BMJ Health & Care Informatics

AIAIAI: AI insights on amassing influence in AI-related publications – an AI-assisted retrospective analysis into AI-related publication

Rotem Lahat,[1] Noa Berick,[2] Majd Hajouj,[2] Tali Teitelbaum,[2] Isaac Shochat [1,2]

To cite: Lahat R, Berick N, Hajouj M, *et al*. AIAIAI: AI insights on amassing influence in AI-related publications – an AI-assisted retrospective analysis into AI-related publication. *BMJ Health Care Inform* 2025;**32**:e101244. doi:10.1136/bmjhci-2024-101244

Received 03 August 2024
Accepted 17 March 2025

[1]Technion Israel Institute of Technology The Ruth and Bruce Rappaport Faculty of Medicine, Haifa, Israel
[2]Hillel Yaffe Medical Center, Hadera, Israel

Correspondence to
Dr Isaac Shochat;
shvecht@gmail.com

ABSTRACT

Objectives This study analyses the trend of artificial intelligence (AI)-related publications in the medical field over the past decade and demonstrates the potential of AI in automating data analysis. We hypothesise exponential growth in AI-related publications, with continuous growth in the foreseeable future.

Methods Retrospective, AI-assisted analysis was conducted using the OpenAI application programming interface for data collection and evaluation. Publications from the top 50 medical journals (Web of Science, Journal Citation Report, 2022) covering 2014 to June 2024. A total of 315 209 papers were initially retrieved with 212 620 remaining after filtering. The outcomes were the total number and percentage of AI-related publications per year, with future trends prediction using statistical models.

Results AI-related publications increased from approximately 500 in 2014 to over 1000 in 2022, with the percentage rising from 2.5% to over 6% in 2024. The analysis identified cardiology and oncology as leading in AI adoption. Predictive models forecast that AI-related publications could reach 10% by 2030 with long-term projections suggesting potential dominance of AI presence by the mid-22nd century.

Discussion The study highlights the significant growth and integration of AI in medical research, with cardiology and oncology at the forefront. AI-assisted data analysis proves efficient and scalable but requires human oversight to maintain credibility.

Conclusions The trajectory of AI-related publications indicates substantial growth and future integration across medical disciplines. Ongoing evaluation of AI's reliability and applicability in medical research remains essential.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Artificial intelligence (AI) has become increasingly prevalent in medical research, with a growing body of literature discussing its applications in diagnostics, treatment planning and automated data analysis.

WHAT THIS STUDY ADDS

⇒ This study provides a quantitative assessment of AI-related publication trends in medical journals, demonstrating a significant increase in AI adoption across specialties and forecasting future growth.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The findings emphasise the expanding role of AI in medical research, highlighting the need for improved methodologies and guidelines to ensure the reliability and credibility of AI-generated insights.

INTRODUCTION

Artificial intelligence (AI) has swept across various sectors like a viral cat video, redefining the boundaries of what technology can achieve. Encompassing a range of technologies, including machine learning, natural language processing and computer vision, AI offers vast capabilities.[1] These technologies enable machines to analyse complex data, recognise patterns and make decisions, offering unprecedented efficiency and innovation across industries.[2]

As in many fields, the trend of AI-related publication has failed to skip medicine. So far, AI has shown significant promise, making waves in diagnostics and medical education. Radiology, for example, has witnessed remarkable advancements with AI algorithms capable of interpreting medical images with high accuracy. These algorithms assist radiologists in detecting abnormalities, thereby improving diagnostic accuracy and patient outcomes.[3] In medical education, AI is providing adaptive learning platforms, virtual simulations and personalised training modules for medical students and professionals.[4 5]

The recent trend of large language model (LLM)-related publications had also allowed the use of AI in assisting the writing of scientific articles, clinical reports and medical documentation. AI can summarise research findings and even conduct literature reviews. If used correctly, this could speed up the

writing process and ensure consistency and accuracy, allowing researchers to focus more on the analysis and interpretation of data.[6]

However, as recent publications demonstrated, LLMs have not always been used correctly.[7] As a statistically driven text generation tool, LLMs have occasionally lacked scientific rigour and accuracy,[8] leading to the publication of misleading information.[9] Without thorough human oversight, these AI-generated texts can propagate errors, potentially compromising the integrity of scientific literature.[10 11]

The use of AI in the medical field has yet to be quantified and comprehensively reviewed. The delicate balance between its massive beneficial potential and underlying risks calls for such analysis, to improve the understanding and application of current and future AI tools in medical research.

