

ORIGINAL ARTICLE

Open Access

Content Analysis of Patient Safety Incident Reports Using Text Mining: A Secondary Data Analysis



Baek, On-Jeon¹ · Moon, Ho Jin² · Kim, Hyosun³ · Shin, Sun-Hwa⁴

¹Doctoral Student, Sahmyook University Graduate School, Seoul, Korea

²Manager, Quality Improvement Department, Sahmyook Medical Center, Seoul, Korea

³Nurse, Quality Improvement Department, Sahmyook Medical Center, Seoul, Korea

⁴Associate Professor, College of Nursing, Sahmyook University, Seoul, Korea

Purpose: This study aimed to identify the main keywords, network structures, and topical themes in patient safety incident reports using text network analysis. **Methods:** The study analyzed patient safety incident reports from a general hospital in Seoul, covering a total of 3,576 cases reported over five years, from 2019 to 2023. Unstructured data were extracted from the text of the incident reports, detailing how the patient safety incidents occurred and how they were managed according to the six-part principles. The analysis was conducted in four steps: 1) word extraction and refinement, 2) keyword extraction and word network generation, 3) network connectivity and centrality analysis, and 4) topic modeling analysis. The NetMiner program was used for data analysis. **Results:** The analysis of degree, betweenness, and closeness centrality revealed that the most common keywords among the top five were "confirmation," "medication," "inpatient room," "caregiver," and "condition." Topic modeling analysis identified three main topic groups: 1) incidents caused by a lack of awareness of fall risk, 2) incidents of non-compliance with basic medication principles, and 3) incidents due to inaccurate patient identification. **Conclusion:** To prevent patient safety incidents, it is necessary to promote a culture of safety in hospitals, standardize patient identification procedures, and provide basic training in medication safety and fall prevention to healthcare staff. Furthermore, empirical research on patient safety practices is necessary to encourage active participation in patient safety activities by patients and family caregivers.

Key Words: Accidental falls; Medication errors; Patient identification systems; Patient safety; Social network analysis

INTRODUCTION

Patient safety involves minimizing the risk of harm or near misses to inpatients while they receive medical care at a healthcare institution [1,2]. Generally, 4 out of 10 patients at medical institutions are exposed to patient safety incidents, and between 3% and 17% of inpatients experience adverse patient safety events. Notably, more than 50% of these incidents are reported to be preventable during the medical service delivery process [3,4]. Therefore, it is crucial to identify the causes of patient safety incidents in advance and to pinpoint and enhance system weaknesses to prevent such incidents [5].

The Korea Patient Safety Reporting and Learning

System was established and is managed by the Korea Accreditation Agency for Medical Institutions (KAAMI). This system was developed to systematically gather data on patient safety incidents following the enactment of the Patient Safety Act (PSA) in July 2016 [6]. Since its inception, the number of reported patient safety incidents has consistently risen: starting with 563 in 2016, increasing to 3,864 in 2017, 9,250 in 2018, 11,953 in 2019, 13,919 in 2020, 13,146 in 2021, and reaching 14,820 in 2022 [7]. However, it is believed that the actual number of patient safety incidents is likely higher. This underreporting can be attributed to a general reluctance within healthcare institutions to openly discuss medical errors and to the voluntary nature of the reporting system, which relies on

Corresponding author: Shin, Sun-Hwa <https://orcid.org/0000-0003-4052-9542>

College of Nursing, Sahmyook University, 815 Hwarang-ro, Nowon-gu, Seoul 01795, Korea.

Tel: +82-2-3399-1595, Fax: +82-2-3399-1594, E-mail: shinsh@syu.ac.kr

Received: Jun 1, 2024 / Revised: Aug 26, 2024 / Accepted: Nov 8, 2024

This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

patient safety personnel and heads of healthcare institutions [8].

Patient safety incident reports are compiled and maintained to track patient safety incidents at each healthcare institution. There are three classifications of patient safety incidents: near misses, adverse events, and sentinel events. A near miss is an error that occurs but does not affect the patient; an adverse event refers to harm caused by a medical intervention unrelated to the patient's underlying disease; and a sentinel event is defined as any occurrence that leads to unexpected death or significant loss of function, not associated with the natural progression of the patient's underlying disease state [9]. KAAMI categorizes patient safety incidents into six levels of harm: near misses, harmless, mild, moderate, severe, and death. According to the 2022 Yearbook of Patient Safety Statistics, the distribution was 35.6% near misses, 25.0% harmless, 26.9% mild, and 12.1% (1,792 cases) moderate, severe, or resulting in death [7].

Previous studies on patient safety incidents have indicated that adverse events occur more frequently among individuals aged 60 years or older, particularly in inpatient and treatment rooms during the daytime working hours of nurses. Among the various types of safety incidents, falls, blood transfusions, and medication errors are most common [8]. The occurrence rates of near misses and adverse events are higher in men than in women, whereas sentinel events occur more frequently in women [10,11]. Regarding the location of these incidents, near misses, adverse events, and sentinel events predominantly take place in inpatient rooms [8,10,11]. Falls are noted as the most frequent type of patient safety incident [10,11]. The reporting and analysis of patient safety incidents serve as essential data for the prevention and reduction of future occurrences. Therefore, the frequency and rate of these incidents are under investigation not only in major countries worldwide but also in Korea [7].

With the growing emphasis on the voluntary reporting of patient safety incidents, stakeholders and organizations dedicated to patient safety are actively developing guidelines, conducting research, and launching campaigns in various forms [12,13]. The National Patient Safety Foundation has issued eight guidelines aimed at enhancing patient safety, which include measures to support healthcare personnel and organizations in improving safety practices and involving patients and their caregivers [1]. The Institute of Medicine has highlighted the necessity of creating a safe medical environment. This involves establishing a system for error reporting, fostering a culture of patient safety within healthcare institutions, and providing train-

ing for healthcare personnel [14]. This underscores the need for the active involvement of governments, hospitals, healthcare providers, and the general public, including patients and family caregivers, in improving patient safety. However, the majority of reports and studies on patient safety incidents continue to focus merely on the frequencies and percentages of incidents, which hampers the identification of contributing factors [8]. A more effective analysis of factors related to patient safety incident reporting and the data reported could lead to the development of appropriate interventions. These interventions could be crucial in preventing and reducing the recurrence of patient safety incidents.

