

Review

Artificial Intelligence and ChatGPT in Abdominopelvic Surgery: A Systematic Review of Applications and Impact

MARTA GOGLIA^{1,2,3}, MARCO PACE¹, MARCO YUSEF¹, GAETANO GALLO⁴,
MATTEO PAVONE⁵, NICCOLÒ PETRUCCIANI¹ and PAOLO AURELLO⁴

¹Department of Medical and Surgical Sciences and Translational Medicine, School in Translational Medicine and Oncology, Faculty of Medicine and Psychology, Sapienza University of Rome, Rome, Italy;

²IHU Strasbourg, Institute of Image-Guided Surgery, Strasbourg, France;

³IRCAD, Research Institute Against Digestive Cancer, Strasbourg, France;

⁴Department of Surgery, Sapienza University of Rome, Rome, Italy;

⁵Dipartimento di Scienze per la Salute della Donna e del Bambino e di Sanità Pubblica, Fondazione Policlinico Universitario A. Gemelli, IRCCS, UOC Ginecologia Oncologica, Rome, Italy

Abstract. *Background/Aim:* The integration of AI and natural language processing technologies, such as ChatGPT, into surgical practice has shown promising potential in enhancing various aspects of abdominopelvic surgical procedures. This systematic review aims to comprehensively evaluate the current state of research on the applications and impact of artificial intelligence (AI) and ChatGPT in abdominopelvic surgery summarizing existing literature towards providing a comprehensive overview of the diverse applications, effectiveness, challenges, and future directions of these innovative technologies. *Materials and Methods:* A systematic search of major electronic databases, including PubMed, Google Scholar, Cochrane Library, Web of Science, was conducted from October to November 2023, to identify relevant studies. *Inclusion criteria* encompassed studies that investigated the utilization of AI and ChatGPT in abdominopelvic surgical settings, including, but not limited to preoperative planning, intraoperative decision-making, postoperative care, and patient communication. *Results:* Fourteen studies met the inclusion

criteria and were included in this review. The majority of the studies were analysing ChatGPT's data output and decision making while two studies reported patient and general surgery resident perception of the tool applied to clinical practice. Most studies reported a high accuracy of ChatGPT in data output and decision-making process, however with an unforgettable number of errors. Conclusion: This systematic review contributes to the current understanding of the role of AI and ChatGPT in abdominopelvic surgery, providing insight into their applications and impact on clinical practice. The synthesis of available evidence will inform future research directions, clinical guidelines, and development of these technologies to optimize their potential benefits in enhancing surgical care within the abdominopelvic domain.

Correspondence to: Marco Pace, MD, Sapienza University of Rome, Department of Medical and Surgical Sciences and Translational Medicine, via di Grottarossa 1035, 00189 Rome, Italy. Tel: +39 3455874837, email: marco.pace@uniroma1.it

Key Words: Artificial intelligence, abdominopelvic surgery, ChatGPT, innovation.

In recent years, the landscape of abdominal surgery has been undergoing a transformative shift, driven by advancements in artificial intelligence (AI) and natural language processing. Among the emerging technologies, ChatGPT, a powerful language model developed by OpenAI, has emerged as a promising tool with the potential to revolutionize the field of surgery. As a novel application of AI, ChatGPT holds the promise of enhancing communication, decision-making, and information retrieval in the complex and dynamic environment of abdominal surgery. Indeed, the ability of ChatGPT to understand and generate human-like text might be integrated in the clinical practice in order to improve interaction between surgical teams, decision support, and patient education. With regard to medical knowledge, an earlier version of ChatGPT was shown to perform at or near the passing threshold of 60% accuracy on the United States Medical Licensing Exam (USMLE) (1, 2).



This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-NC-ND) 4.0 international license (<https://creativecommons.org/licenses/by-nc-nd/4.0>).

This review aims to provide a comprehensive exploration of the current utilization of ChatGPT in abdominal surgery, shedding light on its evolving role, applications, and implications for both surgeons and patients. By delving into the current state of research, practical implementations, and future directions, this review aimed to offer insights into the transformative potential of ChatGPT in optimizing various facets of abdominal surgery.

