Human Factors in Design, Engineering, and Computing, Vol. 159, 2024, 2279–2288
https://doi.org/10.54941/ahfe1005796

Development of a Support System to Improve the Ability to Analyze Incident Reports in Hospitals Using Generative AI: Enhancement of the Awareness of Frontline Risk Managers

Yuriko Imura and Yusaku Okada

Graduate School of Science and Technology, Keio University, Japan

ABSTRACT

A major issue in the analysis of in-hospital incidents is the lack of skill and experience of the analyst. When an analyst with little experience analyzes an incident, they tend to only extract superficial characteristics and list only ineffective measures such as minor work improvements and thorough confirmation. To solve this issue, it is essential to educate the analyst. However, the higher the safety of the organization, the less experience employees have with accidents and serious incidents. Compensating for this experience with only regular education and virtual experiences is too laborious and takes a lot of time for the analyst to grow. Therefore, in this study, we examined strategies to enhance the awareness of analysts through the routine analysis of incident reports in medical settings. Specifically, this involves the development of the Medical Risk-Managers' Awareness Enhancement System in Medical Incident Analysis.

Keywords: Medical incident analysis, Healthcare safety, Medical risk manager

INTRODUCTION

In medical facilities and general hospitals, various information on incidents in patient care is collected and analyzed with the aim of preventing accidents. However, in small and medium-sized hospitals, effective measures are not fully implemented because many of the medical risk managers responsible for analyzing incident reports lack analytical skills and experience. This is because incident analysis by inexperienced analysts tends to extract only superficial features and only lists ineffective measures such as minor operational improvements and thorough checks. In fact, a look at the factors reported to the Japan Institute for Health Care Excellence (85,204 incidents in total, 2023) shows that about half of the factors are "Insufficient confirmation or carelessness of the person concerned", indicating that only superficial factors are extracted. Although there are methods to improve the awareness of analysts through conventional medical safety education to solve such problems, it is very difficult to increase the time and cost of safety education for medical risk managers who are responsible for on-site analysis

for this purpose. It is impossible for small and medium-sized hospitals to implement safety education in their own departments. Therefore, this study proposes measures to improve the awareness of analysts through the addition of an AI-based advice function to the routine incident report analysis work performed by medical risk managers.

CONCEPT

The purpose of this study is to develop strategies that contribute to the improvement of awareness capabilities through incident report analysis. However, in order to devise concrete methods, it is essential to clarify what awareness capabilities are required in the context of medical safety. Therefore, we investigated the objectives and outcomes of past medical safety education programs and conducted interviews with top management at various hospitals regarding the characteristics of personnel considered to have high awareness capabilities. By organizing these results, we identified the following five characteristics related to awareness capabilities:

- Information gathering and sharing
- Addressing feelings of anxiety and discomfort
- Adherence to fundamentals
- Collaboration with others
- Building a sense of purpose

In this study, we will use these five characteristics as indicators of awareness capabilities and measure their fulfilment through incident report analysis tasks.

PROPOSAL OF METHOD

The objective of this study is to enhance the awareness capabilities of analysts by integrating an AI-based advisory function into the incident report analysis process. To achieve this, we first developed an incident report analysis support system and then examined how to incorporate an awareness improvement function into it.

In hospitals, incident report analysis typically involves collecting information related to the incident, noticing contributing factors, and developing improvement measures based on these factors. For analysts to effectively implement improvement activities on-site, it is important that they notice as many factors as possible during the factor extraction stage. Additionally, it is desirable that the incident reporting process itself contributes to the improvement of awareness capabilities by helping analysts notice potential contributing factors. Therefore, in the factor extraction process, it is necessary to design the system in a way that allows analysts to grasp any deficiencies or excesses in their current factor extraction and to foster the improvement of their own awareness capabilities.

Based on these requirements, we explored a method where the content of the extracted factors is evaluated based on two criteria: (1) comprehensiveness of factors and (2) the five characteristic values. Based on

the evaluation results, the system would then provide advice to guide analysts on what additional content should be included (Shown as Figure 1).

