

Use of Generative AI in Medical Writing by Non-Native English Researchers

Utilisation de l'Intelligence Artificielle Générative dans la Rédaction Médicale par des Chercheurs Non-Natifs Anglophones

استخدام الذكاء الاصطناعي التوليدي في الكتابة الطبية من قبل الباحثين غير الناطقين باللغة الإنجليزية

Hamza Gazzeh^{1,2}, Amal Ghribi^{2,3,4}, Youssef Zanina^{2,3,5}, Mohamed Khelil^{2,3}, Ahmed Ben Abdelaziz^{1,2,3}

1. Faculty of Medicine of Sousse, University of Sousse, Tunisia
2. Research Laboratory LR19SP01, Sahloul Hospital of Sousse, Tunisia
3. Directorate of Information Systems, Sahloul Hospital of Sousse, Tunisia
4. Faculty of Medicine of Sfax, University of Sfax, Tunisia
5. Faculty of Pharmacy of Monastir, University of Monastir, Tunisia

ABSTRACT

Introduction: Generative Artificial Intelligence (AI) has increasingly found its way into scientific medical writing, which can be particularly inappropriate in non-native English-speaking countries. This study aimed to determine the occurrence of AI-generated texts in medical publications originating from the Greater Maghreb countries (Libya, Tunisia, Algeria, Morocco, and Mauritania).

Methods: This was a cross-sectional study that gathered all medical publications indexed on MEDLINE, published in the first week of July 2024, with first author affiliated with Greater Maghreb countries. The rate of AI-generated texts was calculated using the AI detection tool: ZeroGPT®. Each article was analyzed in its entirety and each section separately (Abstract, Introduction, Methods, Results, and Discussion). Articles or sections were considered "suspects of AI generation" (sAI-g) if the rate was $\geq 25\%$. Results were presented as medians associated with their corresponding Inter Quartile Range (IQR).

Results: In all, 48 scientific medical articles were published by first authors from the Greater Maghreb countries. Articles were classified as "sAI-g" in 65% of cases, with a median rate of 36.2% [IQR=11.0%-49.4%]. AI-generated text was detected mainly in three sections: "Methods" (sAI-g=86%, median=59.3% [IQR=28.5%-71.7%]), "Abstract" (sAI-g=69%, median=52.2% [IQR=0.0%-90.2%]) and "Introduction" (sAI-g=58%, median=43.2% [IQR=0.0%-79.4%]), while the "Discussion" section had the lowest median rate (sAI-g=30%, median=10.4% [IQR=0.0%-27.9%]).

Conclusion: Scientific medical articles from the Greater Maghreb countries used Generative AI extensively. This requires, on one hand, advancing medical education and mandating dissertations in English to build capacities for non-native English-speaking researchers, and on the other hand, providing training on the responsible use of AI tools and establishing ethical guidelines to uphold academic integrity.

Key words: Generative Artificial Intelligence – ChatGPT – Chatbot – Large Language Models – Communication Barriers – Language Barriers – Medical Writing – Publication – Scientific Misconduct.

RÉSUMÉ

Introduction: L'Intelligence Artificielle (IA) Générative s'intègre de plus en plus à la rédaction scientifique médicale, ce qui pourrait être particulièrement inapproprié dans les pays non anglophones. L'objectif de cette étude a été de déterminer l'occurrence de textes générés par l'IA dans les publications médicales issues des pays du Grand Maghreb (La Libye, la Tunisie, l'Algérie, le Maroc, et la Mauritanie).

Méthodes: Il s'agissait d'une étude transversale ayant colligé toutes les publications médicales indexées sur MEDLINE, publiées durant la première semaine de Juillet 2024, dont le premier auteur a été affilié à un pays du Grand Maghreb. Le taux de textes générés par IA a été identifié par l'outil de détection d'IA ZeroGPT®. Chaque article a été analysé dans son intégralité ainsi que par sections (Résumé, Introduction, Méthodes, Résultats, et Discussion). Le texte analysé a été considéré comme « suspect de génération par IA » (sAI-g) lorsque le taux d'IA a été $\geq 25\%$. Les résultats ont été présentés par les médianes accompagnées de leurs Intervalles Inter Quartiles (IIQ).