From preoperative planning and intraoperative decision-making to postoperative care and patient communication, the integration of ChatGPT introduces a new dimension to surgical practice. The present review tried to explore the impact of this technology on surgical workflows, its role in enhancing precision, and the challenges and opportunities associated with its adoption in the clinical setting. As the surgical community embraces the digital era, understanding the nuances of ChatGPT's integration into abdominal surgery becomes paramount for harnessing its capabilities effectively and ensuring optimal patient outcomes.

As we navigate this dynamic intersection of artificial intelligence and surgery, this systematic review aims to offer a panoramic view of the current landscape, fostering a deeper understanding of ChatGPT's potential and paving the way for further research and innovation in abdominal surgery (3).

Materials and Methods

Search strategy. The systematic review adhered to the guidelines set forth by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) and was prospectively registered with the International Prospective Register of Systematic Reviews (PROSPERO) prior to commencing data extraction. A comprehensive search was conducted across multiple databases, including PubMed, Google Scholar, ScienceDirect, with the search period spanning from October to November 2023. The search was restricted to articles published in English, and no additional filters were applied. The initial database retrieval involved screening by title and abstract, employing specific search terms, such as "chatgpt", "abdominal", "general", "surgery", and "artificial intelligence" (4).

Data extraction. The Rayyan software (Qatar Computing Research Institute, HBKU, Doha, Qatar) was used for the selection process. Following the elimination of duplicate publications, the titles, abstracts, and keywords were independently screened for inclusion by M.G. and M.P. Eligible articles then underwent a full-text review. In instances of disparities in inclusion decisions, a consensus was reached through discussion with a third author (G.G.). The inclusion criteria centred on articles describing the adoption of ChatGPT on the abdominopelvic surgical field. Excluded were articles that lacked the utilization of the tool in this field, as well as abstracts, reviews, meta-analyses, letters, and editorials. Studies reporting on ChatGPT as a tool for article writing as well as in field other than abdominopelvic surgery were excluded. As well as articles regarding other kind of surgery beyond abdominopelvic like plastic surgery or otorhinolaryngology were excluded. Subsequently, relevant data pertaining to authors, journal, field of application, aims and studies details, were extracted for subsequent analysis.

Results

The search strategy identified studies reporting the use of ChatGPT in the field of abdominopelvic surgery. Initially, 112 studies were identified, of which 34 full texts were selected at the level of title and abstract screening. Finally, 14 studies met the criteria for inclusion in the systematic review (Figure 1). Due to the heterogeneity of reports in data collection and study outcomes, a quantitative analysis of the results was not appropriate.

All studies found in the Literature were published in 2023 (100%). They were mainly (57.1%) from the field of general surgery (GS) and other fields of abdominopelvic surgery (43%) including urology, pediatric, bariatric and gynaecologic surgery.

Of the included studies, eight were about the capability of the system to generate data output (50%) (3, 5-9); three (24.4%) considered the decision-making accuracy during tumour board (10) or emergency/intraoperative setting; two (14.2%) about the response quality perception by both medical students and patients (11-14); One (7%) included the ability of data collection and one (7%) data generation (15, 16) (Table I).

Discussion

Summary of main results. This systematic review encompassed a diverse range of investigations into the capabilities and limitations of ChatGPT within the medical domain and specifically the surgical field. The analysis revealed that while ChatGPT demonstrated a comprehensive understanding of established general surgical knowledge, its capacity to generate ground-breaking concepts or discoveries beyond existing paradigms was limited. The potential for ChatGPT to serve as an informative resource for patients was acknowledged, albeit accompanied by concerns regarding information accuracy and the generation of non-existent references. Noteworthy was ChatGPT's ability to respond valuably to clinical queries, although further refinement is imperative. Patient perceptions, particularly among those with lower education levels, exhibited a negative shift following explanations provided by the AI chatbot. The study underscored ChatGPT's feasibility in data collection scenarios, indicating its potential utility in certain research contexts. Additionally, a substantial inconsistency in responses with repeated queries was observed. The results also emphasized the dependence of ChatGPT's output quality on the prompt and highlighted the necessity for content verification, along with secure integration into electronic health records before adoption in healthcare systems. The comprehensive synthesis of findings elucidated both the promise and challenges associated with ChatGPT's integration into clinical medicine, emphasizing the importance of collaboration between AI and human expertise (17, 18).

