

Research Letter

AI-Based Patient Safety Incident Reporting System for Nurses (SAFE-Nurse): Development and Validation Study

Nindya Febriyani Putri¹, BSN, RN, MN; Rini Rachmawaty^{1,2}, BSN, RN, MN, PhD; Nuurhidayat Jafar³, BSN, RN, MN, PhD

¹Nursing Management & Healthcare Quality Research Group, Faculty of Nursing, Hasanuddin University, Makassar, Indonesia

²Department of Nursing Management, Faculty of Nursing, Hasanuddin University, Makassar, Indonesia

³Department of Community Nursing, Faculty of Nursing, Hasanuddin University, Makassar, Indonesia

Corresponding Author:

Rini Rachmawaty, BSN, RN, MN, PhD
Nursing Management & Healthcare Quality Research Group
Faculty of Nursing, Hasanuddin University
Jl. Perintis Kemerdekaan KM. 10
Makassar 90245
Indonesia
Phone: 62 81213916730
Email: rini.rachmawaty@unhas.ac.id

Abstract

This research letter describes the development and initial validation of SAFE-Nurse, an artificial intelligence–assisted patient safety incident reporting system for nurses, which demonstrated excellent content and face validity (average scale-level content validity index [S-CVI/Ave] 0.99; average scale-level face validity index [S-FVI/Ave] 1.00) and is ready to progress to usability testing.

JMIR Form Res 2026;10:e91321; doi: [10.2196/91321](https://doi.org/10.2196/91321)

Keywords: artificial intelligence; digital health; incident reporting; nursing practice; patient safety; validation studies

Introduction

Patient safety incident reporting is essential for health care quality, yet underreporting remains common due to workload, limited feedback, and blame culture among nurses [1,2]. Paper-based reporting further limits reporting efficiency and data management [3]. Although AI has potential to support healthcare processes [4], its application in patient safety incident reporting remains limited.

To address this gap, we developed Smart Assisted Feedback and Education for Nurses on Reporting Safety Events (SAFE-Nurse), an AI-assisted incident reporting system. This study evaluated its content and face validity.

Methods

Study Design

This research and development study used the ADDIE framework to develop SAFE-Nurse. The process comprised needs analysis (Analyze), system design (Design), prototype

development and content and face validity assessment (Develop), pilot implementation (Implement), and usability evaluation using the System Usability Scale (Evaluate). This report focuses on the Develop phase.

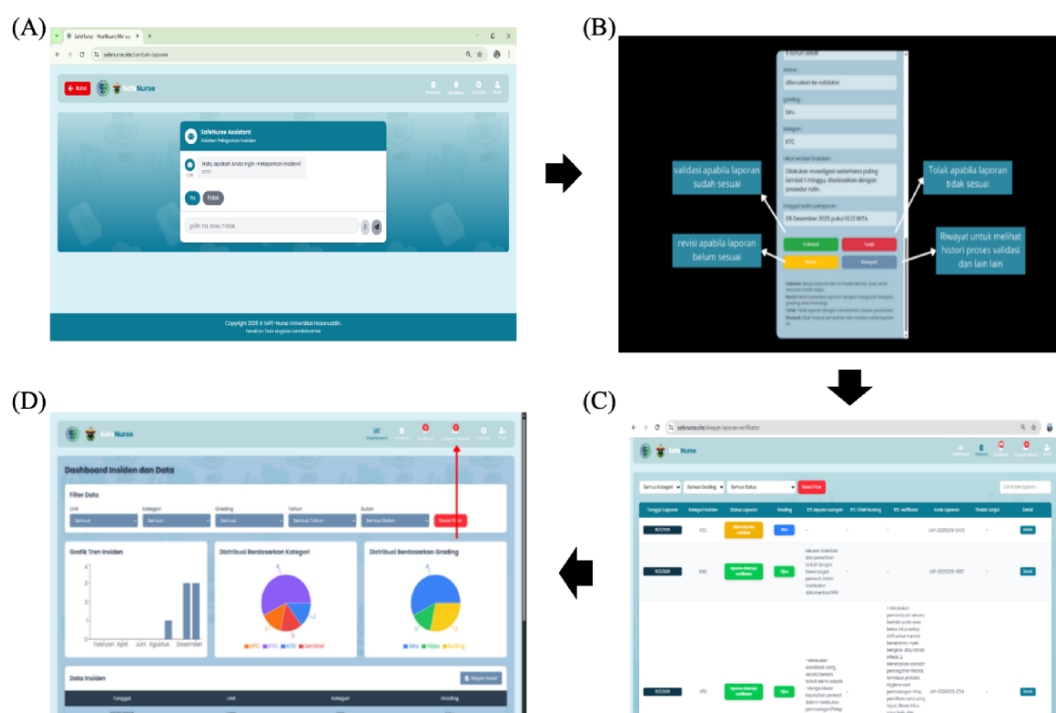
System Description (SAFE-Nurse)