Recently, advancements in information and communication technology have simplified big data analysis, leading to active research on semi-structured and unstructured data. Text mining techniques, which analyze unstructured text data, have particularly garnered significant attention. These techniques involve extracting meaningful information through natural language processing [15]. Among these, topic modeling is employed to examine research trends and social issues by categorizing topics within documents and identifying key terms associated with each topic [16,17]. In one study, topic modeling was applied to online news articles to observe shifts in social issues before and after the implementation of the PSA. This analysis yielded eight topics: medical practice, medical personnel, infection and facilities, comprehensive nursing service, medicines, medicine and medical supplies, system development and establishment for improvement, PSA, and healthcare accreditation. The findings indicated that the emphasis shifted from patient safety awareness, infection control, and healthcare accreditation before the PSA implementation to a focus on patient safety culture and medication afterward [18]. While there have been studies that utilized text network analysis on academic papers and online articles concerning patient safety [18,19], no research has yet been conducted on the analysis of texts from patient safety incident reports written directly by hospitals.

This study aimed to identify keywords and themes by analyzing the patient safety incident reports through the lens of text network analysis. Specifically, it sought to discover keywords and themes associated with patient safety incidents by examining the documented problem situations based on the six-part principles outlined in previous reports. The use of text mining techniques facilitated this analysis, providing a foundation for the prevention of patient safety incidents and the development of intervention plans in hospitals.

METHODS

1. Study Design

This descriptive research study was conducted to identify keywords, network structures, and major topics by analyzing the patient safety incident reports from a general hospital in Seoul with more than 300 beds over a five-year period.

The specific methodology was as follows: first, identifying the general characteristics associated with patient safety incidents; second, examining the co-occurrence frequency of major keywords related to these incidents; third, assessing the network structure and centrality (degree centrality, betweenness centrality, and closeness centrality) within the network of patient safety incidents; and fourth, identifying major topics through topic modeling of the incident reports.

2. Analytical Objects and Research Procedures

This study aimed to analyze the problematic situations described in patient safety incident reports from a general hospital in Seoul. Sahmyook Medical Center, a general hospital with over 300 beds, has a Quality Improvement (QI) department and employs dedicated patient safety staff. The Institutional Review Board of Sahmyook Medical Center approved this study for secondary data analysis (IRB No. 116286-202401-HR-02).

The data collection period was from January 1, 2019 to December 31, 2023. Two nurses responsible for patient safety in the QI department at Sahmyook Medical Center conducted the data extraction. They focused on unstructured data from the main text, specifically occurrences and responses to patient safety incidents, based on the six-part principles outlined in the "Patient Safety Incident Report" compiled over five years. To ensure confidentiality, the extracted data were coded to remove sensitive personal information of both patients and reporters. The data were then organized into a password-protected Excel spreadsheet, accessible only to the research team. Prior to analyzing the data, a spell check was conducted, and typos were corrected by reviewing the incident reports. The analysis encompassed 3,576 cases reported over the 5-year period from 2019 to 2023.

3. Data Analysis

Data analyses were conducted using SPSS version 25.0 (IBM Institute, NY, USA) and NetMiner version 4.5.1.c

(Cyram Incorporated, Seoul, Korea). The general characteristics associated with patient safety incidents were examined by analyzing their frequency and percentage. The analysis then proceeded with several steps: word extraction and purification, keyword extraction and word network creation, analysis of network connection structures and centrality, and topic modeling.

1) Word extraction and refinement

To extract morphemes (the minimum semantic units) from the incident reports (unstructured text), each report was entered as a single row in an Excel spreadsheet. Using NetMiner's morpheme extraction method, pronouns, adverbs, and other non-essential parts of speech were automatically removed to highlight key concepts, leaving only nouns [20]. Additionally, it was determined that single-letter words (in Korean) such as "the," "this," and "that" lack meaningful content, so words with fewer than two letters were excluded using the query function. For network analysis, dictionaries of defined words, thesaurus words, and exception words were created. Text analysis requires distinguishing between singular and plural forms as they are treated as separate entities; thus, word refinement is essential. Defined words are phrases identified when two or more words combine to convey a single meaning [21]. In this study, 146 terms were registered as defined words, including compound nouns. For instance, "position change" was treated as a single term rather than two separate words. Thesaurus words are terms that share the same or similar meanings but are presented differently across reports. These are unified under a single representative term [21]. In this study, 56 words were registered; "bed rail," "bed handrail," and "side rail" were all standardized to "side rail," "body guard" to "bodyguard," "body restraint" to "restraint," and "medication," "drug," and "oral medication" to "medication." Similarly, "drop down," "slip," and "fall" were consolidated under "fall." Exception words, considered irrelevant to the analysis, were also identified [21]. In this study, 31 words were preemptively excluded. Given the focus on patient safety incident reports, terms like "patient," "incident," "report," and "safety" were omitted. Additionally, words pertaining to the six W principles commonly used in these reports—"who," "when," "where," "what," and "how"—were also removed.

2) Extraction of keywords and network formation

Using the distance information between extracted words, an inter-word network was constructed to identify connections among words that appear in close proximity. For network analysis, only words with a moderate level of

connectivity were included. However, the criteria for what constitutes moderate connectivity have not been established, and there are no clear guidelines on the number of words with a specific frequency of occurrence to include in network generation [22]. In previous studies, the selection of words for the network was based on context and meaning, the interpretability of the research results, and the degree of visualization [22]. Therefore, this study visualized the network using the frequency of the top 30 co-occurring words and converted the two-mode network of word-sentence into a one-mode word-word network for analysis. Referring to prior research on text network analysis [23], the conversion to a 1-mode network involved using the inner product coefficient, which accounts for the overall weight of the words in the similarity index. The inner product coefficient facilitates an intuitive assessment by analyzing the size and direction of the words, and it offers the advantage of consistent results regardless of the order of the words [23]. The range of co-occurrence was adjustable by setting the distance between two words. In this study, the window size was set to "3" based on the characteristics of the data analyzed and previous studies, without considering directionality [17]. Taking into account the degree of visualization and contextual interpretation of the study, a network was created based on a co-occurrence frequency of 2.

