Network of Patient Safety Databases

2026 DATA SPOTLIGHT

SI Free-Text Analysis of OTHER Event Type



MAKE AMERICA HEALTHY AGAIN

Patient Safety Listed PSO Organization

Network of Patient Safety Databases
Data Spotlight, 2026

SI Free-Text Analysis of OTHER Event Type

U.S. DEPARTMENT OF
HEALTH AND HUMAN SERVICES
Agency for Healthcare Research and Quality
5600 Fishers Lane
Rockville, MD 20857
https://www.ahrq.gov/

AHRQ Publication No. 26-0025
October 2026
https://www.ahrq.gov/npsd/data/spotlights.html



ACKNOWLEDGMENTS

The Network of Patient Safety Databases Data Spotlight, 2026, is the product of voluntary participation in the Agency for Healthcare Research and Quality (AHRQ) Patient Safety Organization (PSO) program by providers and PSOs nationwide. Many individual providers, hospital facilities and PSOs collaborated to collect and submit the data used in this report to the PSO Privacy Protection Center (PSOPPC). Without the efforts of these dedicated individuals and organizations, AHRQ and the PSOPPC / Network of Patient Safety Databases (NPSD) team would not have been able to produce this report.

Specifically, we thank the Authors from Cormac Corporation: Joanne Campione, Kartik Chaurasiya and Louise Hadden

---

This document is in the public domain and may be used and reprinted without permission. Citation of the source is appreciated. Suggested citation: Campione J, Chaurasiya K, Hadden L. NPSD Data Spotlight: SI Free-text Analysis of OTHER Event Type. (Prepared by Cormac). Rockville, MD: Agency for Healthcare Research and Quality; October 2026. AHRQ Pub. No. 26-0025.

Table of Contents

INTRODUCTION 1
BACKGROUND 1
METHODOLOGY 2
Data Source 2
Data Preparation 3
Free-text Analysis Method 3
Topic Modeling 3
Large Language Model Classification 4
Theme Definitions 5
Key Theme Assignment 9
Theme Overlap Analysis 10
Relationship Between Key Themes And Other Event Characteristics 10
RESULTS 10
Theme Weights From Final LLM Classification 10
Classification Of Key Themes 10
Validation Of Final LLM Classification Results 12
Event Characteristics By Key Theme 13
Report Type 13
Extent of Residual Harm to Patients (Incidents Only) 14
Contributing Factor 15
Location of Event 16
Theme Overlap 17
RECOMMENDATIONS 19
CONCLUSION 20
APPENDIX. Mutually Exclusive Themes Including Theme Overlap 21

INTRODUCTION

This *NPSD Data Spotlight* illustrates the use of super intelligence (SI)-based free-text analysis to identify the most common themes among hospital safety concerns reported as an OTHER Event Type by Patient Safety Organizations (PSOs). It also provides insights into the free-text analysis methods used by the Patient Safety Organization Privacy Protection Center (PSOPPC) to analyze qualitative data. It details the analytic approach, the technical methods involved and the results of the analysis.

The purpose of this SI-based free-text analysis was to better understand the themes and/or topics that PSOs are categorizing as OTHER Event Types.

BACKGROUND

The PSOPPC was created by the Agency for Healthcare Research and Quality (AHRQ) to support the implementation of the Network of Patient Safety Databases (NPSD) as authorized by the Patient Safety and Quality Improvement Act of 2005 (PSQIA). Healthcare providers submit patient safety event data to PSOs, which in turn submit standardized versions of these event reports to the PSOPPC via the Common Formats.[1] AHRQ's Common Formats are a set of standardized definitions and formats that make it possible to collect, aggregate and analyze uniformly structured information about patient safety. The Common Formats also include fields for reporters to enter unstructured free text to provide further detail on events. After aggregation and deidentification of data in the PSOPPC, analyses of the non-identifiable data are published in the NPSD to provide a national-level perspective of patient safety events and their contributing factors.

Events in the hospital setting are reported to the PSOPPC through the AHRQ Common Formats for Event Reporting-Hospital (CFER-H), which collects information for 10 event categories: 1) Anesthesia; 2) Surgery, including other invasive procedures; 3) Blood or Blood Product; 4) Device or Medical/Surgical Supply; 5) Fall; 6) Healthcare-associated Infection; 7) Medication or Other Substance; 8) Perinatal; 9) Pressure Injury/Ulcer and 10) Venous Thromboembolism (VTE).[2] These specific Event Types were created because they were believed to reflect the most common patient safety events occurring in the hospital setting. The last event type response option that a PSO can choose is *OTHER: Please specify*.

While the Common Formats were designed with the intention for reporters to typically choose among these 10 most common Event Types, the most reported Event Type in the NPSD is the OTHER Event Type. More specifically, among 3,087,813 events in the NPSD with report dates during 2016-2025, 59.4% of the events were reported as OTHER.[3]

Therefore, the purpose of this analysis is to explore the free-text data among OTHER Event Type

[1] The AHRQ Common Formats are accessible at https://www.psoppc.org/common-formats.
[2] Event types differ slightly between Common Formats for Event Reporting – Hospital Versions 1.2 and 2.0. These 10 event types appear in at least one of these two versions.
[3] AHRQ Network of Patient Safety Databases 2026 Chartbook (Table 1). Accessible at https://www.ahrq.gov/npsd/data/chartbook/index.html

records to assess themes among these reported events.

Free-text analysis is the scanning of unstructured, qualitative text fields, which allows for the identification and categorization of critical information that is not collected as a structured data element or answer value (i.e., with pre-determined response options). Free-text analysis provides the opportunity to see a more comprehensive picture of an event by reviewing the written text entered in a narrative text field (e.g., an unstructured Common Formats field). Specific to events reported to the PSOPPC, free-text analysis aids in identifying common events, unsafe conditions and contributing factors that may not be immediately obvious in the structured data.

The objectives of this analysis were to 1) detect and define new topics / themes; 2) measure theme prevalence; 3) evaluate the overlap and relationships among the themes; 4) analyze the relationships between themes found in the free text and other factors associated with an event, such as report type (incident, near miss or unsafe condition), contributing factors, harm level, and hospital setting/location; 5) provide recommendations for new Event Types and 6) evaluate the analytic value of the free-text data.

METHODOLOGY

DATA SOURCE

This analysis was based on a sample of patient safety concerns submitted to the PSOPPC using CFER-H V1.2 or CFER-H V2.0. From the PSOPPC, we pulled records with 1) an event report date between July 1, 2024 and June 30, 2025, 2) an Event Type = OTHER and 3) any information in the "description of event" free-text field. This resulted in 134,000 records. We then performed a two-stage sample design to ensure fair representation across both small and large PSOs. **The final sample included 57,234 records across seven PSOs.**

The narrative, free-text used in this analysis were captured from the following fields (also called Data Elements [DE]) in the CFER-H:

- DE15, which guides the user to "Describe the specifics of the incident, near miss, or unsafe condition in your own words, as best you know them." There is a 4,000-character field limit for this DE to be submitted to the PSOPPC.
- DE87, which guides the user to enter "additional event information." This field is captured in CFER-H V1.2 only.
- The "Other, please specify" write-in text field associated with DE21 when the reporter selects OTHER as the Event Type.
- The "Other, please specify" write-in text field associated with DE105 when the reporter selects OTHER as the Contributing Factor.

The above free-text fields were extracted from the PSOPPC database and merged as one text field for each record.

DATA PREPARATION

To prepare and handle the free-text data, the following steps were performed:

1. Data deidentification. This step is for data protection and privacy. Using Amazon Web Services (AWS) Comprehend, the data were stripped of all identifiers and personal information. For example, all names, addresses and dates are replaced with a default "NAME," "ADDRESS" and "DATE," respectively.
2. Text processing. Text processing involves reviewing, cleaning and preparing the free-text fields for analysis.

FREE-TEXT ANALYSIS METHOD

Topic Modeling

1. LDA Technique

We first performed latent Dirichlet allocation (LDA) for topic modeling to identify coarse-grained topics within a corpus. Topic modeling is a general technique used in text analysis to uncover hidden themes or topics within a collection of free text. This first technique found nine key topics: 1) Laboratory / Labeling, 2) Restraints and Residential / Nursing, 3) Radiology / Scans / Imaging, 4) Perinatal, 5) Admission / Discharge, 6) Behavioral / AMA, 7) Blood Glucose Surveillance, 8) Patient Health Record Documentation and Accuracy and 9) IV Therapy.