SAFE-Nurse is an AI-assisted web application for patient safety incident reporting. It uses OpenAI Whisper (whisper-1) for voice-to-text transcription and GPT-4o-mini via an application programming interface (API) to normalize incident narratives, recommend incident categories, and assess chronology completeness. Severity grading is calculated deterministically by the backend using predefined impact and probability scores. AI outputs are reviewed and confirmed by nurses and may be revised by authorized reviewers. Detailed information on the AI pipeline, API inputs and outputs, prompt structure, and severity-grading process is provided in [Multimedia Appendix 1](#).

The system supports 3 roles: staff nurses, head nurses, and quality and patient safety personnel. Staff nurses submit reports and receive AI-assisted guidance. Head nurses review,

validate, revise, or reject reports and review AI-generated classifications. Quality and patient safety personnel perform final verification and monitor incidents through the dashboard (Figure 1).

Figure 1. SAFE-Nurse interfaces: (A) AI-assisted incident reporting interface for staff nurses; (B) validation interface for head nurses; (C) incident reporting dashboard; (D) analytical dashboard for quality and patient safety monitoring.



Study Setting and Participants

Content validity was assessed by 6 experts selected purposively through professional and institutional networks in Makassar and Yogyakarta, Indonesia. Eligible experts had at least 5 years of hospital experience and expertise in nursing informatics, nursing management, or patient safety and quality, with 2 experts from each field. Assessments were conducted at their respective institutions or remotely via Zoom. Experts were invited via email or WhatsApp.

Face validity was assessed among 17 potential SAFE-Nurse users at RSIA Paramount Makassar, Indonesia, selected purposively using maximum variation sampling. Participants included staff nurses (n=7), head nurses (n=7), and quality and patient safety personnel (n=3), recruited through hospital contacts based on their patient safety reporting roles. The content- and face-validity groups were separate and nonoverlapping.

Instrument

A researcher-developed 43-item questionnaire based on SAFE-Nurse system features and study objectives assessed content and face validity (Multimedia Appendix 2). It comprised 5 domains: content relevance and objectives (7 items), clarity and usability (10 items), functionality and workflow (7 items), technical aspects and confidentiality (9 items), and usefulness and impact (10 items). Content validity used a 4-point Likert scale; face validity used a dichotomous “clear”/“unclear” scale with optional comments. All 43

items were retained. The questionnaire was not pretested or pilot-tested.

Validity Procedure (Content and Face Validity)

Participants received study information and access credentials to explore SAFE-Nurse before completing the questionnaire during scheduled in-person sessions. Experts assessed system relevance, whereas end users assessed clarity after interacting with SAFE-Nurse. Participants could also provide comments or suggestions in the open-ended section using the prompt: “Please provide comments or suggestions regarding aspects of the SAFE-Nurse system that require improvement.”

Data Analysis

Data were analyzed descriptively. Content and face validity were assessed using the content validity index (CVI) and face validity index (FVI), respectively. Open-ended comments were summarized by refinement area.