3) Analysis of network connectivity and centrality

Analysis was conducted at both the network and node levels, focusing on the network that was created. At the network level, the size and density of the networks were assessed, while at the node level, the average degree and mean distance between networks were evaluated. Degree centrality is a measure of centrality that quantifies the number of connections a word has with other words in the network. A word with high degree centrality is linked to many other words and is considered a keyword within the network [22]. The centrality index of a network indicates how centrally a word is positioned within the network. A word with high centrality is deemed a core keyword. The value of centrality ranges between 0 and 1, with higher values indicating greater centrality for the word at the center of the network. Betweenness centrality is another centrality measure that reflects the extent to which a word lies on the shortest path between other words, indicating its role as a bridge connecting two words [24]. Closeness centrality measures the contextual similarity of a word to its neighbors based on its proximity to other words in the network; a higher closeness centrality suggests that the word is more contextually similar to its neighbors [22]. In this

study, the top 30 words were identified based on degree centrality, betweenness centrality, and closeness centrality. Centrality was analyzed visually using a spring map. In the visualization, the larger the word size, the higher the centrality index, and the thicker the link, the higher the co-occurrence frequency between words.

4) Topic modeling analysis

Topic modeling is a statistical method that utilizes Latent Dirichlet Allocation (LDA) algorithm to identify specific topics within sentences. This algorithm calculates the probability distribution of word occurrences in a sentence and identifies potential trends in topics. A silhouette coefficient closer to "1" indicates better clustering [25]. Drawing on previous studies that utilized the silhouette coefficient [23], we employed k-means clustering to determine the number of topics. Additionally, based on prior research concerning LDA's input options, we used the Monte Carlo Markov chain method. By varying the α value, β value, and the number of topics, we extracted topics through repeated simulations [23,26]. In this study, after identifying the top 20 topic models with the highest silhouette coefficients through multiple simulations, we narrowed our focus to the top 10. These were exclusively classified using the topic words that appeared in the models. After reviewing the original content of the words and the sentences assigned to the selected topics, the research team convened to name each topic, reflecting its underlying theme. The topic groups were named by referencing the top words for each topic, and the top 10 words were visualized on a topic-keyword map using a 2-mode network of topic-words.

RESULTS

1. General Characteristics Associated with Patient Safety Incidents

This study analyzed 3,576 patient safety incident reports. Table 1 displays the general characteristics associated with these patient safety incidents. The annual distribution of incident reports was as follows: 630 reports (17.6%) in 2019, 698 reports (19.5%) in 2020, 776 reports (21.7%) in 2021, 718 reports (20.1%) in 2022, and 754 reports (21.1%) in 2023. In terms of incident type, falls were the most common, accounting for 1,859 cases (52.0%), followed by medication/transfusion errors, which comprised 1,114 cases (31.1%). The breakdown of incidents included 169 near misses (4.7%), 2,545 harmless events (71.2%), 854 adverse events (23.9%), and eight sentinel events (0.2%). The gen-

Table 1. General Characteristics Associated with Patient Safety Incident Reports (N=3,576)

Variables	Categories	n (%)
Year of PSI	2019	630 (17.6)
	2020	698 (19.5)
	2021	776 (21.7)
	2022	718 (20.1)
	2023	754 (21.1)
Types of PSI	Fall	1,859 (52.0)
	Medication/ transfusion	1,114 (31.1)
	Surgery/ treatment/ procedure	49 (1.4)
	Others	554 (15.5)
Severity of PSI	Near miss	169 (4.7)
	Harmless event	2,545 (71.2)
	Adverse event	854 (23.9)
	Sentinel event	8 (0.2)
Patient's gender	Men	1,863 (52.1)
	Women	1,713 (47.9)
Patient's age (year)	0~19	69 (1.9)
	20~39	179 (5.0)
	40~59	603 (16.9)
	60~79	1,630 (45.6)
	≥ 80	1,088 (30.4)
	Missing	7 (0.2)
Medical department	Internal medicine department	2,449 (68.5)
	Surgical department	923 (25.8)
	Clinical support department	204 (5.7)
Reporter	Nurse	3,350 (93.7)
	Physical therapist	64 (1.8)
	Clinical pathologist	45 (1.3)
	Radiologist	34 (0.9)
	Others	83 (2.3)
Location of PSI occurrence	Inpatient room	1,624 (45.4)
	Hospital bed	783 (21.9)
	ICU or ER	305 (8.5)
	Outpatient examination room or operating room	270 (7.6)
	Restroom or shower room	224 (6.3)
	Hallway, lobby, or stairs	161 (4.5)
	Others	209 (5.8)

ER=emergency room; ICU=intensive care unit; PSI=patient safety incident.

der distribution of patients involved in these incidents was relatively balanced, with 1,863 cases (52.1%) involving men and 1,713 cases (47.9%) involving women. The predominant age group affected was 60~79 years, representing 1,630 cases (45.6%). Among the medical departments, internal medicine reported the highest number of cases, with 2,449 incidents (68.5%). Nurses were the most frequent reporters of incidents, contributing to 3,350 cases (93.7%). The majority of incidents occurred in the "inpatient room," with 1,624 cases (45.4%).

2. Top Keywords for Patient Safety Incidents

After registering the defined, thesaurus, and exception words from the patient safety incident reports, a total of 4,507 words were extracted through morphological analysis. Upon examining the frequency of these words, it was found that words occurring only once constituted 50.2% of the total, while those appearing 10 or more times made up 13.0%, and words with 15 or more appearances accounted for 9.6%. For this study, we excluded single-letter words due to their ambiguous nature and selected words that were at least two letters long and appeared at least 15 times to ensure the results were meaningful. Ultimately, 370 words met these criteria.

The frequency of occurrence and connectivity rankings of the words used in the patient safety incident reports were analyzed, with the top 30 words presented in Table 2. The ten keywords with the highest connectivity were: confirmation (161 occurrences), medication (144 occurrences), inpatient room (131 occurrences), caregiver (124 occurrences), condition (121 occurrences), bed (111 occurrences), examination (93 occurrences), prescription (92 occurrences), nurse (83 occurrences), and restroom (79 occurrences). Notably, "confirmation" emerged as the most frequently appearing word in the patient safety incident reports.