2. BERTopic Modeling

To identify more semantically coherent and fine-grained topics, we applied BERTopic, an open-source Python library and modular topic-modelling framework that combines transformer-based embeddings with clustering techniques to discover topics from text data.

The BERTopic pipeline used the following open-source components: two transformer-based embedding models — Snowflake/snowflake-arctic-embed-l-v2.0 for general semantic representation and EmbeddingGemma-300M-Medical for medical and clinical terminology along with UMAP (umap-learn) for dimensionality reduction and HDBSCAN (hdbscan) for clustering semantically similar documents into topics.

Both LDA and BERTopic were implemented in Python using open-source libraries. The LDA model was developed using Gensim, while BERTopic was implemented using the BERTopic package. This setup enabled a transparent, reproducible and flexible comparison between traditional probabilistic topic modelling and modern embedding-based topic modelling approaches.

As expected, BERTopic found the same themes as LDA and provided an additional set of themes such as Emergency Department Arrival, Workforce Safety (including security staff),

Environmental / Room Cleanliness, Emergency Codes and Response, Allergies, Blood Pressure, Skin-related / Skin Tissue and Trauma / Injury (not Falls-related).

We further defined each of the themes identified by BERTopic. The outputs from both LDA and BERTopic provided the words most strongly associated with each theme. This output facilitated theme definition.

Large Language Model Classification

We used a large language model (LLM) to classify each record into one or more themes. We chose to use Claude Sonnet 4.5 LLM to classify free-text narratives into predefined themes. Claude Sonnet 4.5 LLM is available on AWS Bedrock, which is FedRAMP High Level-Approved. A FedRAMP-authorized platform offers secure access to LLM models and includes data isolation and encryption.

After the first round of LLM, we performed manual validation of a random subset to assess the accuracy of the LLM. Discrepancies were adjudicated and used to further refine prompts and theme definitions.

More specifically, between round one and round two of LLM, we used SI-assisted research and human expertise of hospital patient safety events and concerns to refine the final list of themes (including combining themes, splitting themes, adding themes or deleting themes). We went from 17 themes down to 15 themes.

We split the Behavioral Health Theme by separating Against Medical Advice (AMA) from Mental Health Disorders and/or Substance Abuse (non-violent) because AMA may not always be behavioral-health related. For example, AMA discharge may be caused by dissatisfaction, financial concerns or feeling better prematurely. A theme we had labeled as “Patient Security / Restraints” was dissolved and parts of that theme’s definition were moved to either Workforce Safety or Patient Unsafe.

We dropped Allergies, Blood Pressure, Environmental / Room Cleanliness and Blood Glucose Surveillance because less than 2% of the total records were classified to these themes in round one LLM. (Note that a hypertensive emergency would fall under the Emergency Codes and Response theme.) Lastly, because Pressure Injury and Fall Event Types in the CFER-H only accept incident report types, we added Pressure Injury and Fall-related as themes to investigate if PSOs were using OTHER Event Type to report Pressure Injuries and Falls that were near misses and unsafe conditions. Again, we further enhanced and refined the themes’ definitions. Theme definition refinement utilized references and was facilitated by human review of the free text.

Table 1. List of the Final Themes used for Analysis

Theme Name (Long)	Theme Name (Short)
Patient Health Record Documentation	Health Records
Admission / Emergency Department Arrival / Discharge (not AMA)	Admit / Discharge

Theme Name (Long)	Theme Name (Short)
Laboratory Specimen Handling and Result Communication	Laboratory
Patient Discharge Against Medical Advice (AMA)	AMA
Intravenous (IV) Therapy	IV Therapy
Workforce Safety	Workforce Safety
Radiology / Scans / Imaging	Radiology
Patient Unsafe	Patient Unsafe
Emergency Codes and Response	Emergency Response
Trauma / Injury	Trauma / Injury
Skin-related / Skin Tissue (non-Pressure Injury)	Skin (non-P.I.)
Labor or Delivery	Labor / Delivery
Fall-Related	Fall-related
Mental Health Disorders and/or Substance Abuse (non-violent)	MH SubAbuse
Pressure Injury	Pressure Injury

THEME DEFINITIONS

Table 2. Final Theme Definitions for Large Language Model (LLM) Classification

Theme Short Name	Definition (and *PROMPT* if applicable)
Health Records	This topic captures patient safety event reports concerning the accuracy, completeness and integrity of patient health record documentation, including issues such as incorrect, incomplete or mis-assigned records and metadata (e.g., medical record number [MRN], overlay errors or patient identification mismatches). It includes documentation-related discrepancies such as incorrect patient identifiers, errors in MRN assignment and inconsistencies in clinical records that may affect patient identification and continuity of care. The topic also captures issues related to documentation of informed consent for procedures, including missing, incomplete or improperly recorded consent forms and associated patient identification details. These events reflect breakdowns in documentation practices and health record management that can contribute to clinical errors, delays in care and risks to patient safety due to inaccurate or unavailable information.
Admit / Discharge	This topic encompasses patient safety event reports documenting key transitions in a patient's care journey, including arrival to the emergency department, triage, hospital admission, inter-unit transfers and discharge processes. It includes clinical workflow stages such as triage assessment — where patients are prioritized based on acuity — as well as admission decisions, transfers between care levels (e.g., ED to inpatient or progressive care units) and discharge planning activities. The topic reflects documentation of care transitions, including timing elements (e.g., timestamps, sequencing of events) and coordination across teams as patients move through emergency, inpatient and transitional care settings.

Theme Short Name	Definition (and *PROMPT* if applicable)
	These events occur during periods of increased vulnerability for patients, particularly during handoffs and discharge processes, where gaps in communication or coordination may impact safety and quality of care. *Prompt: This category excludes cases where patients leave against medical advice (AMA) or without being seen.*
Laboratory	This topic reflects hospital event reports concerning problems with laboratory specimen collection, labeling and test result communication, including issues where alert messages — often sent via clinical communication systems — are not properly logged, assigned, or received by the appropriate staff. It highlights coordination breakdowns between clinical teams and laboratory workflows that can affect timely processing of blood samples and test results.
AMA	This topic captures patient safety event reports involving patients who leave care against medical advice (AMA), attempt to leave, discharge themselves or elope from care settings prior to completion of recommended treatment. The topic captures events involving patient refusal behaviors (e.g., “patient refused,” “attempting to leave," "flee”), elopement risk, and unplanned discharge scenarios, as well as outpatient-related events such as no-shows and cancellations (NOS CX) where disengagement from care is documented. These events reflect high-risk periods in care where patient autonomy, adherence challenges and care transitions intersect, particularly during discharge processes, which are associated with increased risk of adverse outcomes such as readmission or incomplete treatment.
IV Therapy	This topic captures patient safety event reports involving intravenous (IV) therapy and infusion processes, including the placement and management of vascular access for administering fluids, medications, blood products or contrast agents. It includes clinical events related to IV access and infusion practices, such as peripheral intravenous (PIV) lines, central lines and central intravenous catheters (CICC), as well as the administration and monitoring of infusions. The topic captures complications associated with IV therapy, including infiltration, edema and infusion-related issues, as well as events where clinicians document concerns with IV site condition, medication delivery or fluid administration. It also includes narratives describing IV-related care processes such as establishing access, monitoring infusion status and managing patient responses to IV therapy, including documentation of terms such as “infiltration,” “edema,” “infusion,” “patient refused” or complications at the vascular access site. These events reflect risks associated with IV therapy as part of medication and treatment delivery workflows, excluding device malfunctions or equipment-related failures.
Workforce Safety	This topic captures patient safety event reports involving workplace violence and safety risks directed toward healthcare staff and security personnel, including physical assault, sexual assault, verbal threats, harassment and other aggressive behaviors from patients, visitors, family members or other staff. It includes incidents and cases involving combative or violent individuals requiring security intervention (e.g., “Code Gray”), as well as the use of restraints, police involvement, contraband searches, weapons detection systems and restricted-access measures implemented to maintain safety within the care environment. The topic also captures events involving workplace safety hazards and unsafe conditions affecting staff, such as faulty equipment, poor ergonomics or environmental risks that may