Ethical Considerations

The study was approved by the Hasanuddin University Research Ethics Committee (No. 659/UN4.18.3./TP.01.02/2025). Written informed consent was obtained, and data were anonymized and kept confidential.

Results

Content and Face Validity

Table 1 summarizes the results. Among 6 experts, item-level CVI ranged from 0.83 to 1.00; scale-level average and

universal agreement CVI were 0.99 and 0.93, respectively. Among 17 end users, item-level FVI ranged from 0.94 to 1.00; scale-level average and universal agreement FVI were 1.00 and 0.98, respectively. Overall, SAFE-Nurse showed strong agreement regarding relevance and clarity.

Table 1. Content validity index (CVI) and face validity index (FVI) of the SAFE-Nurse system.

Domain	Number of items	I-CVI ^a	I-FVI ^b
Content relevance and objectives	7	1.00	1.00
Clarity and usability	10	1.00	1.00
Functionality and workflow	7	1.00	1.00
Technical and confidentiality	9	1.00	1.00
Usefulness and impact	10	0.83-1.00	0.94-1.00
Scale average (S-CVI ^c and S-FVI ^d)	— ^e	0.99	1.00
Universal agreement (S-CVI and S-FVI)	—	0.93	0.98

^aI-CVI: item-level content validity index.

^bI-FVI: item-level face validity index.

^cS-CVI: scale-level content validity index.

^dS-FVI: scale-level face validity index.

^eNot applicable.

Expert Feedback and System Refinement

Expert comments identified areas for refinement in workflow, interface design, reporting, and monitoring. Refinements included reorganizing incident categories and severity grading, simplifying data entry, adding progress indicators, role-specific guidance, notification reminders, and follow-up fields. Integration with the hospital information system (SIMRS) was also recommended.

Discussion

Principal Findings

SAFE-Nurse demonstrated excellent content and face validity, indicating strong consensus among experts and end users regarding its relevance and clarity for patient safety incident reporting [5-7]. These findings support system refinement and progression to usability testing but do not establish reliability, real-world usability, or clinical effectiveness.

The findings are consistent with research demonstrating the potential of AI for patient safety event classification and human-AI collaboration [8]. Generative AI has also

been explored for critical incident analysis, although further evaluation is needed before broader application [9]. Similar nurse-led AI work has used expert review and iterative refinement to improve the structure and completeness of patient safety incident reports [10]. Consistent with these approaches, AI-generated classifications and chronology feedback in SAFE-Nurse are intended to support, rather than replace, human review.

This study has three limitations. First, the expert and end-user groups were relatively small. Second, the 43-item questionnaire was not pretested or pilot-tested, so potential issues related to comprehensibility or respondent burden may not have been identified. Third, clinical performance was not assessed.

Conclusion

SAFE-Nurse demonstrated excellent content and face validity as an AI-assisted incident reporting system. These findings indicate strong consensus regarding the relevance and clarity of SAFE-Nurse and support progression to usability testing; they indicate that SAFE-Nurse is relevant and appropriate for Indonesian hospitals.

Acknowledgments

The authors thank the nursing informatics, nursing management, and patient safety and quality experts who participated in the expert judgment process for their valuable feedback and recommendations during the development and content validation of the SAFE-Nurse system. The authors also acknowledge the Faculty of Nursing, Hasanuddin University, for its academic support and research facilities throughout the study.

Generative AI Use

The authors used generative artificial intelligence (ChatGPT) to assist with language editing, grammar improvement, and manuscript clarity. All AI-assisted outputs were critically reviewed, revised, and approved by the authors, who take full

responsibility for the final content. The figures included in this manuscript are direct screenshots of the functional SAFE-Nurse system and were not generated or modified using AI.

Funding

This research was funded by the Indonesian Ministry of Higher Education, Science and Technology through the Institute for Research and Community Service of Hasanuddin University under the Master's Thesis Research Grant Scheme 2025 with Contract No. 069/C3/DT.05.00/PL/2025.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Technical description of the SAFE-Nurse AI pipeline.

[\[DOCX File \(Microsoft Word File\), 19 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

Content and face validity evaluation instrument for the SAFE-Nurse system.