Our study aims to demonstrate the possibilities and pitfalls of AI use, through an AI-led analysis of the trajectory of AI-related publications in the top 50 medical journals over the past 10 years.

We hypothesise that the use of AI in medical research has grown dramatically over the past decade, with a similar positive trend across different medical fields. We speculate that AI models may falsely predict a full domination of AI-related articles in the foreseeable future, demonstrating some of the AI limitations.

MATERIALS AND METHODS

Our entire research was AI-based, using the capabilities and sage advice from ChatGPT (OpenAI, model 4o). The AI engine used Python programming language to construct scripts for all data collection and statistical analysis. These scripts were deployed through ChatGPT's interface, using Google Colab or locally. The code will be uploaded to an open repository. The only part performed fully manually was the writing of this paper, which was then handed to ChatGPT merely for proofreading and light sparkle.

We included publications from the top 50 journals in medicine (according to the 2022 Web of Science, Journal Citation Report) in our initial analysis. ChatGPT constructed a script, using PubMed Pythonic application programming interface (API), retrieving all accessible abstracts and articles metadata from these journals, dating 10 years ago and up to June 2024.

Our search yielded 315 309 articles from all journals, including each article's title, writers, abstract, and year and month of publication. For reasons known only to the AI deities, four journals yielded no results using our AI-based methods. After eliminating duplicate or erroneous entries, 212 620 individual articles remained. We did not attempt to retrieve the missing data manually to keep our AI-based approach intact.

Our initial method of determining if an article was AI-related was using AI-related keywords in each article's content. Examining the results, however, showed this method to be either too constrictive or too permissive. Therefore, we turned to an AI-based approach, using ChatGPT to write a Python script processing each line in the table as an individual query through OpenAI API with the instruction "Please answer TRUE or FALSE only, is the following article definitely AI-related:". To validate this method, we randomly selected a sample of 20 articles and debated with ChatGPT regarding the rationale and credibility of its decisions. Convinced of its validity, we used this method for the AI relation determination.

We analysed the results over time, plotting the total absolute and relative number of AI-related publications per year and per month. We used linear regression to fit several trendlines, incorporating polynomial features to capture non-linearity and plotted 95% CIs.

To address any variation between different fields in medicine, we plotted the relative number of AI-related articles over the years, stratified by the predominating categories in our data: cardiology, general medicine, oncology, neurology and other specialties. The AI engine added another category of 'unknown' for articles it could not classify. We also plotted a smoothed graph of this data, using a rolling average calculation with a delta of 3 years.

Finally, we created two models to predict the use of AI in future publications and to forecast the date when all articles in the field of medicine will be AI-related. The models we used were autoregressive integrated moving average (ARIMA) model and an error-trend-seasonality (ETS) exponential smoothing model.

This study was conducted and reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology guidelines.

RESULTS

Our initial retrieval yielded 315 209 papers across the field. Trying to retrieve data from *Acta Neuropathologica* resulted in various errors we could not resolve without abusing PubMed API policies. Three journals did not yield any information through the API and were excluded (*JAMA, Lancet—Child & Adolescent Health, Nature reviews Gastroenterology & Hepatology*). Skimming through the data, 2257 records without a specific month were excluded from the per-month analysis, while 57 did without a publication year and were excluded from all time-based analyses. Additionally, the retrieval process for *Psychiatry and Clinical Neurosciences* resulted in sparse information (21 publications) from *Neurosciences (Riyadh, Saudi Arabia)*. Eliminating duplicates left us with 229 588 individual papers, and further elimination of missing or improper data left us with a final dataset of 212 620 papers for analysis.

The total cost of OpenAI line-by-line queries resulted in a cost of US$61.62 ingesting 260 165 API requests and 11 768 851 tokens overall. This cost was influenced by many repeating queries, human errors and initial finetuning.