3. Analysis of Network Connectivity and Centrality of Patient Safety Incidents

In this study, a network was constructed based on the criteria of a co-occurrence frequency of more than three times and a proximity of within two words between terms. This resulted in a network comprising 1,675 nodes and 2,789 links. The network's density was determined to be 0.04, with an average degree of 15.07, a mean distance of 2.36, and a maximum diameter of 5.00. These metrics indicate that each keyword, on average, co-occurs with 15 other keywords, traverses a path of approximately 2.36 steps, and all keywords are interconnected within four steps.

To identify the network characteristics of incident reports, we analyzed the top 30 words based on degree, betweenness, and closeness centralities, as shown in Table 2. The results visualized using a spring map are presented in Figure 1. The top five words across the three centrality analyses were "confirmation," "medication," "inpatient room," "caregiver," and "condition," with similar rankings and compositions noted in each analysis. Words such as "explanation," "surgery," "situation," and "fluid" were in-

Table 2. Top 30 Keywords by Frequency and Centrality Analysis

(N=4,576)

Rank	Frequency		Degree centrality		Betweenness centrality		Closeness centrality	
1	Confirmation	161	Confirmation	0.436	Confirmation	0.140	Confirmation	0.626
2	Medication	144	Medication	0.390	Medication	0.118	Medication	0.604
3	Inpatient room	131	Inpatient room	0.355	Inpatient room	0.101	Inpatient room	0.601
4	Caregiver	124	Caregiver	0.336	Caregiver	0.090	Caregiver	0.592
5	Condition	121	Condition	0.328	Condition	0.084	Condition	0.589
6	Bed	111	Bed	0.301	Bed	0.079	Bed	0.575
7	Examination	93	Examination	0.252	Examination	0.060	Prescription	0.556
8	Prescription	92	Prescription	0.249	Restroom	0.036	Examination	0.555
9	Nurse	83	Nurse	0.225	Prescription	0.035	Nurse	0.551
10	Restroom	79	Restroom	0.214	Discovery	0.035	Discovery	0.548
11	Discovery	76	Discovery	0.206	Nurse	0.028	Occurrence	0.543
12	Injection	73	Injection	0.198	Occurrence	0.028	Injection	0.538
13	Occurrence	73	Occurrence	0.198	Doctor	0.026	Clinical support department	0.527
14	Fall	67	Fall	0.182	Injection	0.023	Doctor	0.527
15	Doctor	61	Doctor	0.165	Time	0.019	Fall	0.527
16	Procedure	59	Procedure	0.160	Error	0.017	Procedure	0.526
17	Error	56	Error	0.152	Explanation	0.017	Time	0.518
18	Time	56	Time	0.152	Performance	0.017	Explanation	0.515
19	Explanation	53	Explanation	0.144	Procedure	0.016	Restroom	0.511
20	Clinical support department	50	Clinical support department	0.136	Fall	0.016	Surgery	0.509
21	Alone	44	Alone	0.119	Fluid	0.013	Situation	0.508
22	Surgery	44	Surgery	0.119	Result	0.013	Performance	0.508
23	Result	42	Result	0.114	Surgery	0.012	Result	0.506
24	Performance	40	Performance	0.108	Treatment	0.011	Error	0.506
25	Floor	40	Floor	0.108	Clinical support department	0.010	Treatment	0.501
26	Treatment	39	Treatment	0.106	Individual	0.009	Patient identification	0.501
27	Situation	38	Situation	0.103	Floor	0.009	Fluid	0.501
28	Chair	35	Chair	0.095	Alone	0.008	Usage	0.495
29	Fluid	35	Fluid	0.095	Situation	0.008	Morning	0.492
30	Movement	33	Movement	0.089	Patient identification	0.008	Admission	0.490

cluded in the rankings due to their high centrality despite their low frequency of appearance. Conversely, terms like "internal medicine department," "before medication," and "after administration" appeared frequently but were excluded from the ranking because of their low centrality.

4. Topic Modeling of Patient Safety Incidents

The number of topics in the incident report was determined using k-means clustering with silhouette coefficients, and the calculations were performed using the default settings of the NetMiner program. To identify the

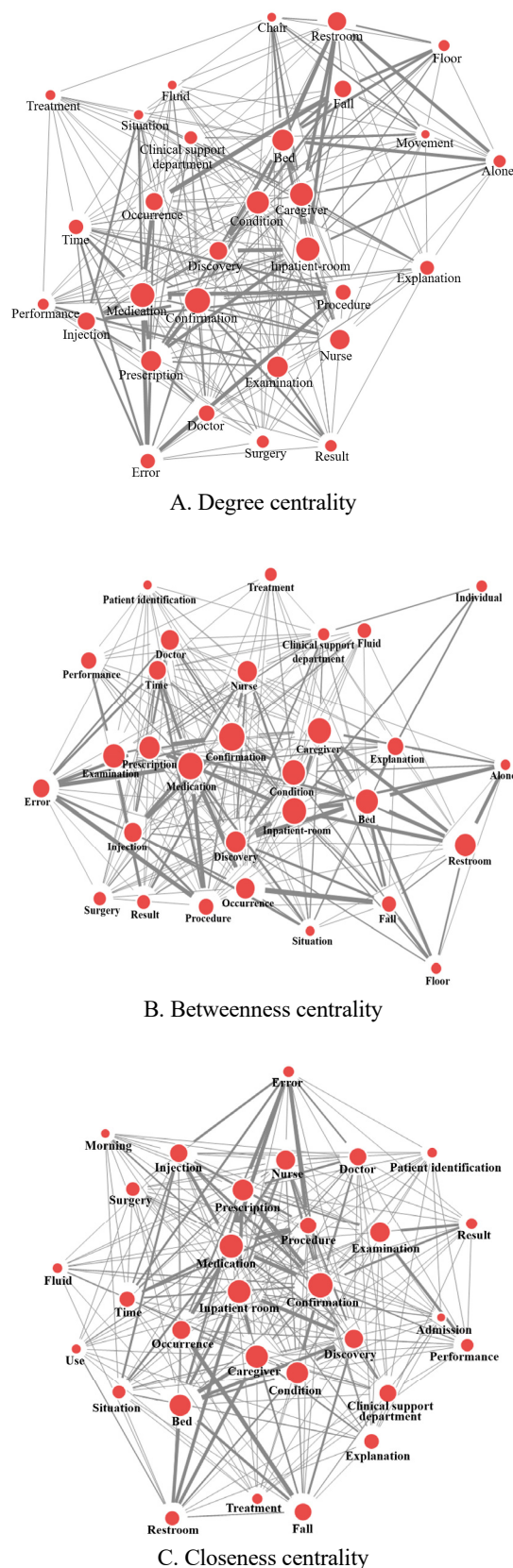


Figure 1. Spring network map of centrality.