Theme Short Name	Definition (and *PROMPT* if applicable)
	result in injury or workers' compensation claims. These events reflect risks to workforce safety in healthcare settings, where staff are at increased risk of violence and injury compared to other industries, particularly in high-stress environments such as emergency departments and behavioral health units. *Prompt: This category excludes events where patients report feeling unsafe (patient-perceived safety concerns), which are classified separately.*
Radiology	This topic pertains to patient safety event reports involving diagnostic imaging processes such as CT and MRI scans, including ordering, performance, interpretation, documentation and communication of imaging results as part of the clinical diagnostic workflow. It captures a range of imaging-related events described in clinical workflows, including routine imaging procedures, documentation of radiology exams and interactions between technologists, radiologists and clinical teams in the delivery of diagnostic care. Within these events, reports may also include issues such as incorrect imaging orders, variability in interpretation or breakdowns in communication of findings, which are recognized contributors to diagnostic error and delays in care. These events reflect the complexity of radiology workflows, where multiple steps, handoffs and communication points between providers influence the accuracy and timeliness of diagnosis and patient management.
Patient Unsafe	This topic captures patient safety event reports involving interventions and care processes intended to protect patients from harm, including the use of restraints when a patient poses an immediate risk to themselves (e.g., self-harm or unsafe behavior). It includes events involving suicidal ideation, suicide attempts, self-harm behaviors and other situations where patients are unable to maintain their own safety and require monitoring, observation or protective interventions. The topic captures safety-related monitoring and interventions such as suicide precautions, high-risk observation (e.g., Lighthouse precautions) and behavioral health assessments conducted to evaluate risk to self during hospitalization. It also captures incidents involving inappropriate or unsafe care delivered to patients by healthcare staff, including verbal abuse, physical or sexual misconduct, neglect or failure to follow established care processes or procedures (e.g., staff described as "lazy" or not adhering to care protocols). It further includes documentation of care quality concerns such as incomplete or missed care, deviations from expected clinical practices and patient safety risks arising from inadequate monitoring or failure to adhere to established standards of care. These events reflect risks to patient safety arising from both patient-related behaviors requiring protective intervention and provider-related factors affecting quality and safety of care delivery. *Prompt: This category excludes cases where restraints are used primarily to protect staff or others (classified under Workforce Safety), as well as events where patients report feeling unsafe without associated care or safety intervention.*
Emergency Response	This topic captures patient safety event reports involving activation of emergency response systems, including hospital emergency codes (e.g., Code Blue, Code Stroke) and Rapid Response Team (RRT) activations, in response to acute patient deterioration or time-sensitive medical emergencies. It includes clinical scenarios such as cardiopulmonary arrest requiring resuscitation (e.g., CPR),

Theme Short Name	Definition (and *PROMPT* if applicable)
	stroke alerts requiring rapid evaluation and intervention and other emergent conditions requiring immediate multidisciplinary response. The topic reflects coordinated workflows where emergency teams — including code teams and RRTs — are mobilized to perform rapid assessment, stabilization and life-saving interventions in high-acuity situations. These events emphasize structured emergency response processes designed to improve patient outcomes by reducing delays in treatment and preventing progression to critical events such as cardiac arrest.
Trauma / Injury	This topic captures patient safety event reports involving trauma or injury cases requiring clinical evaluation, documentation or trauma team activation, excluding falls-related events. It includes events where patients present with injuries requiring structured trauma assessment processes, including evaluation of injury severity, application of trauma criteria and activation of trauma response workflows in emergency or acute care settings. The topic reflects documentation of trauma-related clinical workflows, such as trauma team evaluations, injury classification and coordination of care during emergency response to physical injury. These events emphasize the structured and time-sensitive nature of trauma care, where rapid assessment and clinical decision-making are critical to managing injury severity and improving patient outcomes.
Skin (non-P.I.)	This topic captures patient safety event reports involving skin integrity issues and tissue damage, including skin tears, wounds and related assessment and treatment documentation, excluding pressure injuries or pressure ulcers. It includes clinical events where healthcare staff document skin conditions such as tears, wounds or other tissue damage, along with details such as wound size, location, tissue condition and progression over time. The topic also captures care processes related to wound management, including dressing application, monitoring of skin integrity and evaluation of healing or deterioration in non-pressure-related skin injuries. These events reflect documentation and clinical management of skin and soft tissue conditions that may impact patient comfort, infection risk and overall recovery, outside of pressure-related injury classifications.
Labor / Delivery	This topic captures patient safety event reports related to labor, delivery and the immediate perinatal period, including care processes involving both the mother and newborn. It includes clinical events associated with labor and delivery processes, newborn assessment, delivery outcomes and early neonatal care, including cases requiring neonatal intensive care unit (NICU) involvement. The topic reflects documentation of interactions between obstetric, neonatal and nursing teams, as well as care transitions occurring immediately after birth as part of perinatal clinical workflows. Within these events, reports may also describe complications or safety concerns such as maternal morbidity, obstetric hemorrhage or neonatal complications, which are recognized contributors to adverse outcomes in childbirth.

Theme Short Name	Definition (and *PROMPT* if applicable)
Fall-related	A fall event is defined as a sudden, unintended descent of a patient's body to the ground or another surface (e.g., bed, chair or bedside mat), which may be assisted or unassisted This topic captures patient safety event reports involving falls that occur in healthcare settings, including both assisted and unassisted falls, regardless of whether injury occurs Patient outcomes associated with falls may include injuries such as intracranial injury, fractures, dislocations, lacerations requiring sutures, skin tears, avulsions, hematomas or significant bruising Falls may occur during routine care activities such as ambulation, transfers, toileting or movement within the care environment and are considered preventable in many cases through appropriate risk assessment and safety interventions.
MH / SubAbuse	This topic includes a broad range of psychiatric and behavioral health conditions including, but not limited to, schizophrenia, hallucinations, post-traumatic stress disorder (PTSD), obsessive-compulsive disorder (OCD), bipolar disease, bipolar disorder (BPD), attention deficit hyperactivity disorder (ADHD), mood disorders, eating disorders (anorexia, bulimia, binge eating), panic attack, anxiety disorders, phobias, autism spectrum disorders, substance use disorders (SUD), alcoholism and drug abuse. This also includes related mental health treatment, substance use disorder treatment, unintentional overdose and clinical institute withdrawal assessments for alcohol (CIWA). Patients can exhibit signs of agitation, psychological distress, cognitive distress, emotional distress or psychiatric symptoms. This category encompasses a wide spectrum, from general emotional instability and mood disturbances to specific conditions (e.g., depression, anxiety or behavioral dysregulation).
Pressure Injury	A pressure injury (also referred to as a pressure ulcer) is localized damage to the skin and underlying soft tissue, typically occurring over a bony prominence or as a result of pressure or pressure in combination with shear. A pressure injury event includes cases where a Stage 2 or higher pressure injury (i.e., Stage 2, 3, 4, unstageable, deep tissue pressure injury or mucosal membrane pressure injury) develops during hospitalization (i.e., was not present on admission) or worsens from a lower stage that was present on admission. These events reflect patient safety concerns associated with hospital-acquired conditions, where inadequate prevention, monitoring or care processes contribute to skin breakdown and tissue injury during inpatient care. Pressure injuries are considered largely preventable adverse events in many cases and are used as indicators of patient safety and quality of care in hospital settings.

The second and final round of LLM involved the classification of these 15 themes.

KEY THEME ASSIGNMENT

A Key Theme was assigned when a theme received a LLM classification weight greater than 50% for a record. Thus, that key theme received a majority weight. Note that any record without a greater than 50% weight did not receive a key theme. Key Themes exclude records with a 50%-50% split between two themes.

THEME OVERLAP ANALYSIS

When there was not a key theme, we looked for whether two or more themes overlapped to have a combined rate of 60% or more. For this purpose, we analyzed the weights to assign each combination a unique theme name. For example, if one theme received 50% weight and a second theme received 50% weight, the theme was named "Theme 1 / Theme 2." Next, we reviewed themes with a 40%-40% weighting structure and then a 40%-35% structure, and so on. The higher weighted theme was listed first in the paired / overlapping theme name. This iterative approach allowed for the quantification of mutually exclusive theme names and for an analysis of theme overlap.