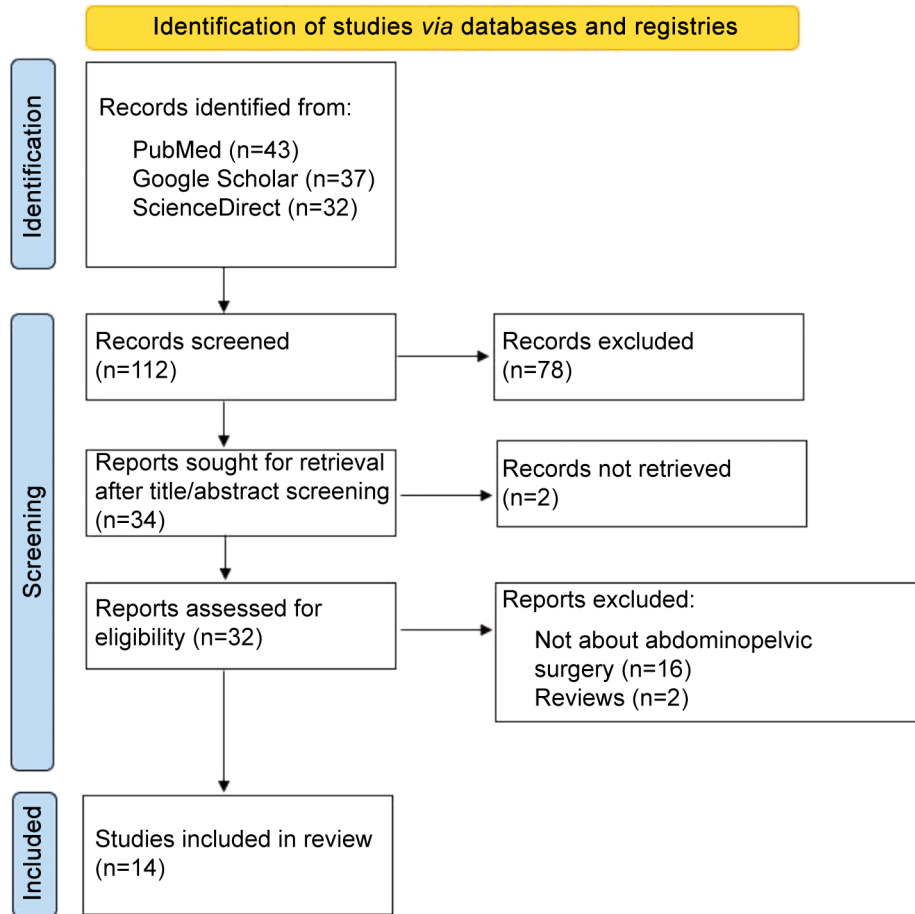


Figure 1. PRISMA flow diagram of study selection.

Indeed, ChatGPT is a tool of great innovation but still difficult to use and incorporate into current clinical practice. The topic most researched and explored by the various authors of all the surgical specialities in question is certainly that of data and information generation. That is, the use of ChatGPT as an interactive colleague that can provide answers to complex clinical questions in real time. For example, Hermann *et al.* queried the system with a questionnaire about cervical cancer, its prevention, diagnosis, treatment, and survival/quality of life (QoL) and rated the system's answers as ranging from 'correct and appropriate answer' to 'completely incorrect and inappropriate answer', finding high rates of 91.7 to 93.8% correct and appropriate answers in the case of prevention and survival/QoL. Modestly lower (71.4% correct answers) in the case of treatment and absolutely insufficient in the case of diagnosis (33% correct answers) (9) Samaan JS *et al.* as well reported the results of the appropriateness of ChatGPT generated answers to bariatric surgery questionnaires. In this case, the

system produced 25% partially correct answers with rare cases of completely incorrect answers (8).

The need behind the use of such software in the surgical field stems from the fact that numerous studies have shown the tendency of patients to search online to answer clinical, surgical, and medical questions. However, the information conveyed *via* online search engines is frequently the subject of fake news and difficult to interpret, a search tool capable of generating correct and appropriate humanoid answers could provide great support for the patient and a foothold for doctor-patient communication, which is still one of the welding points. However, other reported studies investigating ChatGPT data output registered lower rates of accuracy underlying the limitation of large language models and highlighting the need to use them in conjunction with human expertise and judgment (7). Beaulieu-Jones BR *et al.* evaluated ChatGPT responses through two commonly used surgical educational resources: *The Surgical Council on Resident Education* (SCORE) and *Data-B* (1). They

Table I. Details of selected studies.

