


Application of artificial intelligence (AI) in the creation of discharge summaries in psychiatric clinics

The International Journal of
Psychiatry in Medicine
2025, Vol. 60(3) 330–337
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/00912174241284730
journals.sagepub.com/home/ijp


Bertrand Janota, PhD¹  and Krzysztof Janota²

Abstract

Background: The integration of artificial intelligence (AI; ChatGPT 4.0) into medical workflow presents a great potential to enhance efficiency and quality. The use of AI in the creation of discharge summaries is particularly promising. The course of each hospitalization is described in the discharge summary, which is given to each patient and then to his general practitioner at the end of hospital treatment. An analysis of discharge summaries in psychiatric clinics indicates that these documents must fulfill diverse and specific requirements. Nevertheless, AI-generated discharge summaries provided an opportunity to optimize information transfer and alleviate physician workload.

Method: This study evaluated the quality of discharge summaries produced by clinical staff compared to those produced by an AI model (ChatGPT 4.0). Clinicians who wrote the discharge summaries were not informed about the study's purpose or methodology. The completed summaries were subsequently assessed by four attending physicians using predefined criteria. These physicians were also blinded to the study's objectives and were unaware of the individual authors of the summaries. The evaluation criteria included consistency, completeness, and comprehensibility. Additionally, the

¹Clinic for Psychiatry and Psychotherapy, Asklepios Fachklinikum, Lübben, Germany

²Medical Student at Ruprecht-Karl-Universität, Heidelberg, Germany

Corresponding Author:

Bertrand Janota, Asklepios Fachklinikum Lübben Clinic for Psychiatry and Psychotherapy Luckauer Str.17, Lübben 15907, Germany.

Email: b.janota@asklepios.com

time required to prepare these summaries and its impact on overall quality were analyzed.

Results: The results indicated that discharge summaries generated by AI were more efficient than discharge summaries prepared by clinic staff. AI was particularly effective in terms of coherence and information structure.

Conclusion: Further research, training and development is needed to improve the accuracy and reliability of AI-generated discharge summaries.

Keywords

artificial intelligence, ChatGPT 4, medical documentation, discharge summaries, quality

Introduction

The creation of adequate medical documentation, particularly discharge summaries (also known as discharge papers or discharge letters in German-speaking countries), is a critical and essential responsibility of physicians. These documents should include a description of the reason for the hospital admission, whether it was acute or elective, the course of treatment, psychotherapy, any complications that occurred and further recommendations for post-hospital care. Discharge summaries are provided to the patient and their outpatient physician upon discharge. Unfortunately, this task is increasingly time-consuming and often represents an additional burden, especially for young doctors and psychotherapists in training. Outpatient physicians, as indicated by various studies, believe that discharge summaries should be concise, informative, and easy to read.¹⁻³ However, young doctors often struggle with creating these documents, sometimes lacking the time needed to properly analyse and describe each case. This can result in reports filled with empty phrases, standard templates, or delays in producing discharge summaries. Additionally, there is often a lack of basic knowledge about the essential components of these documents: what must be included and what can be omitted.

In psychiatric clinics, where every patient interaction constitutes a form of examination and where proper documentation is even more crucial, maintaining a balance between patient care and documentation can be challenging. The pressure on resident physicians to create swift and accurate documentation is substantial and increasingly difficult to meet. An additional challenge arises from the foreign origins of some physicians, for whom the correct grammatical, orthographical and clinical composition of a discharge summary can be particularly challenging. The prepared discharge summaries are typically reviewed by an attending physician for content, who often also correct grammatical and orthographical errors, a process that can lead to frustration and consumes valuable time.

Recently, there has been growing discussion about the use of artificial intelligence (AI) in daily life, including its potential integration into medical practice.⁴⁻¹¹ Clinical

experience indicates that AI is already being incorporated into medical tasks, particularly in the creation of medical documentation such as discharge summaries. The present study was designed to explore the potential use of AI in generating discharge summaries within a psychiatric clinic.