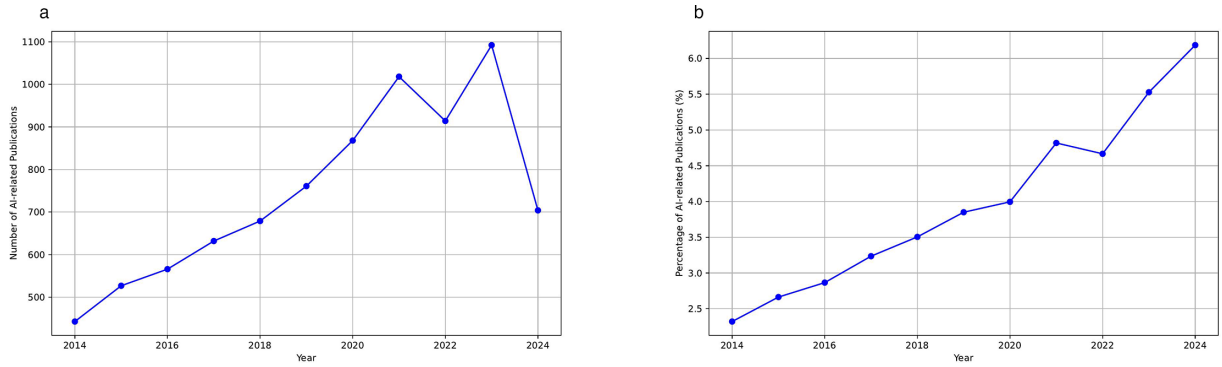

Figure 1 (a) Trend of AI-related publications over the years, (b) Percentage of AI-related publications over the years. AI, artificial intelligence.

Our analysis revealed a significant increase in the number of AI-related publications from 2014 to 2024. The number of publications rose from about 500 in 2014 to over 1000 in 2022, before experiencing a slight dip in 2024 (figure 1a). When looking at the percentage of AI-related publications relative to the total number of publications, the trend becomes even more apparent. The percentage has grown from 2.5% in 2014 to over 6% in 2024 (figure 1b).

Examining the data for a more granular view, plotting the percentage of AI-related publications on a monthly basis shows a steady upward trend with some seasonal variations. Furthermore, we applied a polynomial trend line with a 95% CI to the data. The results suggest that while there are fluctuations, the overall trajectory is a solid upward trend (figure 2).

To understand the impact of AI across various medical fields, we stratified the data by specialty. The analysis showed that neurology and oncology are leading the pack in AI adoption, while other specialties are gradually catching up (figure 3).

Using the ARIMA and exponential smoothing (ETS) models, we forecasted the future of AI-related publications. Our models predict that if the current trends continue, we might see an even higher percentage of AI-related articles, potentially nearing 10% by 2030 (figure 4a,b). Delving further into the future, our models can say, with somewhat dubious probability, that AI-related