Year	Author	Country	Field	Topic	Aim
2023	Lim B <i>et al.</i> (3)	Australia	GS	Data output	To assess ChatGPT, an AI language model, in its ability to shed light on the historical facets of general surgery and pinpoint opportunities for innovation.
2023	Jhon A <i>et al.</i> (13)	USA	GS	Perception	This study explored the perspectives of surgical residents on AI's role in medicine.
2023	Gracias D <i>et al.</i> (12)	Australia	GS	Decision making	This study aims to evaluate the accuracy and comprehensiveness of ChatGPT's knowledge on appendicitis management.
2023	Wang G <i>et al.</i> (5)	China	PD	Data output	To explore controversial issues in pediatric inguinal hernia surgery and compare its responses with the guidelines of the European Association of Pediatric Surgeons (EUPSA).
2023	Kim SH <i>et al.</i> (14)	Korea	UR	Perception	We evaluated changes in patient perceptions regarding AI before and after reading a ChatGPT-written explanation.
2023	Rawashdeh B <i>et al.</i> (15)	USA	UR	Data collection	This feasibility study aimed to evaluate the application and accuracy of ChatGPT in the field of kidney transplantation.
2023	Beaulieu-Jones BR <i>et al.</i> (1)	USA	GS	Data output	To probe ChatGPT performance and stability over time on surgical case questions.
2023	Robinson A <i>et al.</i> (16)	UK	GS	Data generation	This study investigates the use of ChatGPT in constructing operation notes for laparoscopic appendectomy.
2023	Lukac S <i>et al.</i> (10)	Germany	GS	Decision making	To test the ability of ChatGPT to support the multidisciplinary tumor board (MDT) in the planning of the therapy of patients with breast cancer
2023	Podda M <i>et al.</i> (11)	Italy	GS	Decision making	To generate an up-to-date bundle to manage acute biliary pancreatitis using an evidence-based, artificial intelligence (AI)-assisted GRADE method.
2023	Levin G <i>et al.</i> (6)	Istrael	GYN	Data output	To perform a bibliometric analysis to identify ChatGPT-related publications in obstetrics and gynecology
2023	Oh N <i>et al.</i> (7)	Korea	GS	Data output	This study aimed to assess the performance of ChatGPT, specifically the GPT-3.5 and GPT-4 models, in understanding complex surgical clinical information and its potential implications for surgical education and training.
2023	Samaan JS <i>et al.</i> (8)	USA	BS	Data output	To examine ChatGPT accuracy and reproducibility in answering patient questions regarding bariatric surgery.
2023	Hermann CE <i>et al.</i> (9)	USA	GYN	Data output	To quantify the accuracy of ChatGPT in answering commonly asked questions pertaining to cervical cancer prevention, diagnosis, treatment, and survivorship/quality-of-life (QOL)

GS: General surgery; PD: pediatric surgery; UR: urology; BS: bariatric surgery; GYN: gynecology.

registered around 68-71% of correct answers justifying the inaccurate responses with inaccurate information in a complex question (n=16, 36.4%); inaccurate information in fact-based question (n=11, 25.0%); and accurate information with circumstantial discrepancy (n=6, 13.6%). Moreover, they reported a substantial inconsistency in ChatGPT responses with repeat query. However, all the studies considered in this review, and which discuss the use of ChatGPT for data output agree that the accuracy of ChatGPT in answering clinical questions is high but not infallible and that it still has certain limitations, such as the production of non-existent or incorrect references or the difficulty in understanding non-linear questions or in answering repeated questions (3, 5).

On the other hand, some of these studies concern the user's perception of ChatGPT. In particular, the perceptions of general surgery residents regarding the role of artificial intelligence in medicine and those of patients confronted with certain medical explanations provided by ChatGPT were investigated.

In the first case, the trainees welcomed the ChatGPT technology positively as an added value to their training in terms of research information and clinical collaboration for the improvement of both training and patient care. In the second case, on the other hand, opinions were more heterogeneous and not always positive. The authors identified negative opinions especially in patients with a lower level of education (13, 14, 19).

Another aspect of using ChatGPT as a revolutionary support tool for the medical profession is that of decision-making support. In this sense, the use of Chat GPT has been directed in the evaluation of opinions produced by the software when faced with complex clinical decisions, for example within a tumour board in planning the correct therapy for breast cancer, or in the management of complex clinical conditions such as pancreatitis.