RELATIONSHIP BETWEEN KEY THEMES AND OTHER EVENT CHARACTERISTICS

We analyzed the relationships between key themes identified through the SI free-text analysis and other factors collected in the CFER-H's structured fields, such as report type (incident, near miss or unsafe condition), contributing factor, harm level and hospital location.

RESULTS

THEME WEIGHTS FROM FINAL LLM CLASSIFICATION

Among the total sample of 57,234 records, 52.8% (n=30,193) of the records received any weight greater than 0% for only one theme, 32% (n=18,334) for two themes, 4% (n=2,287) for three themes, 0.1% (n=29) for four themes, and two records received weights for five themes. Eleven percent of the total sample had no weights assigned to a theme.

CLASSIFICATION OF KEY THEMES

Eighty-six percent of the sample (49,359 / 57,234) had a Key Theme Classification and 14% (7,875 / 57,234) did not have a Key Theme.

Figure 1 displays each Key Theme's percentage of the 49,359 records.

Figure 1. Distribution of Themes among Records with a Key Theme Classification

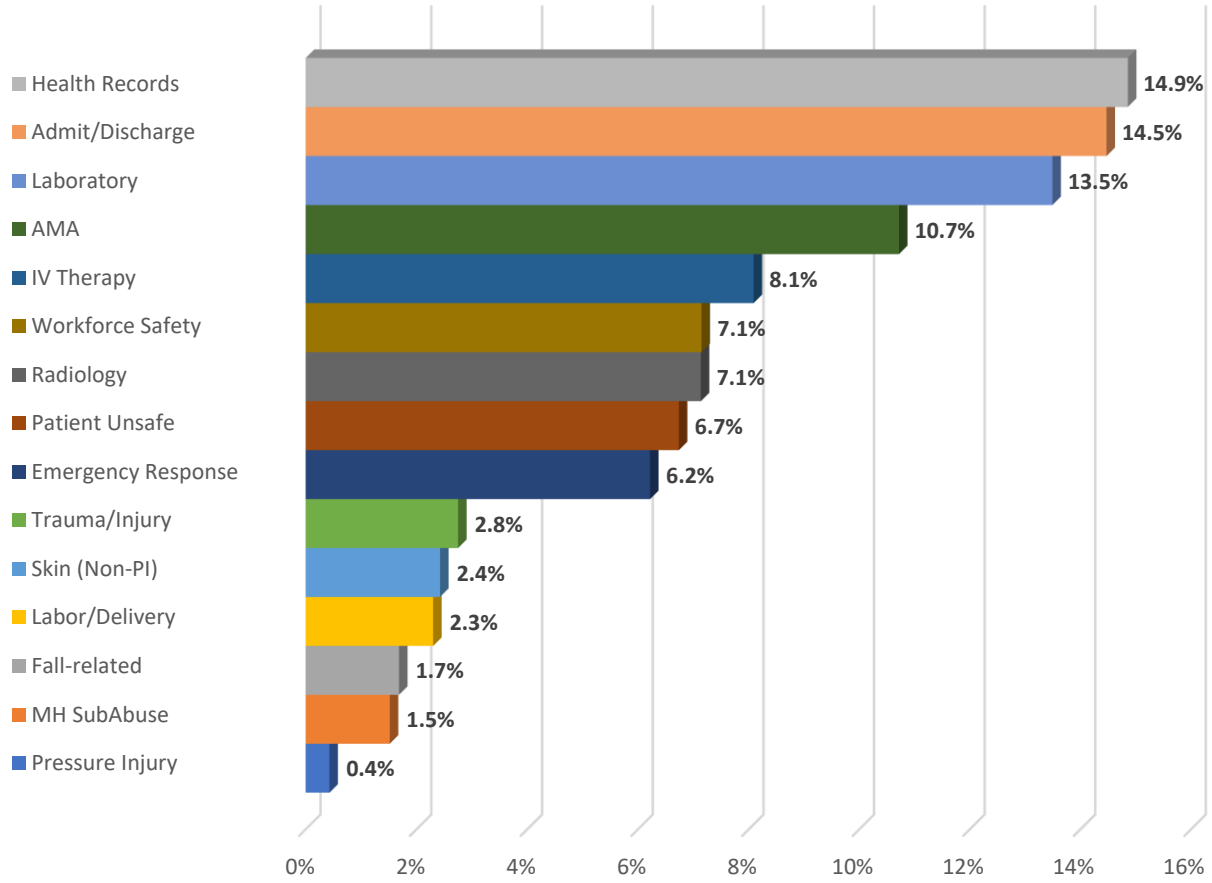


Table 3 shows the number and percentage of records classified as a Key Theme. Among those with a Key Theme Classification, the top themes were 1) Health Records, 2) Admit / Discharge, 3) Laboratory, 4) AMA and 5) IV Therapy.

Table 3. Key Theme Volume across Key Theme Classified Records

Key Theme	# of Records as Key Theme	% of Total
Health Records	7,335	14.9%
Admit/Discharge	7,144	14.5%
Laboratory	6,661	13.5%
AMA	5,291	10.7%
IV Therapy	3,993	8.1%
Workforce Safety	3,529	7.1%
Radiology	3,522	7.1%
Patient Unsafe	3,327	6.7%
Emergency Response	3,070	6.2%
Trauma/Injury	1,359	2.8%
Skin (Non-P.I.)	1,198	2.4%

Key Theme	# of Records as Key Theme	% of Total
Labor/Delivery	1,136	2.3%
Fall-related	832	1.7%
MH SubAbuse	750	1.5%
Pressure Injury	212	0.4%
Total Records with Key Theme	**49,359**	**100%**

VALIDATION OF FINAL LLM CLASSIFICATION RESULTS

Per record, the LLM gave each theme a weight representing the strength of the classification. Weights ranged from 0% (no detection of the theme) to 100% (theme is detected with no other themes overlapping).

To validate the accuracy of the classifications, we performed validation using four different approaches on a random sample of 400 records, which were grouped into four files of 100 records each. For this random selection process, we set the control variables as 1) masked PSO identifier and 2) binary indicator variables (0/1) for each theme. We then used PROC SURVEYSELECT that performs hierarchic serpentine sorting, which alternates between ascending and descending sorting throughout all levels of each control variable. This ensures we have examples of every theme and every PSO in the total sample.

Approach #1: Stability Score

This approach measures the percentage of runs (out of three) where Claude predicted the same top label / theme. For example, for each run on the text "Patient fell down while a nurse was drawing blood for a laboratory specimen" Claude outputted "Fall-related." Because Claude consistently chose the same theme in all three runs, the score / metric is 100% stable.

Approach #2: Multi-Judge Agreement (Argmax)

This methodology quantifies the rate of agreement between Claude and Mistral-Large-3-675B-Instruct (Mistral Large 3) on each record's highest-ranked primary classification. Mistral Large 3 is a different LLM available on Amazon Bedrock. For example, if Claude had 55% Fall-related and Mistral had 60% Fall-related, they had 100% agreement on primary classification.

Approach #3: Multi-Judge Agreement

This approach measures whether Claude and Mistral Large 3 assign the same highest-ranked (primary) classification to each record. For example, for a given record, Claude may assign Fall-related the highest score at 55%, while Mistral assigns Fall-related the highest score at 60%. Although the scores differ, both models select Fall-related as the primary classification; therefore, they have 100% agreement on the record's primary classification.

Approach #4: Manual / Human Review

With this approach, researchers read the text, reviewed the LLM results and agreed or disagreed on the LLM's classification to one or more themes. Note that this approach was also performed after round one of LLM.

Table 2 below shows the results of the validation process. There was high agreement across all four validation approaches with the LLM's categorization results. The LLM's ability to interpret textual meaning and context enabled the defined thematic categories to be applied consistently at scale.