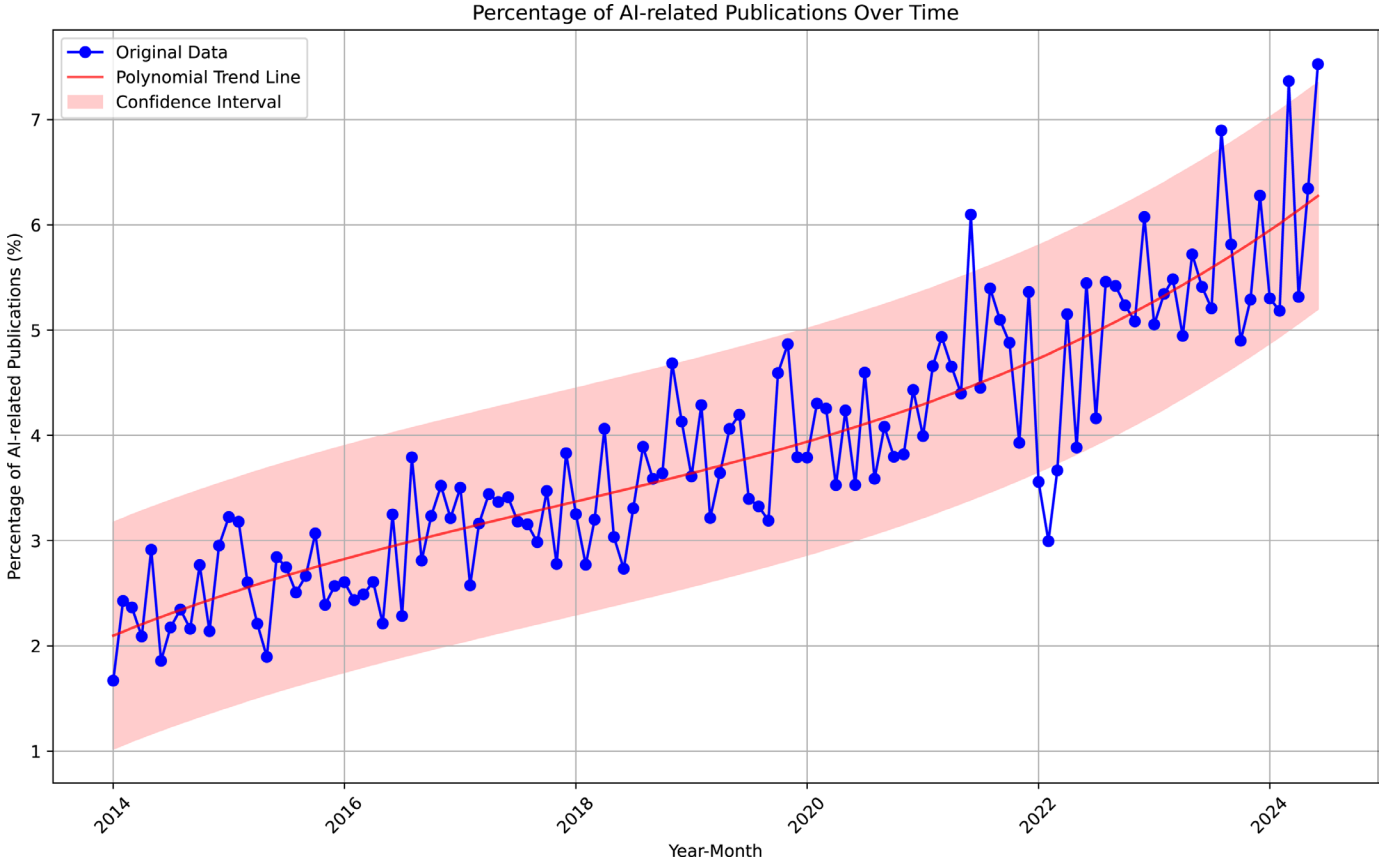

Figure 2 Percentage of AI-related publications over the years.