Résultats: Au total, 48 articles médicaux ont été publiés par des premiers auteurs issus des pays du Grand Maghreb. Ces articles ont été jugés « sAI-g » dans 65% des cas, avec un taux médian de 36,2% [IIQ=11,0%-49,4%]. Les textes générés par IA ont été principalement détectés dans trois sections: « Méthodes » (sAI-g = 86 %, médiane=59,3%, [IIQ=28,5%-71,7%]), « Résumé » (sAI-g=69%, médiane=52,2%, [IIQ=0,0%-90,2%]) et « Introduction » (sAI-g=58%, médiane=43,2%, [IIQ=0,0%-79,4%]). La section « Discussion » a eu le taux médian le plus bas (sAI-g=30%, médiane=10,4%, [IIQ=0,0%-27,9%]).

Conclusion: L'usage de l'IA Générative dans les articles médicaux du Grand Maghreb a été excessif. Ceci exige, d'une part, de renforcer l'enseignement médical et la rédaction des dissertations en anglais pour renforcer les capacités des chercheurs non anglophones, et d'autre part, de former à l'usage responsable des outils d'IA et d'établir des lignes directrices éthiques pour préserver l'intégrité académique.

Mots clés: Intelligence artificielle générative- ChatGPT – Chatbot – Grands modèles de langage arge - Barrières de communication – Barrières linguistiques - Rédaction médicale - Publications – Inconduite scientifique.

Correspondance

Hamza Gazzeh

Faculty of Medicine of Sousse, University of Sousse, Tunisia. Research Laboratory LR19SP01, Sahloul Hospital of Sousse, Tunisia

Email: gazzehamza@gmail.com

المخلص

المقدمة: يشقّ الذكاء الاصطناعي التوليدي طريقه بشكل متزايد في مجال الكتابة العلمية الطبية، مما قد يكون غير ملائم في البلدان غير الناطقة بالإنجليزية بالخصوص. تهدف هذه الدراسة إلى تحديد مدى استخدام النصوص المؤدّة بالذكاء الاصطناعي في المنشورات الطبية الصادرة عن دول المغرب الكبير (ليبيا، تونس، الجزائر، المغرب، موريتانيا).

المنهجية: ضمت هذه الدراسة المقطعية جميع المنشورات الطبية المفهرسة في قاعدة بيانات MEDLINE، والتي نُشرت خلال الأسبوع الأول من شهر جويلية 2024، وكان المؤلف الأول لها منتميا لإحدى دول المغرب الكبير. احتشبت نسبة النصوص المؤدّة بواسطة الذكاء الاصطناعي باستخدام أداة الكشف ZeroGPT®. تم تحليل كل مقال بصغته الكاملة، بالإضافة إلى تحليل كل قسم منه على حدة (المخلص، المقدمة، النتائج، المناقشة). اعتُبرت المقالات أو الأقسام "مشتبها في توليدها بالذكاء الاصطناعي (sAI-g)" إذا كانت نسبة النصّ المؤدّ $\leq 25\%$. عُرضت النتائج كوسيطات مرفقة بنطاقات الرّبيعيّات (IQR).

النتائج: تم نشر 48 مقالا طبيًا من تأليف باحثين من دول المغرب الكبير. كُتبت جميع المقالات باللغة الإنجليزية، ونُشرت في 12 مجلة، منها ست مجلات مصنّفة ضمن الرّبيّع الأول (Q1) أو الثاني (Q2). صنّفت 65% من المنشورات كمشبته في توليدها بالذكاء الاصطناعي (sAI-g)، حيث بلغ المعدل الوسيط 32.6% [IQR=11.0%-49.4%]. كُشفت النصوص المولدة بالذكاء الاصطناعي في ثلاثة أقسام بالأساس: قسم المنهجية (sAI-g=86%, m=59.3%, [IQR=28.5%-17.7%])، وقسم الملخص (sAI-g=69%, m=52.2%, [IQR=0.0%-90.2%])، وقسم المقدمة (sAI-g=58%, m=43.2%, [IQR=0.0%-79.4%]) في المقابل، كان أدنى وسيط في قسم المناقشة (sAI-g=30%, m=10.4%, [IQR=0.0%-27.9%]).

الختامة: كشفت الدراسة عن استخدام واسع للذكاء الاصطناعي التوليدي في المقالات الطبية الصادرة عن دول المغرب الكبير. يتطلب هذا الأمر، من جهة، تعزيز التعليم الطبي باللغة الإنجليزية واشتراط كتابة الأطروحات بها لدعم الباحثين غير الناطقين بها، ومن جهة أخرى، تقديم تدريب على الاستخدام المسؤول لأدوات الذكاء الاصطناعي ووضع إرشادات أخلاقية للحفاظ على النزاهة الأكاديمية.