optimal number of topics, the settings were adjusted to an α value between 0.1 and 0.2, a β value between 0.01 and 0.02, and the number of iterations set to 1,000, resulting in 224 possible combinations. The feasibility of these topic numbers was assessed by analyzing the distribution of the top 10 silhouette coefficients among the 224 combinations. Following a research team meeting, the top 10 topic models based on the silhouette coefficients were identified. Three topic groups were then selected by setting the α value to 0.2, the β value to 0.01, and the number of iterations to 1,000. These groups were chosen because they ranked in the top 5 for silhouette coefficient values and were distinct in content.

The number of potential topics was set at three, and topic modeling was conducted using the LDA technique. The top five keywords for each of the three topics were identified based on their probability distribution. The topics were named by examining the main text of the incident report that contained these keywords. Keywords and topics for each group are presented in Table 3. Additional keywords related to the major topics are displayed in the network map (Figure 2).

1) Topic 1. Incidents caused by a lack of awareness of fall risk

Topic 1 comprised 41.7% of the total sentences, making it the most prevalent. The keywords associated with this topic were fall, bed, occurrence, restroom, and inpatient room. It focused on patient safety incidents related to falls that happened at night, particularly when older patients were getting out of bed, transferring to the restroom, or moving around in bed without the presence of a family caregiver. Consequently, this topic was labeled "incidents caused by a lack of awareness of fall risk."

2) Topic 2. Incidents of non-compliance with basic medication principles

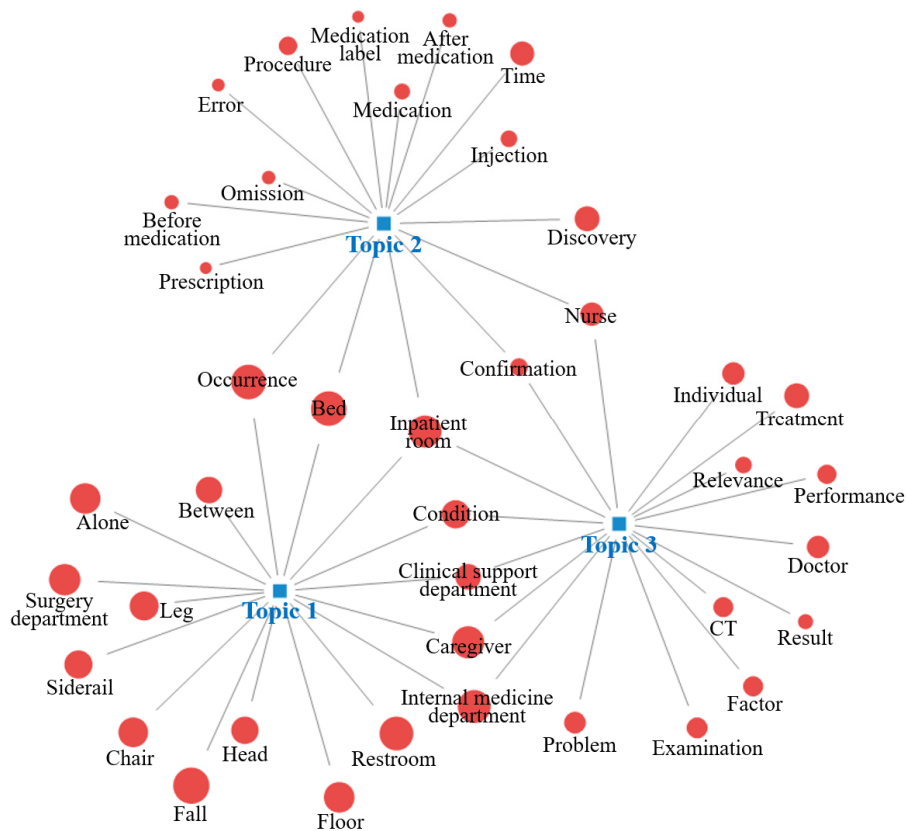
Topic 2 comprised 27.0% of the total sentences, with keywords such as medication, nurse, confirmation, discovery, and error. This topic encompassed patient safety incidents including failure to check the prescription during the medication process, misinterpretation of the prescription, confusion regarding the time of administration, and identification of a medication error prior to administering it to the patient. Consequently, this topic was designated as "incidents of non-compliance with basic medication principles."

3) Topic 3. Incidents due to inaccurate patient identification

Topic 3 comprised 31.3% of the total sentences, with keywords such as examination, confirmation, doctor, con-

Table 3. Topic Groups Based on Topic Modeling

Categories	1st keyword	2nd keyword	3rd keyword	4th keyword	5th keyword	Topic group	Number of sentences (%)
Topic 1	Fall	Bed	Occurrence	Restroom	Inpatient room	Incidents caused by a lack of awareness of fall risk	2,305 (41.7)
Topic 2	Medication	Nurse	Confirmation	Discovery	Error	Incidents of non-compliance with basic medication principles	1,497 (27.0)
Topic 3	Examination	Confirmation	Doctor	Condition	Clinical support department	Incidents due to inaccurate patient identification	1,730 (31.3)
Total							5,532 (100.0)

**Figure 2.** Topic group network by topic modeling.

dition, and clinical support department. This topic included patient safety incidents, including failures to identify a patient during transportation for a test or procedure, while connecting a medical device, and in cases of patients with reduced levels of consciousness. Consequently, this topic was named "incidents due to inaccurate patient identification."

DISCUSSION

This study is meaningful because it is the first to conduct a text network and topic modeling analysis of patient safety incident reports written directly by healthcare workers. The insights gained from this research help elucidate the characteristics of patient safety incidents in hospi-

tals and lay the groundwork for devising strategies to prevent such incidents. The main findings of the study are outlined below.