However, the software, while showing promising capabilities, did not prove infallible and was found to be out of date with the latest therapies and guidelines (10, 11, 20, 21).

ChatGPT's evident capabilities, such as decision-making and data generation, also include that of text generation, which is perhaps the most immediate compared to the software itself. In fact, one study used the tool to generate surgical notes with good results. However, the quality of the reports produced depended strictly on the user's suggestions and indications and therefore simplified and speeded up the work without, however, eliminating human error (16). In the final analysis, given the premises and the incredible improvements over time, it is probably worth considering that ChatGPT may in the future be mentioned among the authors who collaborated in the drafting and generation of a scientific paper, making a substantial contribution not only in structure and methodology but also in content (6).

Results in the context of published literature. As part of the generative pre-trained transformer (GPT) family of models, specifically based on GPT-3.5 architecture, ChatGPT represents a model trained on a diverse range of internet text and capable of understanding and generating coherent and contextually relevant text. "Chat" in ChatGPT refers to its ability to generate human-like responses in a conversational format.

It's important to note that while ChatGPT can produce impressive outputs, it still has many limitations. It may generate incorrect or nonsensical information, be sensitive to input phrasing, and might not always ask for clarification in case of ambiguous queries. OpenAI has made efforts to improve upon these limitations, and it is a part of ongoing research in the field of natural language processing and AI (17, 22, 23).

The systematic review presented in the paper sheds light on the expanding role of artificial intelligence (AI), particularly leveraging ChatGPT, in the realm of abdominopelvic surgery. The integration of AI into surgical practices has the potential to revolutionize patient care and procedural outcomes.

Implications for practice and further research. The utilization of ChatGPT in medicine brings forth significant implications for both clinical practice and future research. While the model showcases a commendable understanding of medical knowledge and offers potential benefits, it also poses challenges that need to be carefully addressed. In practice,

ChatGPT could serve as a valuable resource for disseminating timely and comprehensible medical information to patients, enhancing health literacy. However, the identified issues of inaccuracy and the generation of non-existent or erroneous references underscore the importance of thorough validation and caution in relying solely on AI-generated content for critical medical decisions. The observed negative perception changes among certain patient groups further emphasize the need for tailored communication strategies and consideration of patient demographics in deploying AI-driven tools. Future research should delve into refining ChatGPT's accuracy and reliability, exploring its impact on clinical decision-making processes, patient outcomes, and healthcare provider workflows. Additionally, investigations into ethical considerations, privacy safeguards, and secure integration with electronic health records are paramount to ensure responsible and secure implementation in medical practice. Continuous collaboration between AI developers, healthcare professionals, and researchers is crucial to harness the potential benefits of ChatGPT while mitigating its limitations for safe and effective integration into medical settings (18, 24-27).

Strengths and weaknesses. This is the first systematic review performed according to the PRISMA statement on the adoption of ChatGPT in the abdominopelvic surgery. This study represents an innovative exploration aligning with the freshness of the subject matter. Notably, all existing literature on this topic emerged in 2023, underscoring the contemporary nature of the research. Nevertheless, the very novelty of the subject contributes to a notable diversity in objectives and outcomes, posing a challenge in establishing measurable or comparable parameters across studies. While the systematic review highlights the potential benefits of AI in abdominopelvic surgery, it is crucial to acknowledge the challenges and considerations. Issues, such as data security, model interpretability, and ethical considerations must be addressed to ensure the responsible and effective integration of AI into surgical workflows. Moreover, the review emphasizes the need for continuous model training to enhance stability and consistency in responses (18).

Conclusion

This systematic review provides a comprehensive overview of the current landscape and future potential of AI language model in the surgical scenario. While recognizing the remarkable advancements, the review encourages a cautious and thoughtful approach to address challenges and optimize the collaborative relationship between surgeons and AI technologies. As the field continues to evolve, ongoing research and technological refinement are essential to unlock the full benefits of AI in abdominopelvic surgery.

Declaration

The Authors declare that AI was not used in any part of the manuscript composition process.

Conflicts of Interest

The Authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author's Contributions

PA approved the final version to be published, MP, MG conceived, designed, and wrote the study, MY provided data, MP collected data, GG analysed data, NP critically revised the article.