Table 2. Validation Results

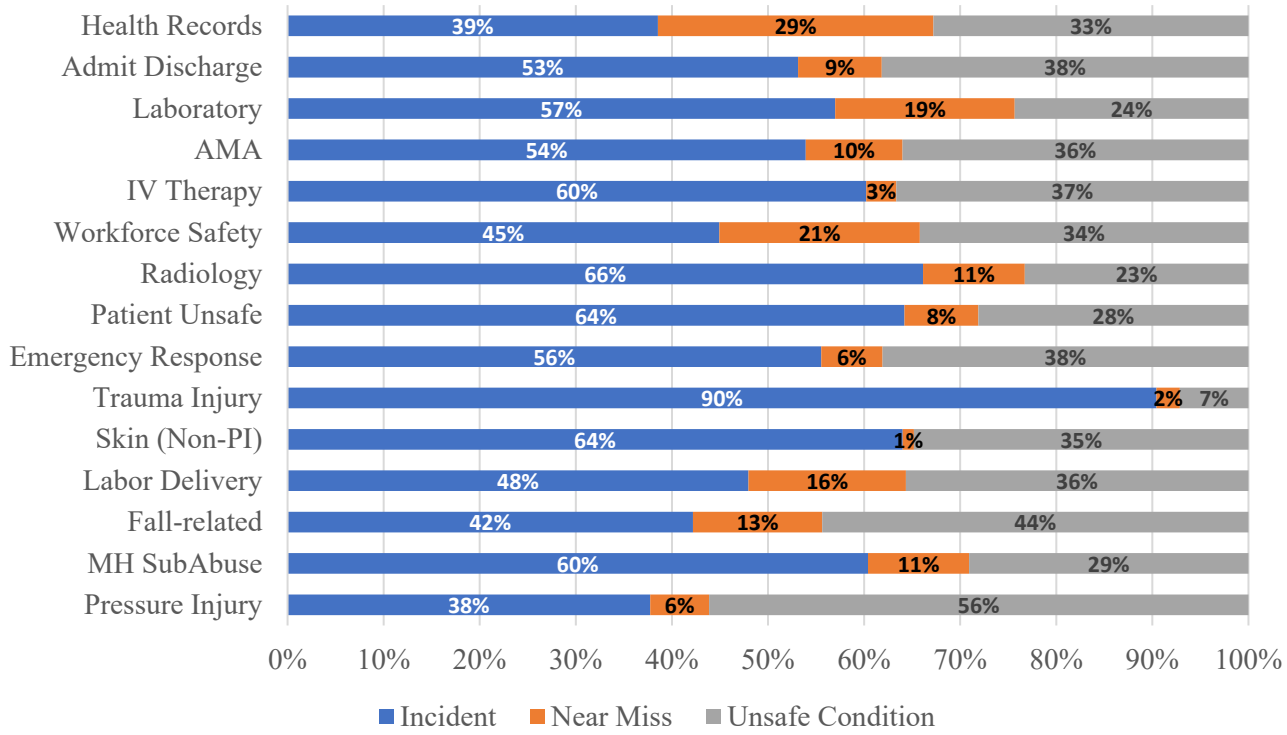
Validation Group	Total Records	Stability Score 3x	Multi-Judge Agreement (Argmax)	Multi-Judge Agreement	Human Agreement %
File 1	100	98.7%	82.0%	94.0%	95%
File 2	100	99.0%	72.0%	81.0%	94%
File 3	100	97.0%	85.0%	89.0%	97%
File 4	100	97.7%	85.0%	96.0%	94%
OVERALL (Macro-Average)	**400**	**98.1%**	**81.0%**	**90.0%**	**95%**

EVENT CHARACTERISTICS BY KEY THEME

Report Type

The figure below shows the distribution of event report type (incident, near miss, or unsafe condition) within each theme. The data sample size for Figure 2 is 49,359. There are no missing responses for Report Type.

Figure 2. Distribution of Report Type within Key Theme



Among the Key Theme classified sample (49,359), the distribution of report type was 54.4% Incident, 13.4% Near Miss and 32.2% Unsafe Condition.

Incident was highest among Trauma Injury and lowest for Pressure Injury. Near Miss was highest among Health Records and lowest for Skin (non-P.I.). Unsafe Condition was highest among Pressure Injury and lowest for Trauma Injury key theme.

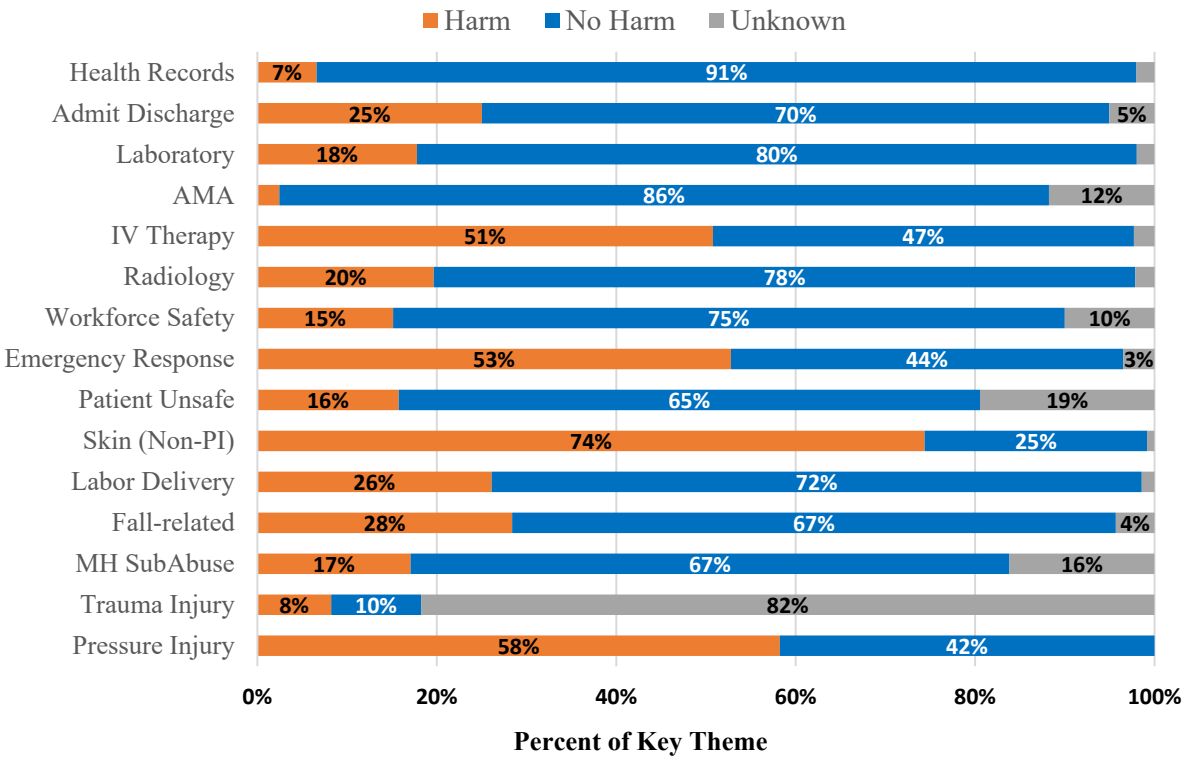
Pressure Injury and Falls are Event Types that PSOs can select in the Common Formats but for Incident Events only. The prevalence of Pressure Injury and the prevalence of Falls-related as key themes were quite low (n=212 and n=832, respectively). However, as suspected, PSOs reported some Pressure Injury and Falls events as an Other Event because they wanted to report an Unsafe Condition or Near Miss.

Extent of Residual Harm to Patients (Incidents Only)

Patient Residual Harm is collected in the CFER-H for only Incident events with the question "After any intervention to reduce harm, what was the degree of residual harm to the patient from the incident (and subsequent intervention(s))?" Here, due to low frequency, the answer values of "Death," "Severe Harm," "Moderate Harm" and "Mild Harm" are combined as "Harm."

The figure below shows distribution of Patient Residual harm among incident events across the key themes. The data sample size for the figure is 25,765. Incident reports with missing reports of Harm are excluded (n=1,093).