Keywords with a high frequency of occurrences identified in the patient safety incident reports of this study included confirmation, medication, inpatient room, caregiver, condition, bed, examination, prescription, nurse, and restroom. Previous studies have not analyzed patient safety incident reports written directly by medical workers, making direct comparisons difficult. However, an analysis of 10 years of "patient safety" articles in online news found that surgery, treatment, examination, center, and professor were frequent keywords [18]. Additionally, an analysis of article abstracts by searching for the topic of "patient safety culture" revealed that safety, awareness, nurse, hospital, and management were frequent keywords [19], which differs from our study. In this study, it was confirmed that "confirmation" co-occurred with the most keywords. The keywords that appeared high in centrality in this study were "confirmation," "medication," "inpatient room," "caregiver," and "condition." This analysis helped identify the main keywords in reports related to patient safety incidents. Furthermore, the keyword "confirmation" exhibited the highest centrality in this study, highlighting the importance of conducting double-checks between healthcare providers before performing high-risk and invasive medical procedures as a critical factor in ensuring patient safety. The keywords derived from this study had a lower network density and a higher degree of connection between keywords compared with previous research [22]. This implies that incident reports containing a specific topic contain representative content and that the topics written in the incident report are limited. Because patient safety incident reports are written based on the 6th principle, which involves a situation where harm to a patient occurred or nearly occurred, they can be seen as reflecting the characteristics of a specific situation. Based on the keywords examined through network connectivity and centrality analyses, we found that patient safety incidents were composed of content such as where they occurred, who found them, and why they occurred.

Topic modeling identified "incidents caused by a lack of awareness of fall risk" as Topic 1, which represented 41.6% of all documents and exhibited the highest rate. This topic encompasses fall accidents involving elderly patients who slip or fall in various hospital locations, with a notable frequency of falls occurring at night or early in the morning when patients with diminished cognitive function are unaccompanied by caregivers. In this study, patients aged 60 years and older constituted over 70% of patient safety in-

cidents, with falls being the most common type. This high frequency of falls aligns with data reported through the Korean Patient Safety Reporting and Learning System, which also indicated falls as the predominant type of incident [7, 8, 10, 11]. This finding is consistent with the practice of hospital patient safety staff entering incident reports into the patient safety reporting and learning system. Previous research has documented falls occurring in various hospital areas, including inpatient rooms, hallways, cafeterias, and restrooms. The severity of these falls can be significant, sometimes leading to sentinel events [10,27]. This study further confirmed that the most frequent site of patient safety incidents was the "inpatient room," with falls primarily occurring in "beds" and "rest/shower rooms." The challenge in these settings is the difficulty for nurses to continuously monitor patients, which increases the risk of adverse events if immediate interventions are not available. To mitigate fall risks within the hospital, it is crucial to provide guidance and education about fall prevention not only to healthcare professionals but also to patients and caregivers. This education should be offered not only at admission but also periodically throughout the hospital stay, tailored to the duration of stay and the severity of the patient's condition. For patients with impaired cognitive function, it is essential to develop and implement specialized, customized fall prevention programs. Additionally, there must be vigilant monitoring of the entire process from fall prevention to response, with concerted efforts to prevent the recurrence of fall incidents.

Topic 2 focused on "incidents of non-compliance with basic medication principles," primarily addressing patient safety incidents involving nurses who were found preparing or administering medication to patients. Statistics from Korean patient safety incident reporting systems indicate that medication errors are the second most common type of patient safety incident [7,8,11]. Similarly, reports from China and Spain identify medication errors as the most frequent patient safety incidents directly associated with hospital care [28,29]. Medication errors temporarily assist in safeguarding patients, aiding medical personnel in ensuring patient safety and providing valuable services [30]. Medication accidents in hospitals, often resulting from unsafe medical practices or errors, highlight the need for comprehensive education and training for healthcare workers who handle medications. Furthermore, a previous study revealed that one-third of medication errors occur during the medication preparation process [30]. To prevent such errors, hospitals can implement measures such as barcodes and drug dosage calculation programs. It

is particularly crucial for nurses preparing medications to possess precise knowledge of the drugs and adhere strictly to medication administration regulations. Consequently, nursing education places significant emphasis on the 5-right principle to enhance safety and prevent medication errors. The findings of this study indicated that nurses were the primary reporters in incident reporting. Topic modeling of online news also showed an increased emphasis on the role of nurses in direct care following the implementation of the PSA, particularly focusing on medication over infection control [18]. This underscores the critical role of nurses in patient safety. The study also found that reports of harmless events and near misses, which did not cause direct harm to patients, accounted for over 75% of the total, demonstrating active participation in voluntary incident reporting. Currently, a diverse group of healthcare professionals, including nurses, physiotherapists, clinical pathologists, radiologists, and physicians, are involved in reporting patient safety incidents. However, due to the nature of these incidents, nurses are the most frequent reporters. Despite this, a system for regular training on patient safety prevention and voluntary reporting is essential for all healthcare workers to maintain vigilance regarding patient safety. Previous studies have also suggested that multifaceted efforts through patient safety research and campaigns are necessary to enhance voluntary reporting of patient safety events [12,13]. To achieve this, it is necessary to promote an environment that encourages voluntary reporting and mandates participation in patient safety incident prevention activities for all healthcare workers in the hospital, as well as for patients and family caregivers.

Topic 3, "incidents due to inaccurate patient identification," addresses errors that occur when patient identification is not accurately performed during all treatment processes, including examinations and procedures. Previous studies have indicated that errors in patient identification often arise in cases involving patients with the same or similar names [31]. Additionally, there is a significant risk that patient identification may be neglected due to emergencies or overwhelming workloads that healthcare workers cannot manage alone [32]. Errors in patient identification can occur in various contexts, such as incorrect labeling during blood collection or specimen handling in the laboratory, inaccurate identification during treatment or testing, and errors in surgical schedules or surgical site marking in the operating room [8,33]. In the medication administration process, primarily conducted by nurses, there have been instances of patient urgency, uncooperative attitudes, excessive workloads that are un-