Method

Based on clinical experience, medical histories and short treatment courses, two fictional patients were created using standard admission forms and supplemented with fictional disease progressions. The documentation resembled patient information in bullet-point form, similar to concise documentation by on-duty doctors. The documentation was given to six clinic staff members, with three receiving Case 1 and three receiving Case 2. The participants were resident physicians (3) and psychotherapists (3). They were asked to write standardized discharge reports and record the time required. None knew the reports' purpose or discussed the task with others, ensuring unbiased assessments. Three participants received data for Case 1, involving a patient with depressive disorders, sleeping pill dependency, and hypertension, totaling 577 words. Similar data was given to ChatGPT 4.0 to create a standardized discharge report. Another group received data for Case 2, involving a patient with paranoid schizophrenia, alcohol dependency, and arterial hypertension, totaling 394 words. ChatGPT 4.0 also received this data to generate a discharge report. ChatGPT 4.0 received a five-minute training with two sample discharge reports. After two weeks, participants submitted their completed reports and discussed difficulties encountered, such as incomplete data, lack of patient knowledge and concerns about the evaluation. All reports, including AI-generated ones, were coded and sent to four attending physicians for evaluation. Each doctor received descriptions of two cases, evaluation criteria, and individual reports. They rated the reports on a scale from 1 to 3 according to specific criteria. The attending physicians were not informed of the exact purpose of the study, did not know who wrote each discharge summary and were not aware that one of the discharge summaries was prepared by AI. They were informed that the purpose of the assessment was to improve the quality of medical records at the clinic.

Evaluation criteria

The evaluation of the report focused on several key criteria. First, the information provided must be clear and concise, ensuring that all details are brief yet comprehensible. Logical flow and coherence are essential, with clear transitions between sections to maintain the reader's understanding. Proper grammar and orthography are required for correct language use, alongside the application of standardized medical terminology. The report should be well-organized in its structure and formatting, ensuring that all essential information is included for completeness. It is important that the content is understandable for outpatient doctors, being clear and precise, while also ensuring that the data is current and accurate. Correct referencing of examinations and

therapies is necessary for accuracy. Additionally, the time taken to prepare the report was assessed, with up to 15 minutes earning 3 points, 15 to 30 minutes earning 2 points, and over 30 minutes earning 1 point.

Results

The lowest total score was 73 points, and the highest was 103 points while the maximal possible score per evaluation was 33 points, thus with a total possible score of 132 points.

Evaluation of discharge summaries

Case 1 (patient with depressive disorders, sleeping pill dependency, and hypertension) The highest score was 101 points for a discharge summary prepared by a psychotherapist, excelling in “Logic and Coherence,” “Medical Terminology,” and “Structure and Formatting,” prepared in 90 minutes. The AI-generated report placed second scoring 97 points, excelling in “Medical Terminology” and “Understandability for the Outpatient Doctor,” prepared in 15 minutes.

Case 2 (patient with paranoid schizophrenia, alcohol dependency, and arterial hypertension) The AI-generated report received the highest score of 103 points, excelling in “Logic and Coherence” and “Medical Terminology,” prepared in 15 minutes. The second-best report, written by the physician, scored 98 points, also excelling in “Logic and Coherence” and “Medical Terminology,” prepared in 40 minutes [Table 1](#).

Discussion

In evaluating Case 1, the discharge letter written by the psychotherapist received the highest rating, scoring 4 points more than the AI-generated summary. This case involved a patient with depressive disorders where psychotherapy was significant. The psychotherapist’s empathy and experience contributed to a higher-quality letter. Notably, the provided patient text comprised 577 words, expanded by the staff member to 1293 words, while ChatGPT 4.0 condensed it to 361 words. The clinic staff member added fictional results from psychological tests and other examinations, potentially leading to a better overall evaluation. This raises the question if the evaluation would be different if the staff member had adhered strictly to the provided data.

The time required to prepare the highest-rated letter was 90 minutes. This raises questions about whether a discharge summary with 1293 characters, written in approximately 90 minutes, meets outpatient doctors’ and workplace time management expectations. The AI-generated discharge summary scored 4 points less than the clinic employee’s but was written in significantly less time, was more compact, and adhered closely to the facts. However, it lacked “emotional/human expression.”

In Case 2, the AI-generated discharge summary received the highest rating. The patient was treated in the admissions ward, focusing on medication management and

Table I. Evaluation table of the individual discharge summaries. AI (Chat GPT 4.0).