Figure 3. Distribution of the Extent of Patient Residual Harm within a Key Theme*



** The figure excludes the labels for percentages less than 3%. For Labor Delivery, the Harm response is for the mother.*

Among all Incident events in the Key Theme classified sample with a response to the harm question, 22% had Patient Residual Harm, 68% No Harm and 9% Unknown Harm.

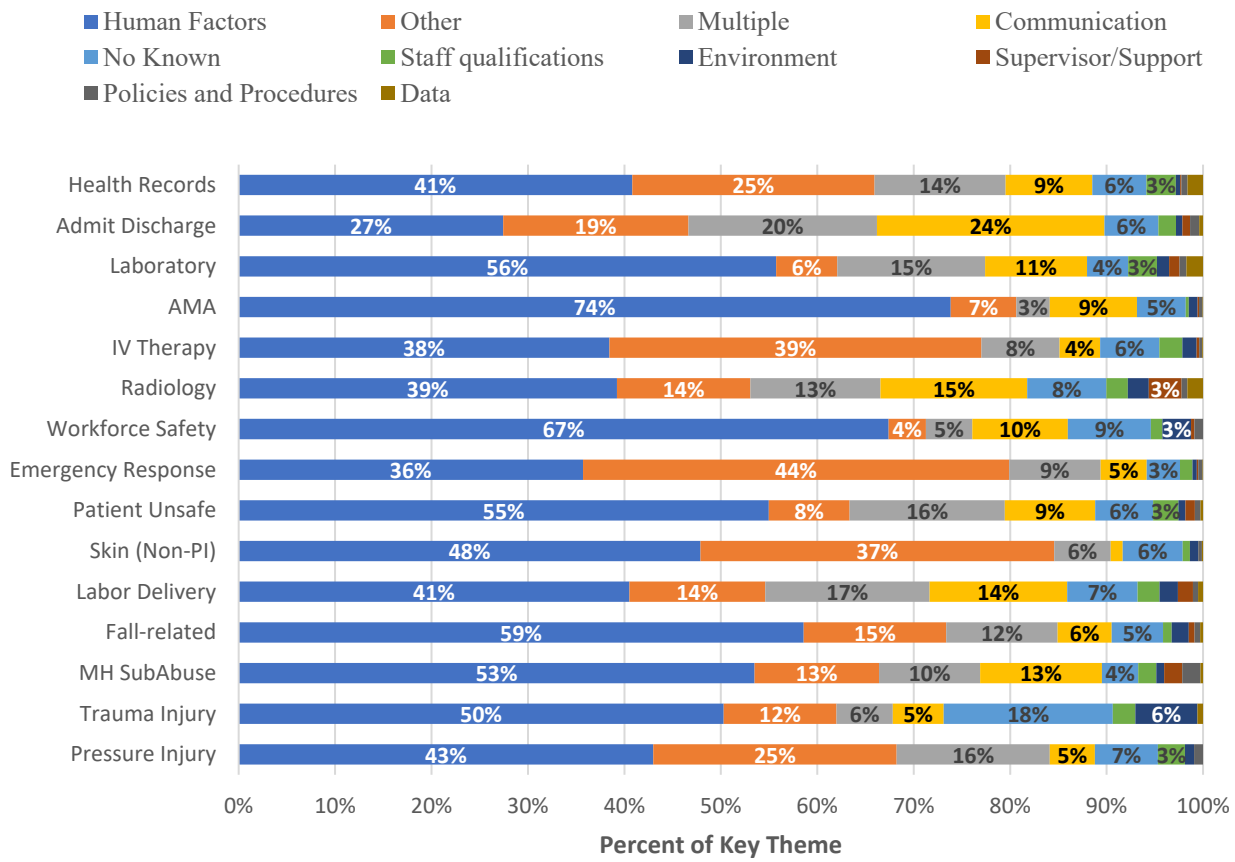
The rate of Patient Residual Harm was highest for Skin (non-P.I.) and lowest for AMA. Unknown harm was very high for Trauma Injury.

Contributing Factor

Contributing Factor is collected in the CFER-H with the question "What factor(s) contributed to the event?"

The figure below shows the distribution of contributing factors across the key themes. The data sample size for the below graph is 25,297. Records with missing reports of Contributing Factor are excluded (n=24,062).

Figure 4. Distribution of Contributing Factor within a Key Theme*



** The figure excludes the labels for percentages less than 3%.*

On average, among the Key Theme classified sample, the most common Contributing Factors were Human Factors (46%), Other (19%), Communication (11%) and Multiple Factors (12%).

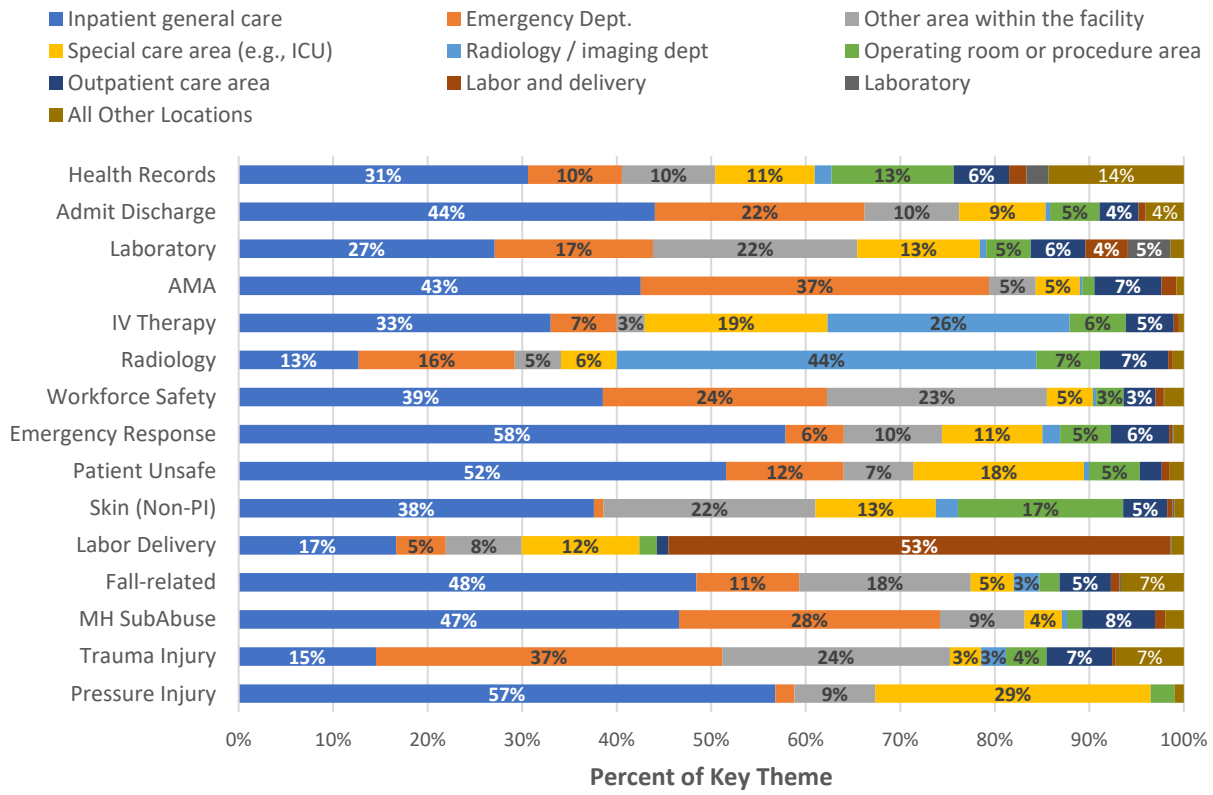
Human Factors had the highest representation among AMA and Workforce Safety key themes. Communication was highest among Admit / Discharge.

Location of Event

Location is collected in the CFER-H with the question "Where did the event or unsafe condition occur?" The figure below shows the response option locations and the distribution of those locations across the key themes.

The data sample size for the above graph is 39,708. Records with missing reports of Contributing Factor are excluded (n=9,651).

Figure 5. The Distribution of Event Location within Key Theme*



** The figure excludes the labels for percentages less than 3%.*

On average, among the Key Theme classified sample, the most common locations were Inpatient General Care (37%), Emergency Department (16%), Other Area within the Facility (12%) and Special Care Area (e.g., ICU, NICU) (11%).