الكلمات المفتاحية: الذكاء الاصطناعي التوليدي - شات جي بي تي - روبوت الدردشة - الحواجز التواصلية - الحواجز اللغوية - الكتابة الطبية - المنشورات العلمية - سوء السلوك العلمي

INTRODUCTION

The use of Generative Artificial Intelligence (Gen-AI) tools in scientific writing has become increasingly prevalent [1]. These tools can assist in various aspects of the research process, including idea generation, content structuring, literature review, data analysis, and editing [2–4]. Furthermore, Gen-AI can significantly improve writing clarity, style, and coherence, especially for non-native English speakers, by correcting grammatical errors and suggesting appropriate phrasing [5].

Despite these potential benefits, the use of Gen-AI tools raises ethical concerns including bias, misinformation [6], and academic integrity issues [7,8]. Indeed, the ability of these tools to generate original-sounding text complicates the detection of plagiarism, making it more challenging to identify such instances of scientific misconduct [9]. These effects would be more anticipated in non-English-speaking countries, given that academic recognition is based on publications in English in prestigious journals with high citation rates [10].

However, the extent of the use of Gen-AI tools in scientific medical publications authored by non-native English-speaking researchers is poorly documented, the use of Gen-AI tools by non-native English-speaking researchers in scientific medical publications is poorly documented, especially regarding their prevalence, researchers' dependency on these tools, and the frequency of their occurrence, particularly in Africa and the Greater Maghreb region. Conducting such a study would raise awareness among all stakeholders in scientific writing in health sciences (authors, supervisors, reviewers, editors, educational managers, and scholarly societies) about the

need to regulate this phenomenon. This would enhance the alignment of publications with research ethics and societal responsibility principles. Hence, this pilot study aimed to determine the occurrence of AI-generated texts in medical writing within publications originating from the Greater Maghreb countries (Libya, Tunisia, Algeria, Morocco, and Mauritania), published online during the first week of July 2024.

METHODS

This was a cross-sectional study that gathered all publications indexed in MEDLINE, published between the first and seventh of July 2024, a convenient period sufficient to collect a statistically significant number of articles, affiliated with a country in the Greater Maghreb region. Only publications written in English or French, with both an abstract and full text available in open access, were included. After conducting a search query on PubMed, articles were excluded if they were identified as non-medical through a screening of titles and abstracts by the authors or if the first author was not affiliated with a country in the Greater Maghreb (Figure 1).

For each article, bibliometric characteristics were collected, including the study design, the article's field based on the specialty of its first author, the publishing journal, and its ranking based on percentiles. The study design was identified using the standardized classification of article types provided by the PubMed platform., and the journal's ranking was obtained from the SCImago Journal Rank (SJR) [11]. Each publication was then analyzed using the AI detection tool ZeroGPT® (<https://www.zerogpt.com>), both in its entirety (full text) and by individual

sections (Abstract, Introduction, Methods, Results, and Discussion). This tool highlights every sentence identified as written by AI (ChatGPT, GPT4 and Gemini) and provides a percentage gauge, labeled as "AI GPT*", showing the overall proportion of AI-generated content in the text [12]. The full text or individual sections were classified by consensus among the researchers as "suspected of AI generation" (sAI-g) when the detected rate was $\geq 25\%$. Data were analyzed using both RStudio® with version 4.1.1 of R® software. Categorical variables were analyzed using absolute frequencies, while quantitative variables were assessed using medians and quartiles. Results were presented in both tabular and graphical formats, with plots created using the "ggplot2" R® package.

RESULTS

After conducting our search query, 93 scientific articles were identified, ultimately resulting in 48 articles selected for analysis, as shown in the flowchart (Figure 1). All articles were written in English. The type of publication was, in 88% of instances, "case reports". The specialty of the first author was surgical in 69% of cases. Articles were published in 12 different journals, six of which were ranked in quartiles Q1 or Q2. The top three journals publishing these articles were "International Journal of Surgery Case Reports" (n=27), "Radiology Case Reports" (n=6), and "Urology Case Reports" (n=6) (Table 1).