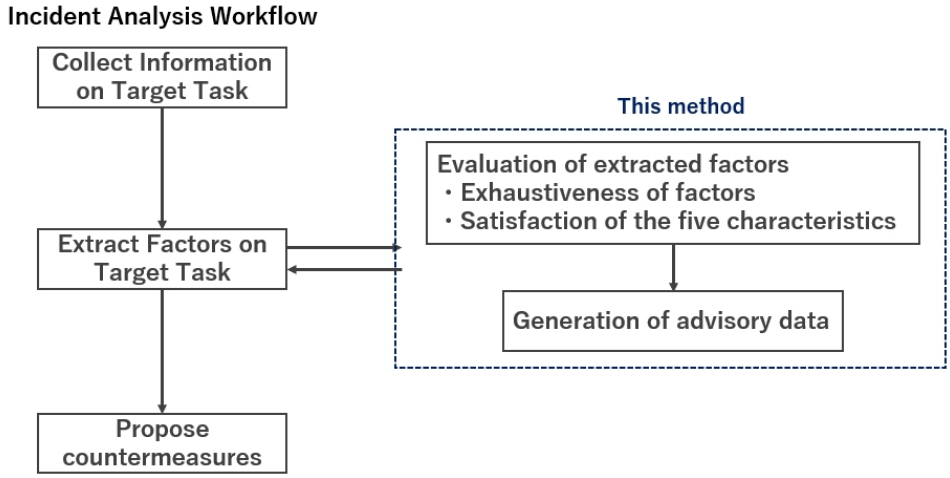

Figure 1: Overview of the proposed system.

EVALUATION SCALE AND SCALE ITEMS FOR SKILLS REQUIRED OF HEALTHCARE PROFESSIONALS

The past research on the evaluation scales of skills required for healthcare professionals was investigated, resulting in the collection of 82 scale items. The collected past studies were limited to empirical research, such as observation, case note analysis, interviews, and surveys. After removing duplicate content, items that could be incorporated into factor analysis were selected, and a feature classification was conducted. As a result, 12 categories were obtained (Table 1).

Table 1. Rating scales/scale items for skills required of healthcare professionals.

Category	Score	Elements
Communication	y_1	.Information exchange .Building mutual understanding .Developing proactive assertions
Decision making	y_2	.Recognizing options .Generating alternatives .Risk assessment
Teamwork	y_3	.Collaboration with others .Task coordination .Consideration for others

(Continued)

Table 1. Continued

Category	Score	Elements
Workload management	y_4	.Workload management .Prioritization .Delegation of tasks .Coping with pressure
Situation awareness management	y_5	.Information gathering .Recognition and interpretation of information .Time awareness .Patient observation .Monitoring progress
Organizational support behavior	y_6	.Building and maintaining a team .Conflict resolution
Professional Consideration	y_7	.Supporting team members
Problem setting ability	y_8	.Recognizing and defining problems .Risk evaluation
Instructional ability	y_9	.Assessing capabilities
Leadership	y_{10}	.Exercising authority and assertiveness .Coordinating team actions
Task management	y_{11}	.Planning and preparing tasks .Setting performance standards .Maintaining and providing standards .Identifying and utilizing resources
Producing	y_{12}	.Understanding team needs .Predicting future situations

ORGANIZATION OF BACKGROUND FACTORS CORRESPONDING TO THE EVALUATION SCALE

The prior research on incident analysis was reviewed, and 265 background factors were collected. These factors were organized and categorized into five categories based on the elements shown in Table 1. For example, factors such as "no set procedures" and "uncertain placement" were categorized under "Operations Management" as they relate to "Maintaining and Providing Standards." Similarly, the factor "lack of information sharing" was categorized under "Communication" as it relates to "Information Exchange."

EVALUATION OF THE SATISFACTION OF FIVE CHARACTERISTICS USING THE EVALUATION SCALE

To clarify the relationship between the five characteristics and the twelve evaluation criteria, a broad review of past research referring to each characteristic was conducted. As a result, the relationships between the twelve evaluation criteria and characteristics, as shown below in Figure 2, were identified.

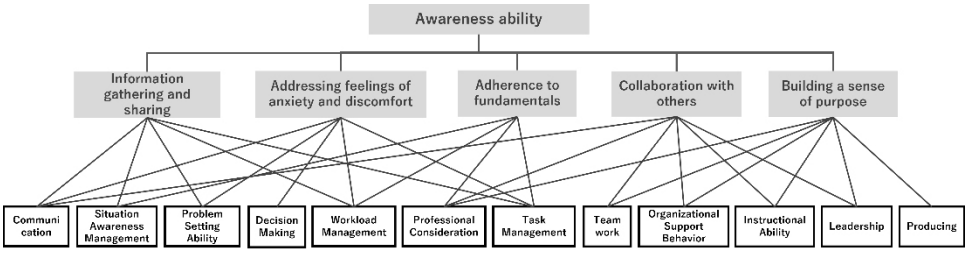

Figure 2: Relationship between 5 characteristics and 12 evaluation criteria.