manageable for nurses alone, or failures to identify patients in emergency situations [32]. Notably, nurses have reported that if they recognize a patient as a familiar face, they might assume the patient's identity and neglect proper identification procedures [32]. In such familiar scenarios, as medical treatment becomes more routine, it becomes increasingly crucial for medical professionals to strictly adhere to fundamental principles. Patient identification is recognized as a critical healthcare practice essential for providing safe care. It must be accurately performed throughout the medical process, from intake and reception through various procedures and operations, to administering medications, blood transfusions, specimen collection, imaging, and all diagnostic and therapeutic procedures [31]. Therefore, accurate patient identification is crucial for various medical professionals, including nurses, doctors, physical therapists, radiologists, and clinical pathologists, who are involved in performing these procedures. A previous study found that a higher level of congruence between nurses' personal and organizational values, along with their organizational and occupational commitment, was associated with better performance in patient verification behaviors [34]. It is vital to promote a safe organizational culture in hospitals to ensure that patient identification is consistently performed at the hospital level. It is also necessary to expand patient identification practices to encompass the entire institution by aligning individual and organizational values [32,34]. Before providing medical services, it is recommended to perform a standardized patient identification procedure that verifies two of the following: the patient's name, registration number, and date of birth. This should be done through open-ended questions and by comparing the patient identification bracelet with the prescription data [31, 35]. The uncooperative attitudes of patients and family caregivers during the patient identification process are a major cause of its non-performance [32]. Therefore, efforts should be made to actively involve patients and family caregivers in the patient identification process through targeted publicity and education, rather than it being solely the responsibility of healthcare workers. Furthermore, the development of these patient safety activities and the cultivation of a safety culture should be extended to all healthcare organizations and continuously developed at the national level.

This study is significant as it explores trends in the patient safety incident reports by analyzing the main keywords and semantic structures of reports compiled over a 5-year period. However, the study's limitation lies in its analysis being confined to reports authored by healthcare

workers at a single general hospital, which necessitates caution when generalizing the findings. Additionally, the focus on quantitative content analysis through text network analysis restricts the depth of analysis possible for each report. To address these limitations, several recommendations are proposed. First, it is essential to broaden the scope of the study to include patient safety incident report analyses from multiple hospitals. Second, although hospitals generate incident reports in various formats for quality management, there is a notable deficiency in research that analyzes the unstructured data within these reports. Future research should incorporate more qualitative studies to thoroughly examine patient safety incident experiences and enhance the analysis of safety incidents. Third, we suggest conducting an empirical study that engages actively in patient safety measures by implementing standardized procedures for patient identification, medication administration, and educational programs aimed at preventing falls, benefiting both patients and their family caregivers.

CONCLUSION

This study conducted a content analysis of patient safety incident reports to identify key words and topics associated with the occurrence of patient safety incidents. The analysis revealed that the most frequently occurring keywords over a five-year period were confirmation, medication, inpatient room, caregiver, and condition. Topic modeling identified three main topics: "incidents caused by a lack of awareness of fall risk," "incidents of non-compliance with basic medication principles," and "incidents due to inaccurate patient identification." Of these, "incidents caused by a lack of awareness of fall risk" were the most prevalent. To prevent patient safety incidents, it is imperative for hospitals to foster a culture of safety, implement standardized patient identification procedures, ensure adherence to basic medication safety principles, and provide fall prevention training for all healthcare workers. Additionally, empirical research into patient safety practices is necessary to encourage active participation in patient safety activities by patients and family caregivers.

CONFLICTS OF INTEREST

Sun-Hwa Shin has been a public relations director at the Korean Society of Adult Nursing since 2024~2025. She was not involved in the review process of this manuscript. She had no other conflicts of interest.

AUTHORSHIP

Study conception and design acquisition - MHJ and SSH; data curation - MHJ and KH; analysis - BOJ and SSH; interpretation of the data - BOJ, MHJ, KH, and SSH; and drafting or critical revision of the manuscript for important intellectual content - BOJ and SSH.

ACKNOWLEDGEMENT

The researchers are grateful to Sahmyook Medical Center officials for providing qualitative data from patient safety incident reports. This research was funded by the National Research Foundation of Korea (No. NRF-2022R1F1A106447).

REFERENCES

- Berwick DM, Shojania KG, Atchinson BK. Free from harm: Accelerating patient safety improvement fifteen years after To Err is Human. National Patient Safety Foundation Report. Boston, MA: National Patient Safety Foundation, 2016, June. Report No: 2015.
- World Health Organization. Patient safety [Internet]. Geneva: World Health Organization; 2019 [cited 2024 March 15]. Available from: <https://www.who.int/teams/integrated-health-services/patient-safety>
- Auraaen A, Slawomirski L, Klazinga N. The economics of patient safety in primary and ambulatory care: flying blind. Organization for Economic Cooperation and Development (OECD). 2018;106:1-57. <https://doi.org/10.1787/baf425ad-en>
- U.S. Department of Health and Human Services, Office of Inspector General. Adverse events in hospitals: National incidence among medicare beneficiaries. OEI Publication No, 06-09-00090 [Internet]. [updated 2010, November; cited 2024 March 15]. Available from: <http://oig.hhs.gov/oei/reports/oei-06-09-00090.pdf>
- Choi EY, Lee HJ, Ock MS, Lee SI. Development of Korean root cause analysis software for analyzing patient safety incidents. Quality Improvement in Health Care. 2018;24(1):9-22. <https://doi.org/10.14371/QIH.2018.24.1.9>
- Ministry of Health and Welfare, Korea Institute for Healthcare Accreditation. Korea Patient Safety Reporting & Learning System: Patient safety related system information. [Internet]. Seoul: Korea Institute for Healthcare Accreditation. [cited 2024 March 15]. Available from: <https://www.kops.or.kr/portal/pag/view/sysGui/systemGuide.do>
- Ministry of Health and Welfare, Korea Institute for Healthcare Accreditation. Korean Patient Safety Incident Report 2022 [Internet]. Seoul: Korea Institute for Healthcare Accreditation.