The median "AI GPT*" rate for the full text was 36.2%, with an Inter Quartile Range (IQR) of [11.0%-49.4%]. The sections with the highest median rates were the "Methods" (59.3%, [IQR=28.5%-71.7%]), "Abstract" (52.2%, [IQR=0.0%-90.2%]), and "Introduction" (43.2%, [IQR=0.0%-79.4%]), while the "Discussion" section had a median rate of 10.4% with an IQR of [0.0% - 27.9%] (Figure 2). Publications were classified as sAI-g in 65% of cases, with the three sections having the highest rate of sAI-g texts being the "Methods" (85.7%), "Abstract" (68.8%), and "Results" (66.7%), respectively (Table 2).

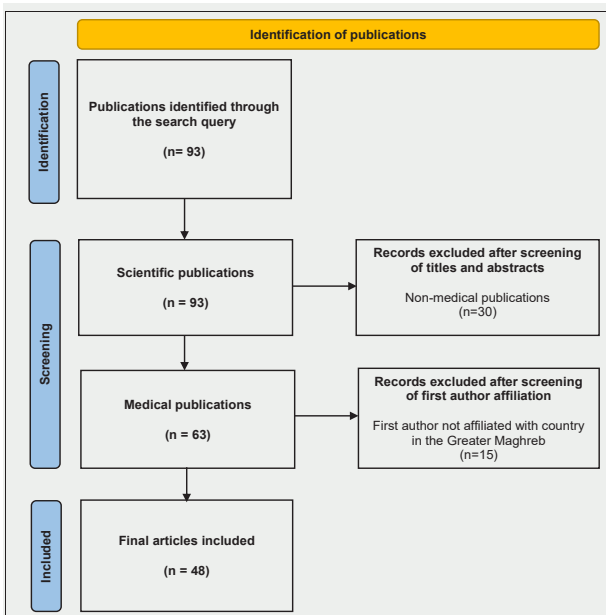


Figure 1. Flowchart of the Study Selection Process: Inclusion of Articles Indexed in MEDLINE, published between the first and seventh of July 2024, and affiliated with a country in the Greater Maghreb Region

Table 1. Bibliometric characteristics of 48 Greater Maghreb’s articles indexed in MEDLINE and published between the first and the seventh of July 2024

	N	%
Affiliation		
Tunisia	24	50.0
Morocco	23	47.9
Libya	1	2.1
Specialty		
Surgical	33	68.8
Urology	11	22.9
Orthopedic Surgery	6	12.5
Gynecology-Obstetrics	5	10.4
General Surgery	4	8.3
Others	7	14.6
Medical	11	22.9
Radiology	4	8.3
Oncology	3	6.3
Others	4	8.3
Fundamental	4	8.3
Study design		
Case reports	42	87.5
Observational	4	8.3
Experimental	1	2.1
Guidelines	1	2.1
Journal		
International Journal of Surgery Case Reports	27	56.3
Radiology Case Reports	6	12.5
Urology Case Reports	6	12.5
Cancer Immunology. Immunotherapy : CII	1	2.1
Primary Care Diabetes	1	2.1
Annals of Medicine and Surgery	1	2.1
BMC Cancer	1	2.1
BMC Psychology	1	2.1
Qatar Medical Journal	1	2.1
Journal of Surgical Case Reports	1	2.1
Translational Oncology	1	2.1
Oxford Medical Case Reports	1	2.1
Journal quartile based on SJR*		
Q1	4	8.3
Q2	2	4.2
Q3	35	72.9
Q4	7	14.6

*SJR = SClmago Journal Rank

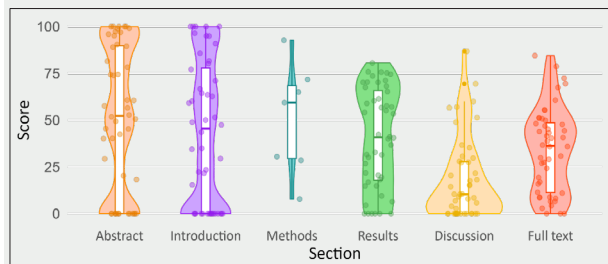


Figure 2. Violin plot of percentage of AI-Generated texts, determined by ZeroGPT® tool, in the full text and sections of 48 Greater Maghreb’s articles published between the first and seventh of July 2024 and indexed in MEDLINE

Table 2. Percentage of full texts and sections classified as suspect of Artificial Intelligence generation from Greater Maghreb's articles published between the first and seventh of July 2024 and indexed in MEDLINE

Classified as sAI-g*	N	n	%
Section			
Abstract	48	33	68.8
Introduction	48	28	58.3
Methods	7	6	85.7
Results	48	32	66.7
Discussion	47	14	29.8
Full Text	48	31	64.6

sAI-g*: suspects of Artificial Intelligence generation (rate of "AI GPT*" in ZeroGPT® tool ≥ 25%)

DISCUSSION

This study reveals that the use of Gen-AI tools in medical publications originating from the Greater Maghreb region was extensive (median "AI GPT*" rate =36.2%, IQR [11.0%-49.4%]). Given that the sample predominantly consisted of case reports, this highlights their susceptibility to AI-generated content. The AI usage was particularly concentrated in the "Abstract", which was the most affected section. In contrast, the "Discussion" section remained relatively spared.