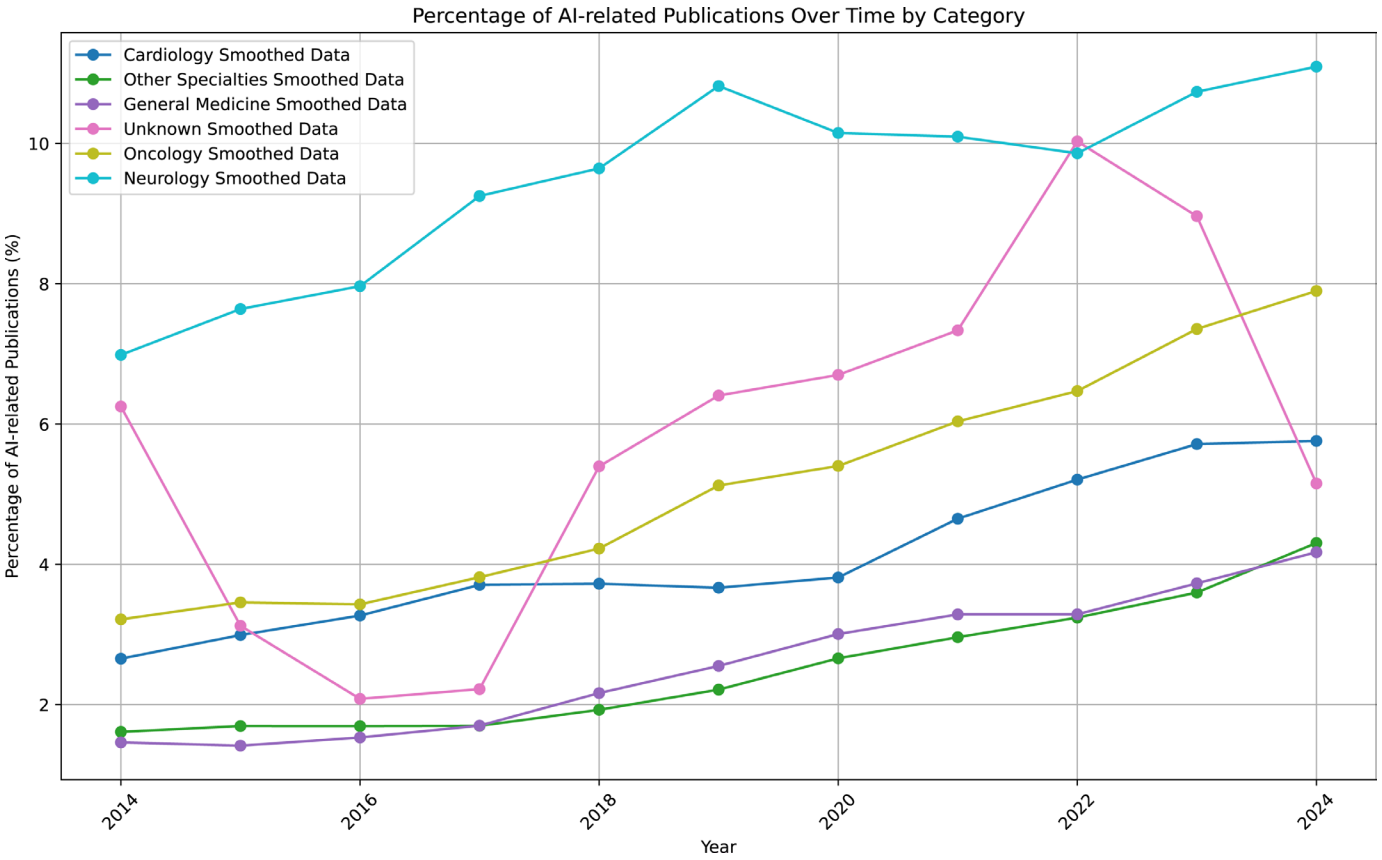

Figure 3 Percentage of AI-related publications by category. AI, artificial intelligence.

publications will conquer the medical field by 30 June, or even earlier, by 31 August, 2184 (figure 5a,b, according to the ARIMA model and ETS mode, respectively).

DISCUSSION

To our best knowledge, this is the first study to thoroughly quantify and analyse the use of AI in medical research. We divide our discussion into two parts: the first part explains the main findings of the study, their meaning and their limitations. While the second part discusses the use of the AI tools in this study and its implications on AI in medical research, addressing both the great potential and risks it may hold.

Analysing the results and their implications

Our analysis highlights the remarkable growth in AI-related publications in the medical field over the past decade. The number of AI-related articles has more than doubled, from approximately 500 in 2014 to over 1000 in 2022, before experiencing a slight dip in 2024. This dip can be attributed to the fact that the analysis is being performed mid-year and is more accurately reflected in the percentage graphs, showing a gradual growth from 2.5% in 2014 to over 6% in 2024. This trend suggests a broader acceptance and incorporation of AI across medical research. This increase might be indicative of AI's expanding role in enhancing

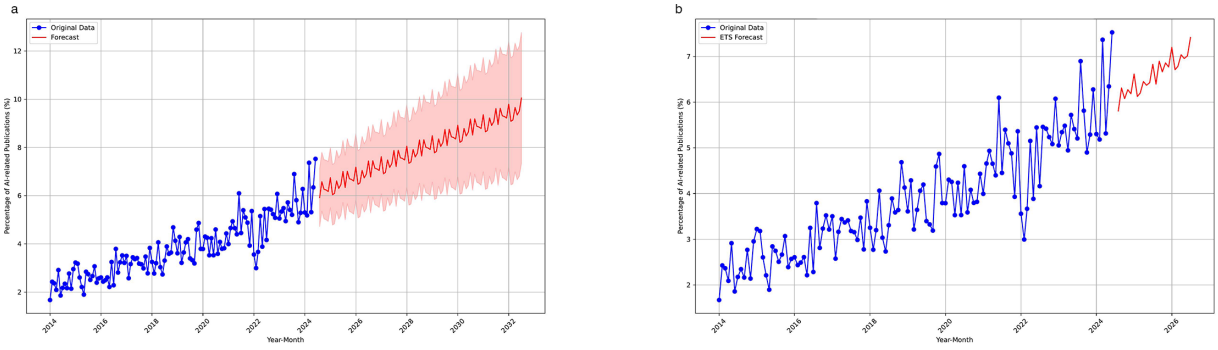

Figure 4 Percentage of AI-related publications over time (a) with ARIMA forecast, (b) with ETS forecast. AI, artificial intelligence; ARIMA, autoregressive integrated moving average; ETS, error-trend-seasonality.

