Medical Teacher

ISSN: 0142-159X (Print) 1466-187X (Online) Journal homepage: www.tandfonline.com/journals/imte20

Re-imagining discharge summary training through artificial intelligence

Chun En Chua, Isaac K. S. Ng, Karina Yuen , Desmond B. Teo & Nathasha Luke

To cite this article: Chun En Chua, Isaac K. S. Ng, Karina Yuen , Desmond B. Teo & Nathasha Luke (27 Feb 2026): Re-imagining discharge summary training through artificial intelligence, Medical Teacher, DOI: 10.1080/0142159X.2026.2634066

To link to this article: https://doi.org/10.1080/0142159X.2026.2634066

View supplementary material

Published online: 27 Feb 2026.

Submit your article to this journal

Article views: 6

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=imte20

MEDICAL TEACHER
https://doi.org/10.1080/0142159X.2026.2634066

NEW WAVE

Re-imagining discharge summary training through artificial intelligence

Chun En Chua[a], Isaac K. S. Ng[b], Karina Yuen[c], Desmond B. Teo[d] and Nathasha Luke[e]

[a]Department of Medicine, Division of Advanced Internal Medicine, National University Hospital, Singapore; [b]Department of Medicine, Division of Rheumatology and Allergy, National University Hospital, Singapore; [c]Artificial Intelligence Centre for Educational Technologies (AICET), National University of Singapore, Singapore; [d]Integrated Programme, Alexandra Hospital, Singapore; [e]Department of Physiology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore

ABSTRACT

What is the educational challenge? Discharge summary (DS) writing is a core competency for junior physicians, yet persistent deficiencies in the quality, accuracy, and timeliness of these clinical documents are well-documented, with downstream repercussions in patient safety and continuity of care. Existing educational interventions rely heavily on faculty-intensive, small-group teaching models, which limits scalability and long-term sustainability. There is therefore a need to develop novel, more resource-efficient approaches to provide high-quality training in DS writing with individualised feedback.

What are the proposed solutions? We propose a new educational model that integrates artificial intelligence (AI)-generated feedback into a structured DS training programme. As a proof-of-concept, we conducted a small-scale evaluation comparing feedback quality from multiple AI platforms and a human trainer using a standardised rubric. Based on these findings, we designed an asynchronous Coursemology-based e-learning module incorporating customised generative-AI (cGen-AI) to generate draft feedback, with human moderation retained as a safety and quality assurance step. This model is currently in the pre-implementation phase.

What are the potential benefits to a wider global audience? This conceptual human-in-the-loop AI model has the potential to deliver scalable, consistent, and individualised feedback while substantially reducing faculty and logistical workload. By enabling asynchronous practice and standardised assessment, it directly addresses sustainability challenges faced by DS training programmes internationally.

What are the next steps? Full implementation and evaluation including reliability, learner acceptance, and educational impact of this model is being planned for an entire medical student cohort to replace the existing small-group, faculty-facilitated sessions. The success of such cGen-AI approach for DS training can also be extended to other similar domains of medical training in the future.

ARTICLE HISTORY
Received 16 January 2026
Accepted 16 February 2026

KEYWORDS
Discharge summary; artificial intelligence; medical education; feedback; assessment

What is the educational challenge?

Discharge summary (DS) writing is a core competency for junior physicians, which is essential for high-quality patient care. Over the past decade, multiple clinical audits have consistently revealed significant deficiencies in the quality, accuracy, and timeliness of DS [1]. These deficiencies are associated with adverse downstream effects, such as poor continuity of care, medical errors, suboptimal adherence to treatments, and follow-up, with compromised patient safety.

To address this problem, we had previously implemented a structured DS training programme for final-year medical students at the National University of Singapore (NUS) [2]. A key component of this programme was interactive, case-based learning with real-time feedback from trained faculty. Through deliberate practice across diverse clinical scenarios, learners developed skills in synthesising problem lists and producing patient-centred DS within standardised frameworks.