Using the results obtained above, the fulfillment degree of each characteristic $z_i (i = 1, ..., 5)$ is defined as follows, with the impact of each characteristic on each criterion represented by β_{ij}:

$$z_i = \sum_{j=1}^{12} \beta_{ij} y_j$$

β_{ij} : Coefficient representing the impact between each characteristic and each criterion $(i = 1, ..., 5)(j = 1, ..., 12)$
z_i : Fulfillment degree of each characteristic $(i = 1, ..., 5)$
y_j : Number of factors belonging to each category $(j = 1, ..., 12)$

FORMULATION OF SATISFACTION LEVELS AND ADVISORY GUIDELINES

To provide advice on addressing deficiencies based on assessed satisfaction levels, it is not practical to list all the missing content. Therefore, it is necessary to provide general direction based on the situation. To do this, satisfaction levels are categorized into stages, and by organizing issues and considerations according to these satisfaction levels, we can outline guidelines for generating advice. The satisfaction levels are defined as $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5 (\beta_1 < \beta_2 < \beta_3 < \beta_4 < \beta_5)$.

Table 2. Guidelines for advice based on satisfaction levels.

Level	Fullfillment z_i	Situations and Problems
Level0	$0 \leqq z_i \leqq \beta_1$	Simply reinforcing awareness, but this is something the person concerned is fully aware of on reflection. Being told this over and over again by several people only increases the mental burden on the individual and, conversely, increases the likelihood of error. If the work was "no confirmation action at all," it makes sense, but if not, it is ineffective without consideration of the current state of the work, including the confirmation action.

(Continued)

Table 2. Continued

Level	Fullfillment z_i	Situations and Problems
Level1	$\beta_1 < z_i \leqq \beta_2$	It makes sense if there is no education about the work, otherwise it will only make the instructor stricter and totally ineffective.
Level2	$\beta_2 < z_i \leqq \beta_3$	Lack of clear guidelines and goals for the measures is likely to make them a formality.
Level3	$\beta_3 < z_i \leqq \beta_4$	Care should be taken to ensure that the rules do not become a skeleton. It is desirable to link this to a survey of employee awareness of the measures.
Level4	$\beta_4 < z_i \leqq \beta_5$	The measures themselves are specific, but they are specific to one part: the bias of extracted causes needs to be eliminated using existing factorial classification methods such as m-shell.4M.
Level5	$\beta_5 < z_i$	In order to obtain understanding and acceptance of safety measures, it is desirable to coordinate not only implementation and instruction of measures, but also with education on safety activities in general.

DEVELOPMENT OF THE SUPPORT SYSTEM

Using the defined evaluation methods, "Medical Risk-Mangers' Awareness Enhancement System in Medical Incident Analysis" was developed to derive the final output from the information items on the incident report (Figure 3). Initially, to support factor extraction in incident analysis, the system was equipped with functions to create a logic tree based on factors inputted in RCA (Root Causes Analysis) format and to classify factors using the SHEL (Software, Hardware, Environment, Liveware) model. Subsequently, to enhance insight capabilities, tools for satisfaction evaluation and advice provision were added.

The functionalities provided to analysts are as follows:

① Incident cause description tool using RCA (Root Causes Analysis)

The analyst inputs the factors in sequence. After that, they add the relationship with the previously input factor. Based on these two pieces of information, the factors are automatically described in a logic tree state, and the relationship between the factors is visualized. In addition, the dependency between the factors is also shown. This dependency is automatically described with reference values based on past cases, which the analyst checks and corrects.

② SHEL Classification tool for incident causes

The input factors are automatically classified into SHEL (Software, Hardware, Environment, Liveware) through morphological analysis. The analyst checks and corrects this.

③ Assessment tool of the Fulfillment of Five Characteristics

The input factors are classified into 5characteristics, and the content is automatically judged to be sufficient to address the problems of each skill (fulfillment). The fulfillment is evaluated on a five-point scale from level 1 to 5.