In all, 65% of the articles had full texts suspected of being generated by AI tools, meaning these articles exhibited AI generation rates exceeding 25%. This observation aligns with the few studies conducted [13–15], which have revealed a significant rise in the use of Gen-AI tools in medical writing, specifically since the launch and widespread adoption of Gen-AI tools such as ChatGPT. However, this rate was significantly higher than that reported by a similar study [13] using the same detection tool, where only 16% of articles had a suspected AI generation rate greater than 20%. Indeed, in the cited study, the ZeroGPT® tool was utilized to assess AI use rates in English articles submitted to the "Orthopaedics and Traumatology: Surgery & Research" journal. The average AI use rate across all articles was 11% ± 6, with 41% of articles exceeding 10% AI use and 15.6% exceeding 20%. This notable difference can partly be explained by the type of publications analyzed, primarily case reports (88%), often written by young residents in surgical or medical specialties. Case reports possess distinct stylistic features that set them apart from other genres of medical discourse. They usually employ a more literary vocabulary [16] to describe the observations related to anamnesis, clinical data, and disease progression, as well as to extract lessons for optimizing diagnostic and therapeutic tasks in clinical settings. This narrative style requires a nuanced understanding of both the medical context and the language in which the report is being written. However, for young authors who received their medical training in non-English languages, crafting such narrative texts can be particularly challenging. This may lead them to rely more heavily on Gen-AI tools for structuring or drafting their work, as an attempt to compensate for their limited familiarity with strict editorial standards, such as authorship criteria, manuscript preparation guidelines,

and ethical reporting standards, which are outlined by organizations like the "International Committee for Medical Journal Editors", as an example.

Despite this probable overestimation, the rate remains remarkably high, which may be attributed to the study's exclusive focus on non-native English-speaking researchers. A study conducted in 2023 [13] showed that Asian geographic origin was a significant risk factor for supposedly higher AI rates. Another study [14] also indicated that there were similarly high AI rates in non-English-speaking countries, including China, Japan, Korea, and South America. This suggests a higher reliance on Gen-AI tools to overcome language barriers [17], as these tools are particularly effective in correcting grammatical errors and improving writing style, thereby making the text clearer and more succinct [5], especially in complex sections where non-native speakers often struggle [18]. However, there is a risk associated with overreliance on these tools, as it can diminish human interaction and critical thinking skills [19]. Such dependence may also perpetuate the dominance of English in scientific discourse, reinforcing linguistic imperialism [20], which refers to the monopolistic dominance of English over native languages, marginalizing non-English linguistic diversity. Other hypotheses that could explain the elevated rate of Gen-AI use include the occurrence of false positives or the use of automatic translators, which may increase the perceived AI generation rate.

The sections with the highest median rates of AI-generated texts were the "Methods", "Abstract", and "Introduction", respectively. However, it is important to note that the elevated rate in the "Methods" section may be biased, as only seven articles included this section (14.5%), while the others were case reports. As for the "Abstract," our findings were consistent with those of Bisi et al. [14] who have identified this section as having the highest prevalence of AI-generated text. Similarly, Miller et al. [15] reported an increase in the prevalence of AI-generated text in MEDLINE-indexed abstracts from 21.7% to 36.7% between 2020 and 2023. One possible explanation for this elevated rate is that the abstract is often the last part written by authors of scientific manuscripts, typically under conditions of writing fatigue and editorial urgency. Additionally, writing abstracts presents several challenges for researchers, particularly non-native English speakers and graduate students, as summarizing extensive research into a concise format while providing adequate background and effectively presenting findings can be particularly difficult [21]. This context may lead inexperienced authors to automate the formulation of the abstract using Gen-AI tools, enabling them to generate this critical section with a single click. However, while such automation facilitates the task, it warrants scrutiny. Authors who have conducted the scientific work themselves are generally better equipped to identify the key messages of a study in a structured manner, effectively conveying the main objective, the type of study, the principal findings, and the study's conclusion.