References

- 1 Beaulieu-Jones BR, Shah S, Berrigan MT, Marwaha JS, Lai SL, Brat GA: Evaluating capabilities of large language models: Performance of GPT4 on surgical knowledge assessments. *medRxiv*, 2023. DOI: 10.1101/2023.07.16.23292743
- 2 Lower K, Seth I, Lim B, Seth N: ChatGPT-4: Transforming medical education and addressing clinical exposure challenges in the post-pandemic era. *Indian J Orthop* 57(9): 1527-1544, 2023. DOI: 10.1007/s43465-023-00967-7
- 3 Lim B, Seth I, Dooreemeah D, Lee CHA: Delving into New Frontiers: assessing ChatGPT's proficiency in revealing uncharted dimensions of general surgery and pinpointing innovations for future advancements. *Langenbecks Arch Surg* 408(1): 446, 2023. DOI: 10.1007/s00423-023-03173-z
- 4 Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hróbjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, McDonald S, McGuinness LA, Stewart LA, Thomas J, Tricco AC, Welch VA, Whiting P, Moher D: The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 372: n71, 2021. DOI: 10.1136/bmj.n71
- 5 Wang G, Liu Q, Chen G, Xia B, Zeng D, Chen G, Guo C: AI's deep dive into complex pediatric inguinal hernia issues: a challenge to traditional guidelines? *Hernia* 27(6): 1587-1599, 2023. DOI: 10.1007/s10029-023-02900-1
- 6 Levin G, Brezinov Y, Meyer R: Exploring the use of ChatGPT in OBGYN: a bibliometric analysis of the first ChatGPT-related publications. *Arch Gynecol Obstet* 308(6): 1785-1789, 2023. DOI: 10.1007/s00404-023-07081-x
- 7 Oh N, Choi GS, Lee WY: ChatGPT goes to the operating room: evaluating GPT-4 performance and its potential in surgical education and training in the era of large language models. *Ann Surg Treat Res* 104(5): 269-273, 2023. DOI: 10.4174/astr.2023.104.5.269
- 8 Samaan JS, Yeo YH, Rajeev N, Hawley L, Abel S, Ng WH, Srinivasan N, Park J, Burch M, Watson R, Liran O, Samakar K: Assessing the accuracy of responses by the language model ChatGPT to questions regarding bariatric surgery. *Obes Surg* 33(6): 1790-1796, 2023. DOI: 10.1007/s11695-023-06603-5
- 9 Hermann CE, Patel JM, Boyd L, Growdon WB, Aviki E, Stasenko M: Let's chat about cervical cancer: Assessing the accuracy of ChatGPT responses to cervical cancer questions. *Gynecol Oncol* 179: 164-168, 2023. DOI: 10.1016/j.ygyno.2023.11.008
- 10 Lukac S, Dayan D, Fink V, Leinert E, Hartkopf A, Veselinovic K, Janni W, Rack B, Pfister K, Heitmeir B, Ebner F: Evaluating ChatGPT as an adjunct for the multidisciplinary tumor board decision-making in primary breast cancer cases. *Arch Gynecol Obstet* 308(6): 1831-1844, 2023. DOI: 10.1007/s00404-023-07130-5
- 11 Podda M, Di Martino M, Ielpo B, Catena F, Coccolini F, Pata F, Marchegiani G, De Simone B, Damaskos D, Mole D, Leppaniemi A, Sartelli M, Yang B, Ansaloni L, Biffi W, Kluger Y, Moore EE, Pellino G, Di Saverio S, Pisanu A: The 2023 MANCTRA acute biliary pancreatitis care bundle: a joint effort between human knowledge and artificial intelligence (ChatGPT) to optimize the care of patients with acute biliary pancreatitis in Western countries. *Ann Surg* 279(2): 203-212, 2024. DOI: 10.1097/SLA.0000000000006008
- 12 Gracias D, Siu A, Seth I, Dooreemeah D, Lee A: Exploring the role of an artificial intelligence chatbot on appendicitis management: an experimental study on ChatGPT. *ANZ J Surg*, 2023. DOI: 10.1111/ans.18736
- 13 St John A, Cooper L, Kavac SM: The role of artificial intelligence in surgery: what do general surgery residents think? *Am Surg* 90(4): 541-549, 2024. DOI: 10.1177/00031348231209524
- 14 Kim SH, Tae JH, Chang IH, Kim TH, Myung SC, Nguyen TT, Choi J, Kim JH, Kim JW, Lee YS, Choi SY: Changes in patient perceptions regarding ChatGPT-written explanations on lifestyle modifications for preventing urolithiasis recurrence. *Digit Health* 9: 20552076231203940, 2023. DOI: 10.1177/20552076231203940
- 15 Rawashdeh B, Kim J, AlRyalat SA, Prasad R, Cooper M: ChatGPT and artificial intelligence in transplantation research: is it always correct? *Cureus* 15(7): e42150, 2023. DOI: 10.7759/cureus.42150
- 16 Robinson A, Aggarwal S Jr: When precision meets penmanship: ChatGPT and surgery documentation. *Cureus* 15(6): e40546, 2023. DOI: 10.7759/cureus.40546
- 17 Wójcik S, Rulkiewicz A, Pruszczyk P, Lisik W, Poboży M, Domienik-Karłowicz J: Beyond ChatGPT: What does GPT-4 add to healthcare? The dawn of a new era. *Cardiol J* 30(6): 1018-1025, 2023. DOI: 10.5603/cj.97515
- 18 Thirunavukarasu AJ, Hassan R, Mahmood S, Sanghera R, Barzangi K, El Mukashfi M, Shah S: Trialling a large language model (ChatGPT) in general practice with the applied knowledge test: observational study demonstrating opportunities and limitations in primary care. *JMIR Med Educ* 9: e46599, 2023. DOI: 10.2196/46599
- 19 Hu JM, Liu FC, Chu CM, Chang YT: Health care trainees' and professionals' perceptions of ChatGPT in improving medical knowledge training: rapid survey study. *J Med Internet Res* 25: e49385, 2023. DOI: 10.2196/49385
- 20 Patil NS, Huang RS, van der Pol CB, Larocque N: Using artificial intelligence chatbots as a radiologic decision-making tool for liver imaging: Do ChatGPT and Bard communicate information consistent with the ACR appropriateness criteria? *J Am Coll Radiol* 20(10): 1010-1013, 2023. DOI: 10.1016/j.jacr.2023.07.010
- 21 Rosen S, Saban M: Evaluating the reliability of ChatGPT as a tool for imaging test referral: a comparative study with a clinical decision support system. *Eur Radiol*, 2023. DOI: 10.1007/s00330-023-10230-0