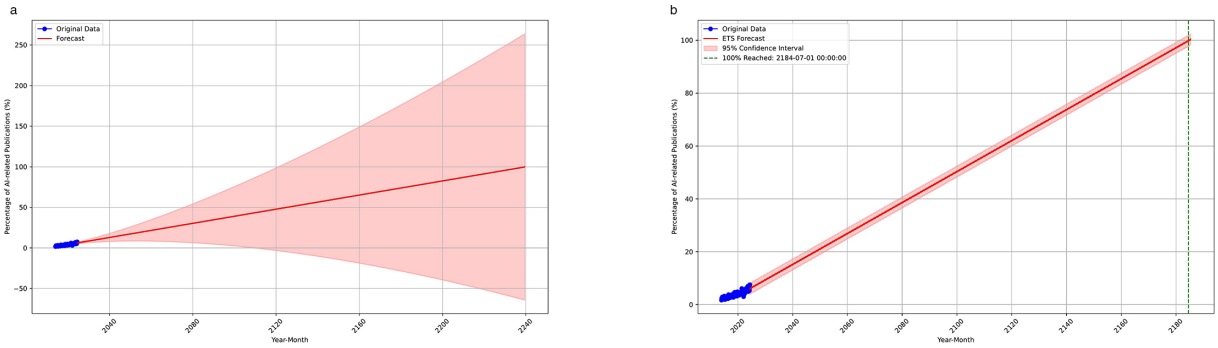

Figure 5 Long-term forecast of AI-related publications (a) with ARIMA forecast, (b) with ETS forecast. AI, artificial intelligence; ARIMA, autoregressive integrated moving average; ETS, error-trend-seasonality.

diagnostic accuracy, personalising medical education and streamlining research processes, but it may also relate to the 'trendiness' of the subject. The seasonal variations observed in the monthly analysis of AI-related publications may point to fluctuating interests and funding cycles within the academic community.

Similar manually performed studies, published in the past, reached similar trendlines and adjunct conclusions,[4 6] although, not venturing as far as our study in projecting total AI domination in medical literature.

Specialty-wise stratification of the data reveals that neurology and oncology are at the forefront of AI adoption. This trend could reflect the high stakes and complex decision-making processes inherent in these fields, where AI can provide substantial benefits. Other specialties are gradually catching up, signalling a widespread recognition of AI's value across the medical spectrum.

Our predictive models, ARIMA and ETS, forecast a continued rise in AI-related publications, potentially reaching 10% by 2030. The ARIMA model suggests that AI-related publications might dominate the medical literature by 30 June 2239, while the ETS model forecasts an even earlier date of 31 August, 2184. While these models offer valuable insights, it is essential to approach long-term predictions with caution, as many pitfalls lie within these prospects. Mainly, both models assume linearity, either directly in the time series data by the ARIMA model, or within its components (level, trend and seasonality) by the ETS model. An assumption which may not apply to the real-world future data. Moreover, both models' performance might degrade over long-term predictions due to accumulating forecast errors. These projections may also highlight one of the shortcomings of the current AI models, where focusing on a specific task might lead to an oversight of the overall picture, reminiscent of the 'paperclip apocalypse' scenario.[12]

DISCUSSING THE METHODS

The methods employed in our study leveraged the capabilities of the OpenAI API, specifically ChatGPT (OpenAI model 4o), to automate the labour-intensive task of line-by-line data analysis. The AI engine used the Python programming language to construct scripts for data collection and statistical analysis, deployed through Google Colab and local environments. While several other open and closed source models exist, and we experimented with a few, none allowed the ease and comfort of the Python API.

One of the key strengths of our approach was the ability of ChatGPT to handle over 200 000 rows of data efficiently. Instead of limiting our results to a set of preselected keywords, the model analysed each row of data individually and judged it based on its internal logic. With some studies showing the previous generation model passing the US medical examinations,[5] the AI ability to process and analyse vast amounts of information quickly could render countless research students redundant. However, this process was not without its challenges. Our initial retrieval yielded 315 209 papers, with several issues encountered during data collection. For instance, errors in retrieving data from certain journals and incomplete records required us to exclude some data to maintain the integrity of our analysis. This emphasises a potential selection bias in our study, as over 100 000 articles were excluded in a non-supervised nor random manner. Additional mistakes, like an incorrect journal or missing information, probably would not have occurred using a research assistant.