[\[DOCX File \(Microsoft Word File\), 17 KB-Multimedia Appendix 2\]](#)

References

1. Krouss M, Alshaikh J, Croft L, Morgan DJ. Improving incident reporting among physician trainees. *J Patient Saf.* Dec 2019;15(4):308-310. [doi: [10.1097/PTS.0000000000000325](#)] [Medline: [27617963](#)]
2. Benevento M, Nicolì S, Mandarelli G, et al. Strengths and weaknesses of the incident reporting system: an Italian experience. *Journal of Patient Safety and Risk Management.* Feb 2023;28(1):15-20. [doi: [10.1177/25160435221150568](#)]
3. Adnyani N, Nugraha INA, Rudiarta I. Evaluasi Pelaporan Insiden Keselamatan Pasien di Rumah Sakit Umum Daerah Karangasem [Article in Indonesian]. *jikep.* 2022;8(4):537-549. [doi: [10.33023/jikep.v8i4.1217](#)]
4. Soularito DS, Arini M, Listiowati E. Digital System Pelaporan Insiden Keselamatan Pasien Di Rumah Sakit [Article in Indonesian]. *Pros Semin Nas Progr Pengabd Masy.* 2022. [doi: [10.18196/ppm.43.635](#)]
5. Yusoff MSB. ABC of content validation and content validity index calculation. *EIMJ.* 2019;11(2):49-54. [doi: [10.21315/eimj2019.11.2.6](#)]
6. Polit DF, Beck CT. The content validity index: are you sure you know what's being reported? Critique and recommendations. *Res Nurs Health.* Oct 2006;29(5):489-497. [doi: [10.1002/nur.20147](#)] [Medline: [16977646](#)]
7. Fajarini M, Rahayu S. Knowledge transfer through mobile application needs questionnaire: validity and reliability. *INJEC.* 2020;5(2):134. [doi: [10.24990/injec.v5i2.296](#)]
8. Chen H, Cohen E, Wilson D, Alfred M. A machine learning approach with human-AI collaboration for automated classification of patient safety event reports: algorithm development and validation study. *JMIR Hum Factors.* Jan 25, 2024;11:e53378. [doi: [10.2196/53378](#)] [Medline: [38271086](#)]
9. Hölzing CR, Rumpf S, Huber S, Papenfuß N, Meybohm P, Happel O. The potential of using generative AI/NLP to identify and analyse critical incidents in a Critical Incident Reporting Systems (CIRS): a feasibility case-control study. *Healthcare (Basel).* Oct 2, 2024;12(19):1964. [doi: [10.3390/healthcare12191964](#)] [Medline: [39408144](#)]
10. Teo KY, Huang L, Woh KCY, Ang SY, Ng JGN. Nurse-led ambient AI scribe for patient safety incident investigation reports (project NARRATE): retrospective pre-post comparative document-quality study. *JMIR Nurs.* Aug 10, 2026;9:e100775. [doi: [10.2196/100775](#)] [Medline: [42574744](#)]

Abbreviations

ADDIE: Analyze, Design, Develop, Implement, and Evaluate

API: application programming interface

CVI: content validity index

FVI: face validity index

SAFE-Nurse: Smart Assisted Feedback and Education for Nurses on Reporting Safety Events

Edited by Stephanie Law; peer-reviewed by Chantelle Recsky, Chenyu Li; submitted 13.Jan.2026; final revised version received 24.Sep.2026; accepted 24.Sep.2026; published 23.Oct.2026

Please cite as:

Febriyani Putri N, Rachmawaty R, Jafar N

AI-Based Patient Safety Incident Reporting System for Nurses (SAFE-Nurse): Development and Validation Study

JMIR Form Res 2026;10:e91321

URL: <https://formative.jmir.org/2026/1/e91321>

doi: [10.2196/91321](https://doi.org/10.2196/91321)

© Nindya Febriyani Putri, Rini Rachmawaty, Nuurhidayat Jafar. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 23.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.