Discharge summary	Case 1				Case 2			
	I	II	III	AI	I	II	III	AI
Time - according to the writer	35 Min	25 Min	90 Min	15 Min	50 Min	40 Min	60 Min	15 Min
Clarity and Conciseness	10	9	9	9	10	10	7	10
Logic and coherence	10	9	11	10	11	11	9	11
Grammar and orthography	8	7	10	10	10	9	6	10
Medical terminology	10	9	11	11	11	11	10	11
Structure and formatting	9	7	11	9	10	9	6	10
Completeness of information	9	6	10	8	8	9	8	10
Comprehensibility and clarity for the outpatient physician	9	8	9	11	10	10	7	10
Comprehensibility for patients	10	10	9	10	9	9	7	9
Currency and accuracy of data	9	7	10	9	9	10	6	10
Correct referencing of Conducted examinations and therapies	8	8	10	7	8	9	6	9
Time expenditure - points	1	2	1	3	1	1	1	3
Total score	93	82	101	97	97	98	73	103
Number of words	727	261	1293	361	450	531	1037	487

Case 1 (fictional patient with depressive disorders, sleeping pill dependency, and hypertension, totaling 577 words).

I – discharge summary written by physician.

II – discharge summary written by physician.

III – discharge summary written by psychotherapists.

AI – discharge summary written by ChatGPT 4.0.

Case 2 (fictional patient with paranoid schizophrenia, alcohol dependency, and arterial hypertension, totaling 394 words).

I – discharge summary written by psychotherapists.

II – discharge summary written by physician.

III – discharge summary written by physician.

AI – discharge summary written by ChatGPT 4.0.

treatment of the acute condition. The worst-rated letter in this case was the longest, suggesting that brevity, precision, and formality are valued higher for acute conditions. Hospitals and outpatient facilities are increasingly using AI to improve process quality, structure, and speed. Integrating AI into Hospital Information Systems can shorten the discharge summary creation processes and impact their quality.^{1,2,14-17} However, specialized training and preparation of AI is required for generating discharge summaries specific to medical areas, notably psychiatry, which often requires a more individualized approach.⁴⁻¹³

Outpatient doctors have clear expectations for hospital discharge summaries, including speed, brevity (max. 2 pages), transparency, use of commonly known abbreviations, highlighting new treatment methods, and clear recommendations for

further care. The study shows that discharge summaries created by ChatGPT 4.0 meet these expectations. AI can optimize the content of discharge summaries, making them understandable for patients with varying medical knowledge levels. This reduces medical staff workload, allowing them to focus on direct patient care. Furthermore, AI can personalize discharge summaries, considering each patient's individual needs and medical history, enhancing their relevance and helpfulness. It also improves coherence and completeness, crucial for effective care. AI can automatically translate discharge summaries into various languages, facilitating communication with diverse patients. Clinical data analysis by AI can tailor the contents to specific diseases and treatments, enhancing informational value. However, there are challenges. AI-based systems might misinterpret or analyze medical data incorrectly, leading to errors. Regardless of AI quality, further control and correction by medical staff will be necessary. AI-generated content may be perceived as less empathetic, especially in psychiatric and psychotherapeutic departments.⁹ Discharge discussions, where the discharge summaries are addressed and patients can ask questions, remain essential. AI systems must implement security measures to protect sensitive medical data. Lastly, excessive automation could lead to job losses for medical staff, such as in writing offices.¹⁸⁻²⁰

Study limitations

The study's findings are subject to several limitations that may affect their generalizability and applicability to real-world clinical practice. Firstly, the study's small sample size, involving only six clinical staff members and four senior doctors for evaluation, limits the representativeness of the broader clinical community. Secondly, the study utilized fictional patient cases, which may not fully capture the complexity and variability of real-life clinical scenarios. Additionally, the AI model, ChatGPT 4.0, was only given a brief five-minute training session with two sample discharge summaries. This limited training might not have been sufficient to optimize its performance in generating discharge summaries, suggesting that a more extensive training period could potentially yield different results.

The study also faced blinding limitations. Although participants and evaluating physicians were blinded to the study's purpose, the use of only two cases might have led to recognition or familiarity, which could have influenced their impartiality.