Moreover, rather than render armies of research assistants obsolete, this approach highlights a new set of skills that would be required. The first author's understanding of Python was key to this study's success, showcasing the importance of programming skills in modern research.

Despite these challenges, the AI-assisted approach significantly reduced the manual labour involved in data collection and analysis. The cost of $61.62 for OpenAI line-by-line queries, involving 260 165 API requests and 11 768 851 tokens, underscores the efficiency and scalability of this method. It is worth noting that human oversight played a crucial role in ensuring the accuracy and reliability of the AI-generated data. Repeating queries, correcting human errors and fine-tuning the process were essential steps in validating the AI's output. However, it is important to note that blind trust in the AI model, without understanding the underlying big data models and tools, might hinder the study's strength

and lead to results such as the overly ambitious 100% forecast in our study. Our study, through exploration of AI-related publication, highlights the potential of combining AI tools with human expertise to enhance research efficiency. This approach not only streamlined the data analysis process but also provided valuable insights into the evolving landscape of AI in medical research.

CONCLUSIONS

As we creep slowly into an unknown future of horrors or wonders, understanding the role of AI in our small facet of the many possibilities of the future can guide us through this increasingly fast-changing world. Our study provides an analysis of the burgeoning trend of AI-related publications in the medical field as well as a hint of one of its many uses in the future to come.

Contributors IS conceived the study idea, refined the concept alongside all coauthors and takes full responsibility as the guarantor for the integrity of the data and the accuracy of the analysis. IS and RL were primarily responsible for coding and data analysis. Literature review and manuscript drafting were conducted by NB, MH and TT, with all authors contributing to the editing and refinement of the final manuscript. All authors reviewed and approved the final version of the manuscript. The study is an AI-based study, using chatGPT (model 4o) created code for the data acquisition, analysis, presentation. Some of the data process was made through OpenAI's API (model 4o) and using chatGPT created code.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval Ethical approval was not required for this study as it was based solely on publicly available data from previously published articles. No human participants, identifiable personal data or experimental interventions were involved.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available. All data used in this study were obtained from publicly available sources. No proprietary or restricted data were used.

ORCID iD

Isaac Shochat http://orcid.org/0000-0003-0153-8956

REFERENCES

1 Kaul V, Enslin S, Gross SA. History of artificial intelligence in medicine. *Gastrointest Endosc* 2020;92:807–12.
2 Hamet P, Tremblay J. Artificial intelligence in medicine. *Metab Clin Exp* 2017;69S:S36–40.
3 Mintz Y, Brodie R. Introduction to artificial intelligence in medicine. *Minim Invasive Ther Allied Technol* 2019;28:73–81.
4 Chan KS, Zary N. Applications and Challenges of Implementing Artificial Intelligence in Medical Education: Integrative Review. *JMIR Med Educ* 2019;5:e13930.
5 Gilson A, Safranek CW, Huang T, *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:e45312.
6 Tran BX, Vu GT, Ha GH, *et al*. Global Evolution of Research in Artificial Intelligence in Health and Medicine: A Bibliometric Study. *J Clin Med* 2019;8:360.
7 Johnson D, Goodman R, Patrinely J, *et al*. Assessing the Accuracy and Reliability of AI-Generated Medical Responses: An Evaluation of the Chat-GPT Model. *Res Sq* 2023.
8 Guo X, Dong L, Hao D. RETRACTED: Cellular functions of spermatogonial stem cells in relation to JAK/STAT signaling pathway. *Front Cell Dev Biol* 2024;11:1339390.
9 Bader R, Imam A, Alnees M, *et al*. REMOVED: Successful management of an Iatrogenic portal vein and hepatic artery injury in a 4-month-old female patient: A case report and literature review. *Radiol Case Rep* 2024;19:2106–11.
10 Biswas S. ChatGPT and the Future of Medical Writing. *Radiology* 2023;307:e223312.
11 Kitamura FC. ChatGPT Is Shaping the Future of Medical Writing But Still Requires Human Judgment. *Radiology* 2023;307:e230171.
12 Bostrom N. Ethical issues in advanced artificial intelligence. *Machine Ethics and Robot Ethics* 2020;10:69–75.