- [cited 2024 March 15]. Available from: <https://statistics.kops.or.kr/biWorks/dashBoardMain.do#>
8. Kim NY. Analysis of patient safety incident in Korea. *Journal of Korean Academy of Nursing Administration*. 2020;26(2): 151-9. <https://doi.org/10.1111/jkana.2020.26.2.151>
 9. Davies JM, Hebert PC, Hoffman C. *The Canadian patient safety dictionary*. Ottawa, Ont: Royal College of Physicians and Surgeons of Canada; 2003.
 10. Shin SH, Won MH. Trend analysis of patient safety incidents and their associated factors in Korea using national patient safety report data (2017-2019). *International Journal of Environmental Research and Public Health*. 2021;18(16):8482. <https://doi.org/10.3390/ijerph18168482>
 11. Jeon Y, Jeong J. Factors affecting level of patient safety incidents in Korean hospitals: using Korean patient safety incidents data 2018-2020. *Humanities and Social Sciences*. 2022; 13(6):2609-22. <https://doi.org/10.22143/HSS21.13.6.179>
 12. Baek SM, Jeon SH. A review of patient safety activities of nurses in Korea. *Journal of Learner-Centered Curriculum and Instruction*. 2021;21(6):657-67. <https://doi.org/10.22251/jlcci.2021.21.6.657>
 13. Cho HA, Shin H. A systematic review of published studies on patient safety in Korea. *Journal of Korean Academy of Dental Administration*. 2014;2(1):61-82. <https://doi.org/10.22671/JKADA.2014.2.1.61>
 14. Institute of Medicine. *Crossing the quality chasm: a new health system for the 21st century*. Washington, DC: The National Academies Press; 2001.
 15. Kim DH. *Window of Statistics* [Internet]. Statistics Training Institute, Daejeon; 2017 [cited 2024 March 15]. Available from: https://sti.kostat.go.kr/window/2017b/html/2017_win_3.html
 16. Lee HA, Kim JH. The trends and prospects of the medical sociology in Korea focusing on the National Journal Articles from 2002-2020. *Health and Social Sciences*. 2020;55(1):165-203. <https://doi.org/10.37243/kahms.2020.55.165>
 17. Lee SS. A study on the application of topic modeling for the book report text. *Journal of Korean Library and Information Science Society*. 2016;47(4):1-18.
 18. Kim N, Lee NJ. An analysis of changes in social issues related to patient safety using topic modeling and word co-occurrence analysis. *The Journal of the Korea Contents Association*. 2021; 21(1):92-104. <https://doi.org/10.5392/JKCA.2021.21.01.092>
 19. Baek SM, Moon IO. The study on the patient safety culture convergence research topics through text mining and CONCOR analysis. *Journal of Digital Convergence*. 2021;19(12):359-67. <https://doi.org/10.14400/JDC.2021.19.12.359>
 20. Cyram Corporation. *Social network analysis using NetMiner: text network analysis*. Gyeonggi-do: Cyram Corporation; 2021.
 21. Lee SS. A content analysis of journal articles using the language network analysis methods. *Journal of the Korean Society for Information Management*. 2014;31(4):49-68. <https://doi.org/10.3743/KOSIM.2014.31.4.049>
 22. Park CS, Park EJ. Identification of knowledge structure of pain management nursing research applying text network analysis. *Journal of Korean Academy of Nursing*. 2019;49(5):538-49. <https://doi.org/10.4040/jkan.2019.49.5.538>
 23. Seo Y. *Identification of knowledge structure of accidental falls nursing research: text network analysis and topic modeling* [dissertation]. Seoul: Chung-Ang University; 2021. p. 1-151.
 24. Lee SS. *Network analysis methodology*. Seoul: Nonhyung Publisher; 2012.
 25. Mehta V, Caceres RS, Carter KM. Evaluating topic quality using model clustering. Orlando, FL: IEEE Symposium on Computational Intelligence and Data Mining (CIDM); 2014. p. 178-85. <https://doi.org/10.1109/CIDM.2014.7008665>
 26. Park CS. National and international trends in research on moral distress in nursing: text-network analysis and topic modeling. *Korean Journal of Medical Ethics*. 2023;26(2):133-49. <https://doi.org/10.35301/ksme.2023.26.2.133>
 27. Ahn S, Kim DE. Factors affecting the degree of harm from fall incidents in hospitals. *Journal of Korean Academy of Nursing Administration*. 2021;27(5):334-43. <https://doi.org/10.11111/jkana.2021.27.5.334>
 28. Gao X, Yan S, Wu W, Zhang R, Lu Y, Xiao S. Implications from China patient safety incidents reporting system. *Therapeutics and Clinical Risk Management*. 2019;15:259-67. <https://doi.org/10.2147/TCRM.S190117>
 29. Aranaz-Andres JM, Aibar-Remon C, Vialler-Murillo J, Ruiz-Lopez P, Limon-Ramirez R, Terol-Garcia E. Incidence of adverse events related to health care in Spain: results of the Spanish National Study of adverse events. *Journal of Epidemiology and Community Health*. 2008;62(12):1022-9. <https://doi.org/10.1136/jech.2007.065227>
 30. Cloete L. Reducing medication errors in nursing practice. *Nursing Standard*. 2015;29(20):50-9. <https://doi.org/10.7748/ns.29.20.50.e9507>
 31. Yum HK. Concept and importance of patient identification for patient safety. *Journal of the Korean Medical Association*. 2015;58(2):93-9. <https://doi.org/10.5124/jkma.2015.58.2.93>
 32. Park B, Kim JH, Choi MK, Nam S, Kim HS, Park Y, et al. Registered nurses' experiences of missing patient identification in general wards. *Journal of Korean Association for Qualitative Research*. 2021;6(2):96-109. <https://doi.org/10.48000/KAQRKR.2021.6.96>
 33. Zeeshan MF, Dembe AE, Seiber EE, Lu B. Incidence of adverse events in an integrated US healthcare system: a retrospective observational study of 82,784 surgical hospitalizations. *Patient Safety in Surgery*. 2014;8:23(1-10).

<https://doi.org/10.1186/1754-9493-8-23>

34. Kim YM, Kang SW, Kim SY. Factors related to nurses' patient identification behavior and the moderating effect of person-organization value congruence climate within nursing units. *Journal of Korean Academy of Nursing*. 2014;44(2):198-208. <https://doi.org/10.4040/jkan.2014.44.2.198>
35. Kim KS, Kim JA, Kwon SH, Song MS. Development of the nursing process based performance measurement tool for medication management and blood transfusion. *Journal of Korean Clinical Nursing Research*. 2010;16(1):177-96. <https://doi.org/10.22650/JKCNr.2010.16.1.177>