The "Discussion" section appears to be the least affected by Gen-AI tools (median rate =10.4%, IQR=[0.0%-

27.9%]), despite its somewhat elevated rate of AI-generated content (sAI-g=30%). This section, particularly relevant to original clinical research papers, seems to still resist full automation. This could reflect the efforts of research teams in non-English-speaking countries to analyze their own results in a contextually appropriate way, considering professional, contextual, and cultural aspects. Furthermore, the complexity and critical nature of this section [22,23], which requires authors to interpret and evaluate their findings, aligns with the broader observation that critical thinking potential remains a key factor in sections that demand deeper analysis [24]. This might explain why researchers are more cautious and less reliant on AI for this section.

Study Limitations

This study possesses several strengths. Firstly, it is a pilot study conducted in non-English-speaking countries, offering valuable insights into the use of Gen-AI tools in medical writing among researchers for whom English is neither the native nor the official language. Secondly, the articles analyzed were all published following a formal peer-review process, ensuring a certain level of scientific rigor. Thirdly, they are indexed in MEDLINE, a highly reputable reference database in the field of medical research. This foundation can serve as a stepping stone for future, more extensive investigations in this area. However, some limitations must be acknowledged. The reliance on a single AI detection tool, ZeroGPT®, presents a notable constraint, as it was not subjected to reproducibility tests to assess its reliability. Although reproducibility seems to be reasonably good for longer texts, such as those analyzed in this study, there is currently no formal scientific evaluation of this detection software. Additionally, the sensitivity of ZeroGPT® can be limited, as it may not always be very accurate in detecting low-level AI-generated content. This could lead to an inflated estimation of the prevalence of AI-generated text and may undermine the accuracy of the conclusions. As large language models continue to evolve, the emergence of various AI detection tools underscores the necessity for future research to evaluate and compare their reliability in identifying AI-generated content. Another limitation of this study is the lack of consideration for ephemeral chat usage, such as the temporary mode in ChatGPT, which leaves no detectable trace. This feature, prioritizing privacy and security, could lead to an underestimation of AI-generated content, as interactions in ephemeral mode would not be captured by detection tools like ZeroGPT®. Future studies should address this gap to provide a more accurate assessment of AI usage in scientific writing. As a preliminary measure, the significant AI GPT* rate threshold of 25% was chosen, according to the plagiarism limits accepted for academic dissertations by scientific institutions in Tunisia [25]. However, there is currently no established “acceptable” AI-generated content rate in the literature. Another limitation of this study was the short timeframe of scientific publications analyzed, which spanned only one week. Such a limited period may not fully capture the broader trends and variations in Gen-AI tool usage in

medical writing. Further research over extended periods would be necessary to gain more comprehensive insights. Additionally, the sample composition was predominantly dominated by case reports, a descriptive and narrative form of publication that may not represent the broader spectrum of scientific medical writing. Expanding future studies to include a more diverse range of publication types would offer a clearer understanding of AI's impact across various formats of medical writing.

CONCLUSION

This pilot study confirms the extensive use of Gen-AI tools by non-English-speaking researchers in medical writing. While these tools offer significant benefits, they also present risks to the scientific integrity of publications and the credibility of research institutions. Those concerns about academic misconduct, data privacy, and algorithmic bias often fuel resistance and stigma toward these technologies [26]. Therefore, it is imperative to develop a roadmap involving all stakeholders in scientific publishing, particularly within health sciences: authors, reviewers, editors, academic managers, and scholarly societies. This roadmap should fundamentally emphasize three critical areas: enhancing researchers' proficiency in the English language, promoting scientific integrity throughout the publication process, and fostering the appropriate use of AI tools as supportive aids. AI tools should be framed as facilitators rather than replacements, helping researchers improve clarity, refine structure, and streamline technical aspects of writing without compromising originality or ethical standards. Establishing clear guidelines on their use, coupled with training programs to ensure ethical and effective integration, will enable researchers to harness these technologies responsibly, thereby advancing the quality and credibility of scientific output (Box 1).

Availability of Data and Materials: The data and materials used in this study are publicly available and the publications are provided with free open access.

Ethical Approval and Consent to Participate: Ethical approval was not required as the study utilized publicly available data, per established research guidelines.

Competing Interests: The authors declare no conflict of interest.