As part of this educational quality improvement (QI) initiative, we successfully implemented multiple pedagogical, assessment, and training strategies that resulted in significant and sustained improvements in the quality of DS written by junior doctors during regular clinical audits. In the existing training model, five groups of five students are facilitated by three trainers per session. Students complete preparatory readings and discuss simulated case scenarios during a 1.5-hour facilitated session. Each group enters its responses onto an online answer-sharing platform (Padlet), after which the issue lists are reviewed in a flipped-classroom format with group discussion and trainer feedback. These

CONTACT Nathasha Luke nathasha@nus.edu.sg 21 Lower Kent Ridge Road, 119077, Singapore
Supplemental data for this article can be accessed online at https://doi.org/10.1080/0142159X.2026.2634066.
© 2026 Informa UK Limited, trading as Taylor & Francis Group

sessions are repeated every three months to accommodate new students rotating through the Department of Medicine at the National University Hospital.

While educationally effective, this approach is time and labour-intensive. Faculty must block clinical time to facilitate sessions and programme continuity is vulnerable to staff turnover. New faculty requires training and sustained commitment without guaranteed protected time. Sustaining these educational interventions over time remains challenging, as improvements often regress without continuous reinforcement [3]. These constraints limit both scalability and long-term sustainability.

What are the proposed solutions?

To overcome these challenges, we propose converting existing case scenarios into an interactive e-learning module hosted on the NUS Coursemology platform, which is a learning management system developed in-house allowing the flexibility to build the human-in-the-loop feature and customise the settings to fit trainers' needs. The module retains the open-ended components that require students to formulate problem lists in DS. Student responses are analysed by an artificial intelligence (AI) system integrated into Coursemology, which generates structured feedback aligned with predefined rubrics. The generated output is then reviewed and moderated by faculty members to safeguard against inaccuracies or hallucinations within AI-generated feedback before being released to the students.

This blended approach allows learners to practise DS writing asynchronously at their own time, while retaining standardised assessments and overcoming manpower and logistical constraints. Faculty involvement in this model shifts from live facilitation to targeted moderation of AI-generated feedback, enabling a significantly lower trainer-to-student ratio and greater flexibility in scheduling.

To inform platform selection, we compared five feedback providers: ChatGPT, Gemini, Copilot, a customised generative-AI (cGen-AI) system, and a human trainer. Feedback quality was assessed using a 1–5 scale across seven domains—accuracy, completeness, detail, organisation, specificity, clarity, and relevance—yielding a maximum score of 35. Two blinded human raters independently evaluated all feedback, demonstrating excellent inter-rater reliability (intraclass correlation coefficient (ICC) 0.90).

Fifty students' responses were evaluated. Among the AI platforms, ChatGPT produced the highest-rated feedback (85%) over the seven domains specified, followed by cGen-AI (63%). Human trainer feedback scored lowest (23%). Human trainer was a medical consultant with more than 15 years of clinical practice and heavily involved in medical education. Human trainer was allowed to grade at his own time and target. While human feedback accurately identified key errors, it was often less comprehensive across multiple feedback domains. Despite using identical rubrics, scoring consistency varied across AI platforms, with lower inter-rater reliability for grading student answers (ICC 0.425–0.675). The data was stored on a secure platform, and no patient or participant-identifiable information was included, as all case scenarios used in the study were simulated.

Based on these findings, we integrated a cGen-AI system into the Coursemology platform and conducted small-scale test runs with selected students. Qualitative feedback from participants was positive, particularly regarding the clarity, specificity, and perceived usefulness of the feedback. The learning model was widely accepted in terms of usability by the trainers and learners.

To our knowledge, this is the first study to integrate cGen-AI into Coursemology for DS training with a focus on patient safety and real-world clinical practice. The customised ChatGPT grading system provides personalised feedback, with the added advantage that human trainers can review and refine grading before release, unlike most AI systems that provide direct, static feedback.

What are the potential benefits to a wider global audience?

The use of cGen-AI to generate structured feedback offers substantial efficiency gains by reducing faculty workload while delivering individualised feedback for learners. Rather than providing generic group feedback, faculty can also focus on higher-order educational tasks, such as reviewing and refining AI-generated responses. Prior studies have demonstrated that AI systems can apply standardised assessment rubrics consistently, a task that can be challenging for human educators, particularly at scale [4].

Our findings highlight the potential of AI-enabled systems to deliver effective, individualised feedback on diagnostic and clinical reasoning skills across diverse medical scenarios. This asynchronous learning model allows students to engage flexibly while receiving timely, personalised guidance that may be difficult to sustain in traditional small-group teaching environment. Maintaining a human-in-the-loop approach ensures that feedback remains accurate, contextualised, and aligned with learning objectives,

representing an optimal blended model during this transitional phase of AI adoption.