④ Advice Presentation tool for Improving Fulfillment

Based on the calculated fulfillment and the contribution to the improvement of non-technical characteristics, "characteristics of the factor group that should be added", "identification of factors that should be detailed", "pointing out redundant factors", and "pointing out factors whose content is unclear" are performed. When the fulfillment is level 2 or lower, the items that should be implemented are forcibly made clear.

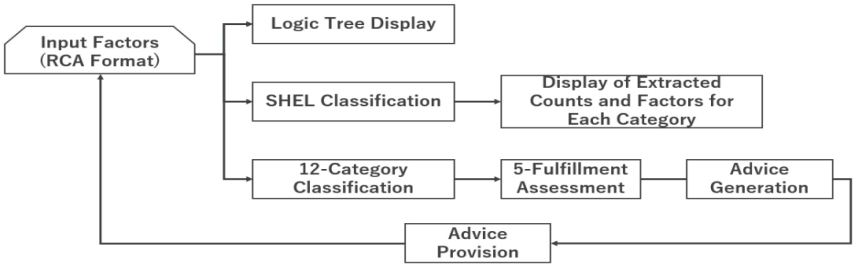

Figure 3: Configuration of medical risk-mangers' awareness enhancement system in medical incident analysis.

VERIFICATION EXPERIMENT

Four medical risk managers were grouped into one team, and six such groups performed factor analysis on the provided incidents. The number of factors extracted before and after using the support system is shown in Figure 4. The satisfaction results are presented in Table 3, and an example of satisfaction results for Group B is shown in Figure 5.

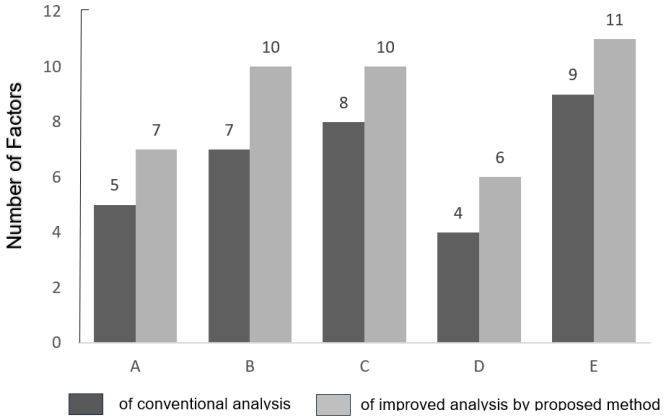

Figure 4: Comparison the result of conventional analysis and of improved analysis by proposed method.

Based on Figure 4, Table 3, and Figure 5, the following conclusions can be drawn before and after support:

Table 3. Fulfillment before and after support.

GROUP	A		B		C		D		E	
METHOD	Conventional	Improved	Conventional	Improved	Conventional	Improved	Conventional	Improved	Conventional	Improved
Information gathering and sharing	1	2	2.5	4	2	3	2	3	4.5	5.5
Addressing feelings of anxiety and discomfort	2.5	4	2	2.5	1.5	2	1.5	2	1	1.5
Adherence to fundamentals	2	2	1.5	3	2.5	2.5	2	2	2	2
Collaboration with others	0.5	1	2	3	0.5	1.5	0.5	2	3.5	3.5
Building a sense of purpose	0.5	0.5	1.5	2.5	1	1.5	0	0.5	1.5	1.5

- The number of factors extracted increased by approximately 18% to 30%. The effect of improvements was confirmed. The three items—"Information Collection and Sharing," "Moderate Anxiety/Discomfort," and "Collaboration with Others"—showed improvement across all groups.
- In Group B, the fulfillment levels increased in all evaluation items after support, with particularly notable improvements in "Information Collection and Sharing" and "Building a Sense of Achievement."