Funding: No funding was received for this work.

Declaration: The authors' contributions were constituted by conception and study design (H.GA, A.BA), data collection (H.GA, A.GH), data analysis (H.GA, A.GH, Y.ZA), writing (H.GA), reviewing (all authors), and supervision (A.BA).

AI Usage: The authors declare the use of chatbot to assist with grammar checking and language refinement during the preparation of this article.

Previous Presentation Declaration: The results of this article were previously presented at the 49th Congress of the Tunisian Society of Medical Sciences (STSM), which took place in Tunis on September 13 and 14, 2024. Additionally, the corresponding abstract was published on October 8, 2024, in *La Tunisie Médicale*. (Reference: Gazze H, Ghribi A, Zanina Y, Khelil M, Ben Abdelaziz A. Misuse of Generative Artificial Intelligence in Scientific Publications by Non-Native English speakers. *Tunis Med.* 2024;102(10):834)

Box 1. Roadmap for Streamlining the Use of AI tools by Non-native English speaking Researchers to Uphold Scientific Integrity

- 1- **Enhancing Researchers' Proficiency in Scientific English:** Strengthen medical English education in health science faculties (medicine, pharmacy, dentistry, nursing, etc.), with a focus on scientific communication at conferences and the preparation of dissertations.
- 2- **Establishing Scientific Manuscript Translation Services:** Provide professional translation services for scientific manuscripts through academic research structures (universities) and healthcare institutions (university hospitals), as well as through health science journals in non-English-speaking countries.
- 3- **Promoting Native Language Publications:** Encourage the development of articles and academic dissertations in the native languages of the authors, and support the creation of international journals based on the languages used in higher education.
- 4- **Ensuring Authenticity in Academic and Editorial Work:** Require the verification of scientific work within academic and editorial contexts by employing AI detection software, alongside plagiarism filters, to ensure authenticity, with consensually appropriate thresholds.
- 5- **Establishing Guidelines for Ethical AI Use in Manuscript Preparation:** Develop guidelines for authors and journals defining the appropriate use of AI tools, including the requirement for AI usage disclosure, to uphold scientific integrity scientific integrity in manuscript preparation and submission.
- 6- **Training in Responsible Use of AI Tools:** Provide training on the capabilities and limitations of AI tools, emphasizing their role as aids in writing and the importance of critically assessing AI-generated outputs.
- 7- **Conduct Further Studies:** Extend this pilot study into a broader comparative investigation of the use of AI tools in scientific publications, both before and after their widespread adoption. Such investigation should utilize multiple Gen-AI detection tools and concentrate on countries in Africa where English is neither the native language nor the official language of medical education.