Internationally, many educational initiatives targeting DS quality face similar sustainability challenges, including frequent turnover of junior physicians and reliance on scheduled, resource-intensive teaching sessions [5]. These programmes are often implemented on a small scale to minimise resistance, potentially limiting their educational impact. Our proposed model removes the need for fixed teaching sessions and allows larger cohorts of learners to engage in DS practice with individualised, moderated AI feedback.

DS teaching programmes utilising 'trained' AI models with pre-defined script inputs and assessment rubrics can be easily transferable to other medical educational institutions, with ease of cross-institutional/international collaboration, and sharing of best practices (e.g. grading metrics, feedback frameworks) in DS training. The success of such an AI-embedded model of teaching and assessment of DS writing can also be readily extended to other pedagogical areas in undergraduate and postgraduate medical training.

What are the next steps?

Among the platforms evaluated, ChatGPT consistently delivered the highest-quality feedback, offering more specific and actionable suggestions than human trainers. While cGen-AI systems powered by other platforms scored lower, they offered operational advantages by embedding rubrics and prompts directly into the workflow, allowing students to receive feedback simply by uploading their work. Importantly, iterative training of these systems led to progressive improvements in feedback quality.

Building on these findings, we are developing a cGen-AI platform powered by ChatGPT-4.1mini and fully integrated into the Coursemology e-learning environment. Students will be able to directly submit DS on the platform and receive AI-curated, human-moderated feedback.

The next phase of this educational model implementation will involve an entire cohort of Year 5 NUS medical students and house officers. This will enable comprehensive quantitative analysis of AI-generated feedback, trainer moderation patterns, inter-rater reliability, as well as qualitative evaluation of student acceptance and perceived educational value. We are also intending to continue longitudinal audits of DSs written at our institution after implementation of the training program to assess organisational performance—which is the highest tier of Kirkpatrick model. Ultimately, our goal is to refine the AI system to the point where minimal human moderation is required, enabling sustainable, high-quality feedback at scale. Beyond this, we are working with other like-minded clinical educators to study how such AI-embedded educational models can be implemented in other important domains of medical education that require structured writing, clinical reasoning, and formative feedback.

Author contributions

CRediT: **Karina Yuen**: Resources, Software.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Ethical approval

Not applicable.

Funding

No funding required.

Notes on contributors

Chun En Chua, Division of Advanced Internal Medicine, Department of Medicine, National University Hospital, Singapore.

Isaac K. S. Ng, Division of Rheumatology and Allergy, Department of Medicine, National University Hospital, Singapore.

Karina Yuen, Artificial Intelligence Centre for Educational Technologies (AICET), National University of Singapore, Singapore.

Desmond B. Teo, Integrated Programme, Alexandra Hospital, Singapore.

Nathasha Luke, Department of Physiology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore.

ORCID

Chun En Chua http://orcid.org/0000-0002-0338-5589
Nathasha Luke http://orcid.org/0000-0002-0911-216X

References

1. Chua CE, Teo DB. Writing a high-quality discharge summary through structured training and assessment. Med Educ. 2023;57(8):773–774. doi: 10.1111/medu.15102
2. Ng IKS, Choe R, Ong SS, et al. Postgraduate pedagogical training and interventions to promote high-quality discharge summaries. Clin Teach. 2025;22(5):e70150. PMID:40696270. doi: 10.1111/tct.70150

3. Kumah A, Nutakor HS, Issah AR, et al. Achieving sustainability of quality improvement projects. Glob J Qual Saf Healthc. 2025;8(1):50–52. doi: 10.36401/JQSH-23-48
4. Zuckerman M, Flood R, Tan RJB, et al. ChatGPT for assessment writing. Med Teach. 2023;45(11):1224–1227. doi: 10.1080/0142159X.2023.2249239
5. Wong AM, Gopal P. Improving the secondary diagnoses capture rate in SingHealth Community Hospital discharge summaries: a quality improvement project made successful by change management principles. BMJ Open Qual. 2024;13(3):e002747. doi: 10.1136/bmjoq-2024-002747