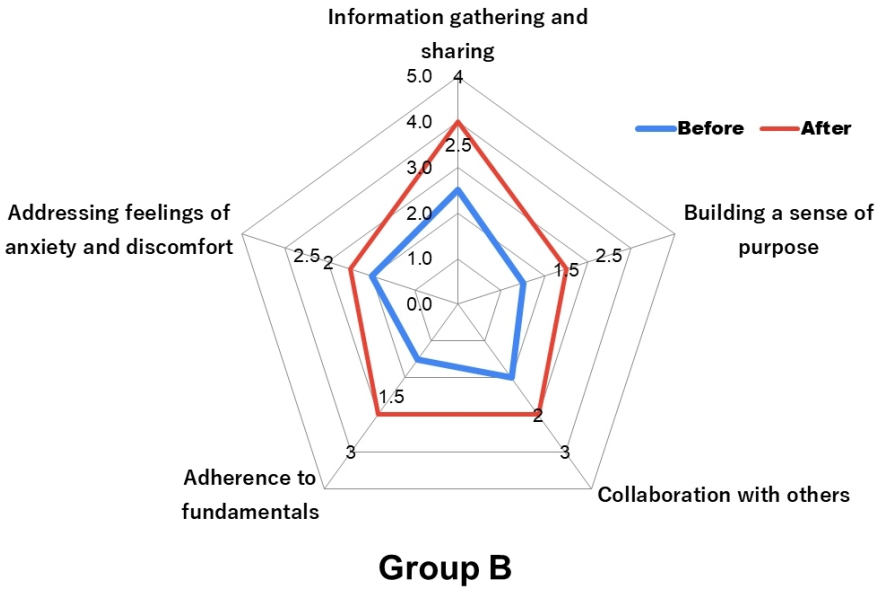

Figure 5: Fulfillment levels in group B.

These results suggest that incident analysis using the support system contributes to an increase in the number of factors extracted and has a significant effect on improving fulfillment levels. Additionally, it was revealed that there are significant variations in the contribution to fulfillment levels

depending on the characteristics. The background of this issue may include differences in the ease of noticing factors, where characteristics strongly associated with easily noticeable factors tend to show higher scores, and insufficient data for determining βij. Therefore, it is considered that focusing on the difficulty of factor extraction and reflecting this in the calculation formula could further improve accuracy in the future.

CONCLUSION

In this study, a prototype was developed using LLM (Large Language Models) to assist in the extraction of factors related to awareness capabilities and the formulation of countermeasures, regardless of the user's knowledge or experience. Based on the adaptability evaluations by several healthcare safety managers who used this system and the subsequent survey results, the following three conclusions were drawn:

1. **No Increase in Work Time**: When comparing incident analysis tasks performed using the support system with traditional incident analysis tasks, no increase in work time was observed.
2. **Enhanced Proactive Attitude:** Analysts were found to have adopted a more proactive attitude towards incident analysis.
3. **Wider Range of Awareness**: Verification results of fulfillment changes indicated that there was an objective widening of "awareness."

These results suggest that incorporating support for enhancing awareness capabilities into incident analysis tasks is feasible. Moving forward, we aim to make the system more useful on the ground by expanding the sample size for validation, identifying and addressing any issues, improving classification accuracy through LLM, and evaluating and refining the advice data generated.

REFERENCES

Flin R, Glavin R, Maran N, et al.: Anaesthetists' Non-Technical Skills (ANTS) System Handbook v1.0. University of Aberdeen, 2012.

Flowerdew L, Brown R, Vincent C, Woloshynowych M. Identifying nontechnical skills associated with safety in the emergency department: A scoping review of the literature. Ann Emerg Med 2012: 59: 386–394.

Gordon M, Darbyshire D, Baker P. Non-technical skills training to enhance safety: A systematic review. Med Educ 2012: 46: 1042–1054.

Irwin A, Weidmann AE. A mixed methods investigation into the use of non-technical skills by community and hospital pharmacists. Res Social Adm Pharm 2015: 11: 675–685.

Mitchell L, Flin R, Yule S, Mitchell J, Coutts K, Youngson G. Development of a behavioural marker system for scrub practitioners' non-technical skills (SPLINTS system). J Eval Clin Pract. 2013 Apr;19(2): 317–23.

Sako T., Noro C., Ito K. (2007)."WHO Global Competency Model", Mie Prefectural College of Nursing.

Swain A. D., and Guttmann H. G. (1983). "Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications Final Report," NUREG/CR-1278, U. S. Nuclear Regulatory Commission.

Undre S, Sevdalis N, Healey AN, Darzi A, Vincent CA. Observational teamwork assessment for surgery (OTAS): refinement and application in urological surgery. World J Surg. 2007.

Wu WA. Reaching common ground: The role of shared mental models in patient safety. J Patient Saf Risk Manag 2018: 23: 183–184.

Yule S, Flin R, Paterson-Brown S, Maran N, Rowley D. Development of a rating system for surgeons' non-technical skills. Med Ed 2006;40: 1098–104.