Health Record Theme events took place across all locations, with a third in the inpatient general care area. Admit / Discharge, AMA, and MHSubAbuse were mainly in the inpatient general care area and in the emergency department. As expected, IV Therapy and Radiology had the highest percentage of events that took place in the radiology / imaging department.

THEME OVERLAP

The full results of the theme overlap analysis and the assignment of mutually exclusive paired / overlapping theme groups can be found in the Appendix. After including records with two or more themes assigned with a combined rate of 60% or more, the total number of records with any theme classification was 50,739, representing 88.6% of the full sample, as shown in Table 4 below.

Table 4. Summary of LLM Classifications including Paired / Overlapping Themes

	# of Records	% of Total
Any Theme Classification*	50,739	88.6%
• *Key Theme Classified*	*49,359*	*86.2%*
• *Classified only as a Paired / Overlapping Theme*	*1,380*	*2.4%*
No Theme Classification	6,498	11.4%
Total Sample	**57,234**	

** The sum of Key Theme Classified and Classified only as a Paired / Overlapping Theme*

Compared with Key Theme classification, the number of records with theme overlap was low (49,359 vs 1,380 records). This limited overlap suggests that the LLM received clear, distinct theme definitions.

The most frequently observed theme overlaps were:

- Admit / Discharge with Health Records,
- Admit / Discharge with Emergency Response,
- Laboratory with Health Records
- Radiology with IV Therapy.

The details of the top five themes are provided below. Note that this detail *includes* the assignment of the theme as part of a paired / overlapping theme.

1. Admission / Emergency Department Arrival / Discharge (not AMA). Admit / Discharge was classified as a theme for 7,616 records representing 15.0% of the 50,739 records with any theme classification. The most common overlapping themes were Emergency Response, Health Records and Radiology. There were 7,144 records with Admit / Discharge as the Key Theme (93.8% of 7,616).
2. Patient Health Record Documentation. Health Records was classified as a theme for 7,583 records representing 14.9% of the 50,739 records with any theme classification. The most common overlapping themes were Admit / Discharge, Laboratory and Radiology. There were 7,335 records with Health Records as the Key Theme (96.7% of 7,583).
3. Laboratory Specimen Handling and Result Communication. Laboratory was classified as a theme for 6,706 records representing 13.2% of the 50,739 records with any theme classification. There were 6,661 records with Laboratory as the Key Theme (99.3% of 6,706).
4. Patient Discharge Against Medical Advice (AMA). AMA was classified as a theme for 5,351 records representing 10.5% of the 50,739 records with any theme classification. The most common overlapping theme was Mental Health Substance Abuse. There were 5,291 records with AMA as the Key Theme (98.9% of 5,351).
5. Intravenous (IV) Therapy. IV Therapy was classified as a theme for 4,028 records representing 7.9% of the 50,739 records with any theme classification. The most common overlapping theme was Laboratory. There were 3,993 records with IV Therapy as the Key Theme (99.1% of 4,028).

RECOMMENDATIONS

The most common themes found in the free text of OTHER Event Types were: 1) Admission / Emergency Department Arrival / Discharge (not AMA), 2) Patient Health Record Documentation, 3) Laboratory Specimen Handling and Result Communication, 4) Patient Discharge Against Medical Advice (AMA), 5) IV Therapy, 6) Workforce Safety, 7) Radiology / Scans / Imaging, 8) Patient Unsafe and 9) Emergency Codes and Response. Each of these nine themes had over 5% prevalence among the 49,359 records with a key theme.

Contributing Factors

The top two themes (Health Records and Admit / Discharge) are part of every hospitalization and are common hospital processes. These themes may reflect process breakdowns that might be better suited as contributing factors rather than Event Types.

Interestingly, as shown in Figure 2, over half of the records with a key theme of Admit / Discharge were incident report types and, as shown in Figure 3, 25% of the incident events with the Admit / Discharge theme included residual patient harm. Therefore, more research is needed to understand whether breakdowns during transitions of care reach the patient and what exactly the adverse outcome is to the patient.

In the theme Patient Unsafe, the definition includes: "It also captures incidents involving inappropriate or unsafe care delivered to patients by healthcare staff, including verbal abuse, physical or sexual misconduct, neglect or failure to follow established care processes or procedures. It further includes documentation of care quality concerns such as incomplete or missed care, deviations from expected clinical practices and patient safety risks arising from inadequate monitoring or failure to adhere to established standards of care." This language was added to the definition of Patient Unsafe because during manual validation we noticed that many of the cases included text about a staff member failing to follow standards of care.

Therefore, we recommend the addition of these three contributing factors to the CFER-H:

1. Health Information that includes breakdowns in patient health record documentation management (including EHR-associated and other Information Technology errors) and patient consent documentation.
2. Team Factors that include breakdowns in communication, planning and coordination associated with patient care transition including admissions, discharges and handoffs / transfers.
3. Staff member failure to adhere to established care processes or procedures (as a new sub-factor under the current Human Factors contributing factor category).

Event Types

Several of the top themes found in this analysis mirror current CFER-H Event Types. For example, the definition of the current Event Type titled Medication or Other Substance is "an event or unsafe condition involving the processing and/or administration of medications, biological products, nutritional substances, medical gases, contrast media or radiopharmaceuticals" and the definition of current Event Type titled Blood or Blood Product is "an event or unsafe condition involving the processing and/or administration of blood or a blood

product." We see similar descriptions in the definitions for Laboratory, IV Therapy and Radiology, and Emergency Codes and Response.

Workforce Safety was found to be the sixth most common theme under OTHER Event Type. Thus, we recommend that PSOs report healthcare workforce events and unsafe conditions to identify trends and to provide insight into how to improve workforce safety.

A previous free-text analysis estimated that 16.5% of all hospital-based patient safety concerns (all Event Types) include a behavioral health-related issue.[4] Thus, we were not surprised that this analysis detected the themes of AMA, Patient Unsafe and Mental Health Disorders and/or Substance Abuse.

Therefore, we recommend that the following themes be considered as new CFER-H Event Types:

1. Laboratory Specimen Handling and Result Communication
2. Intravenous (IV) Therapy Related
3. Radiology / Scans / Imaging Related
4. Behavioral Health —patient unsafe due to the behavior of the patient
5. Workforce Safety — an event or unsafe condition for a staff member(s)
6. Emergency Codes and Response Related

Other

Lastly, the results suggest that CFER-H should allow Falls and Pressure Injury Event Types to accept more than only incident report types.

CONCLUSION

Our analysis searched for 15 themes in the free-text associated with CFER-H OTHER Event Type. Using LLM, 86.2% of the sample (49,359 / 57,234 total) was classified to a Key Theme (weight > 50%) including 30,193 (53% of total) classified as having only one theme (weight = 100%). Eleven percent had no theme classification (weight = 0% for all 15 themes) and 2.4% of the sample was classified as paired / overlapping themes. Although the LLM was consistent, it still had the flexibility to assign weights below 50% to more than one theme when appropriate.

This *NPSD Data Spotlight* demonstrates the effectiveness of using SI-based free-text analysis methods to identify and categorize OTHER Event Types. Furthermore, this analysis illustrates how free-text fields complement and enhance the structured elements of Common Formats by providing critical context and supporting a more complete understanding of patient safety events. As Common Formats increasingly leverage free-text data and text-based analytic methods, guidance should encourage reporters to enter information that clearly addresses the intended purpose of each field, including descriptions of the event, patient outcomes and contributing factors. This approach has the potential to increase the collection of richer, higher-quality data that strengthens the Network of Patient Safety Databases.

[4] Source: 2025 NPSD Data Spotlight: Behavioral Health AI Free-text Analysis. Accessible at https://www.ahrq.gov/sites/default/files/wysiwyg/npsd/data/spotlights/npsd-spotlight_behavioral-health.pdf

APPENDIX. MUTUALLY EXCLUSIVE THEMES INCLUDING THEME OVERLAP

Table 1 below lists the total LLM classifications (n = 50,739) *including* paired / overlapping themes (with a combined rate of 60% or more).

The values in the last column are the percentage of each Theme in which the classification resulted in a Key Theme (i.e., weight > 50%). On average, Key Theme assignments represent 97.7% of all the theme classifications.

