plagiarism) have not changed significantly.[23] However, as users become more familiar with these technologies, their understanding and awareness of potential ethical issues grow.

While ethical concerns and limitations surrounding GPTs/LLMs are becoming more pronounced in the urologic community, there is a decreased interest in requiring the disclosure of GPTs/LLMs in academic writing. Leading journals, publishing houses and international bodies recognize these same ethical concerns and regard proper disclosure of LLMs a requirement for better understanding and countering the pitfalls of this technology.[24, 25] To standardize these disclosures and to preserve the integrity of academic research, the CANGARU initiative is underway and will leverage a large Delphi Consensus model to establish consistent guidelines for the use and reporting of LLMs in academic publication.[26, 27]

In the one-year follow-up survey, there was no increase in calls for regulation despite an increase in ethical concerns. Concurrently, GPTs/LLMs usage in various aspects of clinical decision making has decreased over time. This may potentially indicate self-regulation among respondents – an understanding of the dangers of unchecked GPTs/LLMs usage and self-imposing restrictions to protect patient care.[28] Among respondents who believe GPTs/LLMs outputs should not be used in clinical care due to lack of validation, 30% increased their LLM usage in clinical settings over the past year. This seemingly paradoxical finding bolsters the idea that self-regulation is becoming common practice among urologists given the lack of any legal or operational regulatory ecosystem.[29]

Finally, when evaluating patients' perspectives according to participants, no significant changes were observed compared to the previous year's survey. Most participants still believe that patients are likely to use GPTs/LLMs for self-treatment and self-medication, which could potentially cause them harm. Only a few participants agreed that patients might accept GPTs/LLMs' suggestions. This is likely related to the fact that LLMs lack human empathy, which could be a source of mistrust for patients.[30] Overall, clinicians believe patients would prefer GPTs/LLMs assist in care delivery rather than guide care decisions.

Limitations of the study

This study has some limitations. First, not everyone who answered the first survey also answered the second one. Despite the baseline characteristics of respondents between the first and second surveys being similar, we cannot state that the population of the second survey is representative of the first one. This is because significant differences emerged when comparing respondents who left their email with those who did not leave their email for recontacting. Similarly, a few baseline characteristic differences emerged when comparing those who left their email after the first survey but did not complete the second survey *versus* those who completed both surveys. However, both survey participant samples from the two timepoints were diverse with participation from urologists of all backgrounds.

Secondly, more than half of the respondents came from developed countries and are working in tertiary academic centers. While this may not fully represent the global urological community, it may provide a snapshot of those most likely to have exposure to GPTs/LLMs with views on their integration into healthcare and academics.

Third, the survey was redeployed one year after the initial distribution. There is no guideline or validated tool suggesting the optimal interval for redistributing a prospective cross-sectional survey. The one-year interval was decided based

on internal considerations, allowing for technological advancement and viewpoint maturation.

Finally, other limitations associated with web surveys include the difficulty in calculating an accurate response rate, the lack of standardized, validated tools specifically designed for online surveys, and the potential bias towards interested respondents.[31]

Conclusions

The survey demonstrates a noticeable shift within a year in the adoption and usefulness of AI and LLMs like ChatGPT among urology professionals in both academic and clinical settings. Over the year, the proportion of urologists recognizing the utility of AI has remained stable, but its use increased for academic brainstorming and research. However, ethical concerns and the potential for misinformation continue to be significant barriers. Still an urgent need remains: comprehensive guidelines and oversight mechanisms to ensure AI's ethical integration into urology.

References

1. Lee P, Bubeck S, Petro J. Benefits, Limits, and Risks of GPT-4 as an AI Chatbot for Medicine. N Engl J Med 2023;388:1233–9.

2. Piazza P, Checcucci E, Puliatti S, Belenchòn IR, Veccia A, Rivas JG, *et al.*; Uro-Technology, Social Media Working Group of the Young Academic Urologists Working Party of the European Association of Urology. The Long but Necessary Journey Towards Optimization of the Cause-Effect Relationship Between Input and Output for Accountable Use of ChatGPT for Academic Purposes. Eur Urol Focus 2023;9:1065–7.

3. Open AI. Introducing ChatGPT; 2022 [Internet]. Available from: https://openai.com/index/chatgpt/ [cited 2025, Mar 25].