REFERENCES

1. Fatima A, Shafique MA, Alam K, Fadlalla Ahmed TK, Mustafa MS. ChatGPT in medicine: A cross-disciplinary systematic review of ChatGPT's (artificial intelligence) role in research, clinical practice, education, and patient interaction. *Medicine (Baltimore)*. 2024;103(32):e39250.
2. Khalifa M, Albadawy M. Using artificial intelligence in academic writing and research: An essential productivity tool. *Comput Methods Programs Biomed*. 2024;5:100145.
3. Dergaa I, Ben Saad H, Glenn JM, Ben Aissa M, Taheri M, Swed S, et al. A thorough examination of ChatGPT-3.5 potential applications in medical writing: A preliminary study. *Medicine*. 2024;103(40):e39757.
4. Dergaa I, Chamari K, Zmijewski P, Saad HB. From human writing to artificial intelligence generated text: examining the prospects and potential threats of ChatGPT in academic writing. *Biol Sport*. 2023;40(2):615–22.
5. Giglio AD, Costa MUPD. The use of artificial intelligence to improve the scientific writing of non-native english speakers. *Rev Assoc Med Bras*. 2023;69(9):e20230560.
6. Methnani J, Latiri I, Dergaa I, Chamari K, Ben Saad H. ChatGPT for Sample-Size Calculation in Sports Medicine and Exercise Sciences: A Cautionary Note. *Int J Sports Physiol Perform*. 2023;18(10):1219–23.
7. Liu H, Azam M, Bin Naeem S, Faiola A. An overview of the capabilities of ChatGPT for medical writing and its implications for academic integrity. *Health Info Libraries J*. 2023;40(4):440–6.
8. Doyal AS, Sender D, Nanda M, Serrano RA. Chat GPT and Artificial Intelligence in Medical Writing: Concerns and Ethical Considerations. *Cureus*. 2023;15(8):e43292.
9. Steponenaite A, Barakat B. Plagiarism in AI Empowered World. In: Antona M, Stephanidis C, editors. *Universal Access in Human-Computer Interaction*. Cham: Springer Nature Switzerland; 2023. p. 434–42.
10. Di Bitetti MS, Ferreras JA. Publish (in English) or perish: The effect on citation rate of using languages other than English in scientific publications. *Ambio*. 2017;46(1):121–7.
11. Scimago Journal & Country Rank [Internet]. [Last visit: 2024 Oct 9]. Available from: <https://www.scimagojr.com/>
12. AI Detector - Trusted AI Checker for ChatGPT, GPT4 & Gemini [Internet]. [Last visit: 2024 Oct 9]. Available from: <https://www.zerogpt.com/>
13. Maroteau G, An J-S, Murgier J, Hulet C, Ollivier M, Ferreira A. Evaluation of the impact of large language learning models on articles submitted to *Orthopaedics & Traumatology: Surgery & Research (OTSR)*: A significant increase in the use of artificial intelligence in 2023. *Orthop Traumatol Surg Res*. 2023;109(8):103720.
14. Bisi T, Risser A, Clavert P, Migaud H, Dartus J. What is the rate of text generated by artificial intelligence over a year of publication in *Orthopedics & Traumatology: Surgery & Research*? Analysis of 425 articles before versus after the launch of ChatGPT in November 2022. *Orthop Traumatol Surg Res*. 2023;109(8):103694.
15. Miller LE, Bhattacharyya D, Miller VM, Bhattacharyya M. Recent Trend in Artificial Intelligence-Assisted Biomedical Publishing: A Quantitative Bibliometric Analysis. *Cureus*. 2023;15(5):e39224.
16. Lysanets Y, Morokhovets H, Bieliaieva O. Stylistic features of case reports as a genre of medical discourse. *J Med Case Rep*. 2017;11(1):83.
17. Zenni RD, Andrew NR. Artificial Intelligence text generators for overcoming language barriers in ecological research communication. *Austral Ecology*. 2023;48(7):1225–9.
18. Hwang SJ, Lim JS, Lee RW, Matsui Y, Iguchi T, Hiraki T, et al. Is ChatGPT a "Fire of Prometheus" for Non-Native English-Speaking Researchers in Academic Writing? *Korean J Radiol*. 2023;24(10):952.
19. Waluyo B, Kusumastuti S. Generative AI in student English learning in Thai higher education: More engagement, better outcomes? *Soc Sci Humanit Open*. 2024;10:101146.
20. Al-Ahmadi M, "The Use of Artificial Intelligence Large Language Models in Scientific Research: Blessing or Curse? Ethical or immoral?" (2023). *AMCIS 2023 TREOs*. 118. [Last visit: 2024 Oct 10]. Available from: https://aisel.aisnet.org/treos_amcis2023/118
21. Ben Abdelaziz A, Sakly N, Melki S, Noura S, Ben Abdelaziz A, Babba O, et al. Preparation of the Structured Summary of a Biomedical Scientific article. *Tunis Med*. 2021;99(7):706–13.
22. Jenicek M. How to read, understand, and write 'Discussion' sections in medical articles. An exercise in critical thinking. *Med Sci Monit*. 2006;12(6):SR28-36.
23. Ben Abdelaziz A, Sakly N, Melki S, Noura S, Ben Abdelaziz A, Babba O, et al. The 5x5 approach in scientific biomedical writing. *Tunis Med*. 2021;99(6):585–600.
24. Mlika M, Naceur A, Kpazai G, El Mezni F, Zakhama L, Labbene I, et al. Is critical thinking potential correlated to critical appraisal skills of medical literature? *Tunis Med*. 2023;101(7):602–8.
25. Faculty of Medicine of Tunis-Home [Internet]. *FMT*. 2025 [Last visit: 2025 Jan 8]. Available from: <https://fmt.rnu.tn/wp-content/uploads/2021/12/02.-Procedures-de-la-these-de-Doctorat-en-medecine.pdf>
26. Dergaa I, Fekih-Romdhane F, Glenn J, Fessi M, Chamari K, Dhahbi W, et al. Moving Beyond the Stigma: Understanding and Overcoming the Resistance to the Acceptance and Adoption of Artificial Intelligence Chatbots. *New Asian J Med*. 2023;1:29–36.