- 22 Lebharr MS, Velazquez A, Goza S, Hoppe IC: Dr. ChatGPT: Utilizing artificial intelligence in surgical education. *Cleft Palate Craniofac J*: 10556656231193966, 2023. DOI: 10.1177/10556656231193966
- 23 Alfertshofer M, Hoch CC, Funk PF, Hollmann K, Wollenberg B, Knoedler S, Knoedler L: Sailing the seven seas: a multinational comparison of ChatGPT's performance on medical licensing examinations. *Ann Biomed Eng*, 2023. DOI: 10.1007/s10439-023-03338-3
- 24 Seth I, Bulloch G, Lee CHA: Redefining academic integrity, authorship, and innovation: the impact of ChatGPT on surgical research. *Ann Surg Oncol* 30(8): 5284-5285, 2023. DOI: 10.1245/s10434-023-13642-w
- 25 Ayoub NF, Lee YJ, Grimm D, Divi V: Head-to-head comparison of ChatGPT versus google search for medical knowledge acquisition. *Otolaryngol Head Neck Surg*, 2023. DOI: 10.1002/ohn.465
- 26 Emile SH, Horesh N, Freund M, Pellino G, Oliveira L, Wignakumar A, Wexner SD: How appropriate are answers of online chat-based artificial intelligence (ChatGPT) to common questions on colon cancer? *Surgery* 174(5): 1273-1275, 2023. DOI: 10.1016/j.surg.2023.06.005
- 27 Tomioka K, Aoki T, Kobayashi N, Tashiro Y, Kumazu Y, Shibata H, Hirai T, Yamazaki T, Saito K, Yamazaki K, Watanabe M, Matsuda K, Kusano T, Fujimori A, Enami Y: Development of a novel artificial intelligence system for laparoscopic hepatectomy. *Anticancer Res* 43(11): 5235-5243, 2023. DOI: 10.21873/anticancer.16725

Received January 10, 2024

Revised February 27, 2024

Accepted March 11, 2024