4. Yang L, Xu S, Sellergren A, Kohlberger T, Zhou Y, Ktena I, *et al.* Advancing Multimodal Medical Capabilities of Gemini. ArXiv; 2024 [Internet]. Available from: https://arxiv.org/pdf/2405.03162 [cited 2025, Mar 25].

5. Checcucci E, Rodler S, Piazza P, Porpiglia F, Cacciamani GE. Transitioning from "Dr. Google" to "Dr. ChatGPT": the advent of artificial intelligence chatbots. Transl Androl Urol 2024;13:1067–70.

6. Cei F, Cacciamani GE. Re: Assessment of Artificial Intelligence Chatbot Responses to Top Searched Queries About Cancer. Eur Urol 2024;86:278–9.

7. Hu K. ChatGPT sets record for fastest-growing user base - analyst note. Reuters; 2023 [Internet]. Available from: https://finance.yahoo.com/news/chatgpt-sets-record-fastest-growing-190911828.html?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAALfuL71H0QXQnOpDXOMzOqmf7eg-_eeYfdzI26_NqskhFFZYjc0V_ogh_DdWCnctstiL2hLrdcLPaGdyJMjvynM6-nQrPpzSBL-nokMNzBfLRxiB1KLZYoskI-7qK9jfGPwip6i7h8fH2VPzI7KGsZ7NS9nkmVik_IV_S1u1d6Ky [cited 2025, Mar 25].

8. Eppler M, Ganjavi C, Ramacciotti LS, Piazza P, Rodler S, Checcucci E, *et al.* Awareness and Use of ChatGPT and Large Language Models: A Prospective Cross-sectional Global Survey in Urology. Eur Urol 2024;85:146–53.

9. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004;6:e34.

10. Checcucci E, Verri P, Amparore D, Cacciamani GE, Fiori C, Breda A, *et al.* Generative Pre-training Transformer Chat (ChatGPT) in the scientific community: the train has left the station. Minerva Urol Nephrol 2023;75:131–3.

11. Rivero Belenchón I, Checcucci E, Gómez Rivas J, Puliatti S, Taratkin M, Kowalewski KF, *et al.*; Uro-Technology and SoMe Working Group of the Young Academic Urologists Working Party of the European Association of Urology. Comment on "Artificial intelligence to predict oncological outcome directly from hematoxylin and eosin-stained slides in urology: a systematic review". Minerva Urol Nephrol 2022;74:810–2.

12. Cocci A, Pezzoli M, Lo Re M, Russo GI, Asmundo MG, Fode M, *et al.* Quality of information and appropriateness of ChatGPT outputs for urology patients. Prostate Cancer Prostatic Dis 2024;27:159–60.

13. Hershenhouse JS, Mokhtar D, Eppler MB, Rodler S, Storino Ramacciotti L, Ganjavi C, *et al.* Accuracy, readability, and understandability of large language models for prostate cancer information to the public. Prostate Cancer Prostatic Dis 2024. [Epub ahead of print]

14. Tang YD, Dong ED, Gao W. LLMs in medicine: the need for advanced evaluation systems for disruptive technologies. Innovation (Camb) 2024;5:100622.

15. Majzoub A, Al Rumaihi K, Al Ansari A. The world's contribution to the field of urology in 2015: A bibliometric study. Arab J Urol 2016;14:241–7.

16. Foreign-born students and workers in the U.S. science and engineering enterprise. National Science Foundation; 2020 [Internet]. Available from: https://www.christophertsmith.com/uploads/3/7/4/4/37446679/foreign-born-students-and-workers-in-us-scieng-enterprise-2020-nsf.pdf [cited 2025, Mar 25].

17. Davis R, Eppler M, Ayo-Ajibola O, Loh-Doyle JC, Nabhani J, Samplaski M, *et al.* Evaluating the Effectiveness of Artificial Intelligence-powered Large Language Models Application in Disseminating Appropriate and Readable Health Information in Urology. J Urol 2023;210:688–94.

18. Rao A, Pang M, Kim J, Kamineni M, Lie W, Prasad AK, *et al.* Assessing the Utility of ChatGPT Throughout the Entire Clinical Workflow: Development and Usability Study. J Med Internet Res 2023;25:e48659.

19. Caglar U, Yildiz O, Meric A, Ayranci A, Yusuf R, Sarilar O, *et al.* Evaluating the performance of ChatGPT in answering questions related to benign prostate hyperplasia and prostate cancer. Minerva Urol Nephrol 2023;75:729–33.