Conclusions

In conclusion, while artificial intelligence can be a valuable tool in the creation of hospital documentation, particularly discharge summaries, it is still essential to consider potential challenges and limitations to ensure effective and safe health care delivery. It is also important to find a balance between the use of technology and the maintenance of traditional medical humanism and ethics in patient care.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Bertrand Janota  <https://orcid.org/0009-0003-3759-1510>

References

1. Spießl H. Der Kurzarztbrief: Zeitgewinn. *Dtsch Ärztebl.* 2004;101:2164-2568.
2. Erdogan-Griese B. Arztbrief: mehr als eine ungeliebte Pflicht. *Rheinisches Arzteblatt.* 2010; 12:23-24. doi:10.1055/s-0044-101096
3. Unnewehr M, Schaaf B and Friederichs H. Arztbrief: die Kommunikation optimieren. *Dtsch Ärztebl.* 2013;110:1672.
4. Kalbe P. KI-basierte Arztbrieferstellung – Entlastung für Ärztinnen und Ärzte durch generative Sprachmodelle. *Digitalisierung.* 2024.
5. Draese N. Arztbrief schreibt künftig die KI. *Ingenieur.de.* (Accessed May 2, 2024).
6. LohKI in der Medizin S.: Bald schreibt sie den Arztbrief. Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS. (Accessed May 2, 2024).
7. Rosenberg GS, Magnéli M, Barle N, et al. ChatGPT-4 generates orthopedic discharge documents faster than humans maintaining comparable quality: a pilot study of 6 cases. *Acta Orthop.* 2024;95:152-156. doi:10.2340/17453674.2024.95.152
8. Khalid N, Qayyum A, Bilal M, Al-Fuqaha A and Qadir J. Privacy-preserving artificial intelligence in healthcare: Techniques and applications. *Comput Biol Med.* 2023;158: 106848. doi:10.1016/j.compbiomed.2023.106848
9. Abbasian M, Khatibi E, Azimi I, et al. Foundation metrics for evaluating effectiveness of healthcare conversations powered by generative AI. *NPJ Digit Med.* 2024;7(1):82. doi:10.1038/s41746-024-00704-0
10. Patel SB and Lam K. ChatGPT: the future of discharge summaries? *Lancet Digit Health.* 2023;5:e107-e108. doi:10.1016/S2589-7500(23)00050-2
11. Powles J and Hodson H. Google DeepMind and healthcare in an age of algorithms. *Health Technol.* 2017;7:351-367. doi:10.1007/s12553-017-0185-0
12. Beyaert EW, Ubbink MM and Kaasenbrood AJ. Mediocre quality of discharge reports concerning psychiatric patients. *Ned Tijdschr Geneesk.* 1996;140:365-368.
13. Deloitte. Entlassmanagement im Krankenhaus. Available at: <https://www2.deloitte.com/de/de/pages/life-sciences-and-healthcare/articles/entlassmanagement-im-krankenhaus.html> (Accessed May 1, 2024).

14. Jatkowski A. So schreibt man einen optimalen Arztbrief. Available at: <https://www.arztbriefschreiben.de/so-schreibt-man-einen-optimalen-arztbrief/> (Accessed May 2, 2024)
15. Amering M, Denk E, Griengl H and Sibitz I. ["Because it concerns me"--patient-directed discharge letters]. *Psychiatr Prax*. 1998;25:172-174.
16. Bertrand D, Francois P, Bosson JL, Fauconnier J and Weil G. Quality assessment of discharge letters in a French university hospital. *Int J Health Care Qual Assur Inc Leadersh Health Serv*. 1998;11:90-95.
17. Bechmann S. Studie zu Arztbriefen. Heinrich-heine-universität. 2019. Available at: <https://www.hhu.de/die-hhu/presse-und-marketing/aktuelles/pressemeldungen-der-hhu/news-detailansicht/studie-zu-arztbriefen> (Accessed 23 January 2024).
18. Bures D, Hosters B, Reibel T, et al. Die transformative Wirkung von künstlicher Intelligenz im Krankenhaus: Der Mensch steht im Mittelpunkt. *Inn Med*. 2023;64:1025-1032. Published online 2023 Oct 18. PMID: PMC10602990, PMID: 37853060. doi:10.1007/s00108-023-01597-9
19. Küster D and Schultz T. Künstliche Intelligenz und Ethik im Gesundheitswesen – Spagat oder Symbiose? *Bundesgesundheitsblatt - Gesundheitsforsch - Gesundheitsschutz*. 2023;66: 176-183. Published online 2023 Jan 17. PMID: PMC9892090, PMID: 36650296. doi:10.1007/s00103-022-03653-5
20. Budde K, Dasch T, Kirchner E, et al. Künstliche Intelligenz: Patienten im Fokus. *Dtsch Ärztebl*. 2020;117:7. A-2407/B-2032.