NOTE: Frequencies < 10 and their respective percentages are suppressed.

Table 1. Theme Frequency including Overlap

Mutually Exclusive Theme Name	Key Theme	Freq	% of Total Sample	Key as % of Theme
AMA	AMA	5291	9.24	98.9%
AMA / Emerg Response	None	*	*	
AMA / Fall-related	None	*	*	
AMA / IV Therapy	None	*	*	
AMA / Labor Delivery	None	*	*	
AMA / Laboratory	None	*	*	
AMA / MH SubAbuse	None	28	0.05	
AMA / Trauma Injury	None	*	*	
AMA and THREE Other	None	*	*	
AMA and TWO Other	None	12	0.02	
Admit Discharge	Admit Discharge	7144	12.48	93.9%
Admit Discharge / AMA	None	*	*	
Admit Discharge / Emerg Response	None	91	0.16	
Admit Discharge / Fall-related	None	*	*	
Admit Discharge / Health Records	None	63	0.11	
Admit Discharge / IV Therapy	None	28	0.05	
Admit Discharge / Labor Delivery	None	15	0.03	
Admit Discharge / Laboratory	None	51	0.09	
Admit Discharge / MH SubAbuse	None	29	0.05	
Admit Discharge / Patient Unsafe	None	37	0.06	
Admit Discharge / Pressure Injury	None	*	*	
Admit Discharge / Radiology	None	59	0.10	
Admit Discharge / Skin (non-P.I.)	None	*	*	
Admit Discharge / Trauma Injury	None	13	0.02	
Admit Discharge / Workforce Safety	None	*	*	
Admit Discharge and THREE Other	None	*	*	
Admit Discharge and TWO Other	None	44	0.08	
Emergency Response	Emergency Response	3070	5.36	97.9%
Emerg Response / Fall-related	None	*	*	

Mutually Exclusive Theme Name	Key Theme	Freq	% of Total Sample	Key as % of Theme
Emerg Response / IV Therapy	None	*	*	
Emerg Response / Labor Delivery	None	*	*	
Emerg Response / Laboratory	None	13	0.02	
Emerg Response / MH SubAbuse	None	*	*	
Emerg Response / Trauma Injury	None	*	*	
Emerg Response / Admit Discharge	None	*	*	
Emerg Response / Patient Unsafe	None	*	*	
Emerg Response and FOUR Other	None	*	*	
Emerg Response and THREE Other	None	*	*	
Emerg Response and TWO Other	None	29	0.05	
Fall-related	Fall-related	832	1.45	98.1%
Fall-related / MH SubAbuse	None	*	*	
Fall-related / Skin (non-P.I.)	None	*	*	
Fall-related / Trauma Injury	None	*	*	
Fall-related and THREE Other	None	*	*	
Fall-related and TWO Other	None	*	*	
Health Records	Health Records	7335	12.82	98.5%
Health Records / AMA	None	*	*	
Health Records / Admit Discharge	None	*	*	
Health Records / Emerg Response	None	*	*	
Health Records / Fall-related	None	*	*	
Health Records / IV Therapy	None	16	0.03	
Health Records / Labor Delivery	None	*	*	
Health Records / Laboratory	None	42	0.07	
Health Records / MH SubAbuse	None	*	*	
Health Records / Pressure Injury	None	*	*	
Health Records / Radiology	None	*	*	
Health Records / Skin (non-P.I.)	None	*	*	
Health Records / Trauma Injury	None	*	*	
Health Records / Workforce Safety	None	*	*	
Health Records and TWO Other	None	17	0.03	
IV Therapy	IV Therapy	3993	6.98	99.1%
IV Therapy / Admit Discharge	None	*	*	
IV Therapy / Laboratory	None	*	*	
IV Therapy / MH SubAbuse	None	*	*	
IV Therapy / Skin (non-P.I.)	None	*	*	
IV Therapy and TWO Other	None	20	0.03	
Labor Delivery	Labor Delivery	1136	1.98	99.8%
Labor Delivery and TWO Other	None	*	*	
Laboratory	Laboratory	6661	11.64	99.3%
Laboratory / Emerg Response	None	*	*	
Laboratory / Health Records	None	*	*	

Mutually Exclusive Theme Name	Key Theme	Freq	% of Total Sample	Key as % of Theme
Laboratory / IV Therapy	None	*	*	
Laboratory / Labor Delivery	None	*	*	
Laboratory / Patient Unsafe	None	*	*	
Laboratory and THREE Other	None	*	*	
Laboratory and TWO Other	None	36	0.06	
MH SubAbuse	MH SubAbuse	750	1.31	95.2%
MH SubAbuse / IV Therapy	None	*	*	
MH SubAbuse / Patient Unsafe	None	*	*	
MH SubAbuse and THREE Other	None	*	*	
MH SubAbuse and TWO Other	None	33	0.06	
Patient Unsafe	Patient Unsafe	3327	5.81	94.2%
Patient Unsafe / AMA	None	*	*	
Patient Unsafe / Emerg Response	None	11	0.02	
Patient Unsafe / Fall-related	None	*	*	
Patient Unsafe / Health Records	None	32	0.06	
Patient Unsafe / IV Therapy	None	13	0.02	
Patient Unsafe / Labor Delivery	None	*	*	
Patient Unsafe / Laboratory	None	15	0.03	
Patient Unsafe / MH SubAbuse	None	65	0.11	
Patient Unsafe / Pressure Injury	None	*	*	
Patient Unsafe / Radiology	None	17	0.03	
Patient Unsafe / Skin (non-P.I.)	None	*	*	
Patient Unsafe / Trauma Injury	None	*	*	
Patient Unsafe / Workforce Safety	None	19	0.03	
Patient Unsafe and THREE Other	None	*	*	
Patient Unsafe and TWO Other	None	*	*	
Pressure Injury	Pressure Injury	212	0.37	99.5%
Pressure Injury / Health Records	None	*	*	
Radiology	Radiology	3522	6.15	93.9%
Radiology / AMA	None	12	0.02	
Radiology / Admit Discharge	None	*	*	
Radiology / Emerge Response	None	*	*	
Radiology / Emergency Response	None	45	0.08	
Radiology / Fall-related	None	*	*	
Radiology / Health Records	None	35	0.06	
Radiology / IV Therapy	None	28	0.05	
Radiology / Labor Delivery	None	*	*	
Radiology / Laboratory	None	15	0.03	
Radiology / Skin (non-P.I.)	None	*	*	
Radiology / Trauma Injury	None	10	0.02	
Radiology / Workforce Safety	None	*	*	
Radiology and THREE Other	None	*	*	

Mutually Exclusive Theme Name	Key Theme	Freq	% of Total Sample	Key as % of Theme
Radiology and TWO Other	None	56	0.10	
Skin (Non-P.I.)	Skin (Non-P.I.)	1198	2.09	99.4%
Skin (non-P.I.) / Health Records	None	*	*	
Skin (non-P.I.) / Patient Unsafe	None	*	*	
Skin (non-P.I.) and FOUR Other	None	*	*	
Skin (non-P.I.) / MH SubAbuse	None	*	*	
Skin (non-P.I.) / Pressure Injury	None	*	*	
Skin (non-P.I.) / Trauma Injury	None	*	*	
Trauma Injury	Trauma Injury	1359	2.37	99.7%
Trauma Injury / Health Records	None	*	*	
Trauma Injury / IV Therapy	None	*	*	
Trauma Injury and THREE Other	None	*	*	
Workforce Safety	Workforce Safety	3529	6.17	97.4%
Workforce Safety / AMA	None	40	0.07	
Workforce Safety / Emerg Response	None	*	*	
Workforce Safety / Emerge Response	None	*	*	
Workforce Safety / Fall-related	None	*	*	
Workforce Safety / IV Therapy	None	*	*	
Workforce Safety / Labor Delivery	None	*	*	
Workforce Safety / Laboratory	None	*	*	
Workforce Safety / MH SubAbuse	None	32	0.06	
Workforce Safety / Patient Unsafe	None	*	*	
Workforce Safety and TWO Other	None	*	*	
Total (Key Themes and Paired Themes)		**50,739**	**88.6%**	