20. Odekunle FF, Odekunle RO, Shankar S. Why sub-Saharan Africa lags in electronic health record adoption and possible strategies to increase its adoption in this region. Int J Health Sci (Qassim) 2017;11:59–64.

21. Meng X, Yan X, Zhang K, Liu D, Cui X, Yang Y, *et al.* The application of large language models in medicine: A scoping review. iScience 2024;27:109713.

22. Rodler S, Ganjavi C, De Backer P, Magoulianitis V, Ramacciotti LS, De Castro Abreu AL, *et al.* Generative artificial intelligence in surgery. Surgery 2024;175:1496–502.

23. Lozic E, Stular B. ChatGPT v Bard v Bing v Claude 2 v Aria v human-expert. How good are AI chatbots at scientific writing? arXiv; 2023 [Internet]. Available from: https://arxiv.org/pdf/2309.08636 [cited 2025, Mar 25].

24. COPE Council. COPE position statement; 2023 [Internet]. Available from: https://publicationethics.org/guidance/cope-position/authorship-and-ai-tools [cited 2025, Mar 25].

25. Ganjavi C, Eppler MB, Pekcan A, Biedermann B, Abreu A, Collins GS, *et al.* Publishers' and journals' instructions to authors on use of generative artificial intelligence in academic and scientific publishing: bibliometric analysis. BMJ 2024;384:e077192.

26. Cacciamani GE, Collins GS, Gill IS. ChatGPT: standard reporting guidelines for responsible use. Nature 2023;618:238.

27. Else H. Should researchers use AI to write papers? Group aims for community-driven standards. Science 2024. [Epub ahead of print]

28. Collier R. Professionalism: the privilege and burden of self-regulation. CMAJ 2012;184:1559–60.

29. Ong JC, Chang SY, William W, Butte AJ, Shah NH, Chew LS, *et al.* Ethical and regulatory challenges of large language models in medicine. Lancet Digit Health 2024;6:e428–32.

30. Wang C, Liu S, Yang H, Guo J, Wu Y, Liu J. Ethical Considerations of Using ChatGPT in Health Care. J Med Internet Res 2023;25:e48009.

31. Singh S, Sagar R. A critical look at online survey or questionnaire-based research studies during COVID-19. Asian J Psychiatr 2021;65:102850.

Conflicts of interest

The authors certify that there is no conflict of interest with any financial organization regarding the material discussed in the manuscript.

Authors' contributions

ChatGPT 4.0 open AI was used for grammar correction in the introduction, methodology, and discussion and conclusions Francesco Cei and Giovanni E. Cacciamani supervised the GPT and take full responsibility for the information provided. Conception and design: Giovanni E. Cacciamani, Francesco Cei. Acquisition of data: Francesco Cei, Conner Ganjavi, Ethan Layne, Tesniem Hussari, Rafael Rafik Gevorkyan, Michael Eppler. Analysis and interpretation of data: Giovanni E. Cacciamani, Francesco Cei. Drafting of the manuscript: Giovanni E. Cacciamani, Francesco Cei, Ethan Layne, Tesniem Hussari, Severin Rodler. Critical revision of the manuscript: Pietro Piazza, Severin Rodler, Enrico Checcucci, Juan Gomez Rivas, Karl F. Kowalewski, Ines Rivero Belenchón, Stefano Puliatti, Mark Taratkin, Alessandro Veccia, Loïc Baekelandt, Pieter De Backer, Jeremy Teoh, Bhaskar Somani, Marcelo Wroclawski, Andre De Castro Abreu, Alberto Briganti, Andrea Salonia, Francesco Montorsi, Francesco Porpiglia, Declan Murphy, David Canes, Inderbir Gill, Giovanni E. Cacciamani. Statistical analysis: Francesco Cei, Giovanni E. Cacciamani. Supervision: Andre De Castro Abreu, Alberto Briganti, Andrea Salonia, Francesco Montorsi, Inderbir Gill, Giovanni E. Cacciamani. All authors read and approved the final version of the manuscript.

History

Article first published online: April 15, 2025. - Manuscript accepted: March 5, 2025. - Manuscript revised: February 20, 2025. - Manuscript received: August 20, 2024.

Supplementary data

For supplementary materials, please see the HTML version of this article at www.minervamedica.it