2026 JULY 19–22 VANCOUVER, BC
JOINT AAPM | COMP MEETING

Quantitative Risk Assessment in Radiation Oncology via LLM-Powered Root Cause Analysis of Incident Reports

Yuntao Wang, MS[1]; Siamak P. Nejad-Davarani, PhD[1]; Elizabeth Bossart, PhD[1]; Matthew T. Studenski, PhD[1]; Mariluz De Ornelas, PhD[1]; Yunze Yang, PhD[1*]
[1]Department of Radiation Oncology, University of Miami

ABSTRACT

Background
Incident-learning systems provide empirical safety data, but free-text narratives are difficult to analyze systematically at scale. Large language models may enable automated, taxonomy-aligned root-cause analysis.

Methods
An LLM analyzed 254 institutional incident reports and assigned pathway categories, contributory factors, responsible parties, and TG-100 severity scores. OLR and ARM were used to identify severity predictors and systemic relationships.

Results
Pre-treatment Planning was the most frequent pathway. Individual and Procedural Factors were the most frequent contributory factors. OLR identified significant severity predictors, while ARM revealed recurrent workflow-cause relationships.

Conclusion
LLM-powered analysis provides a scalable framework for incident-derived quantitative risk assessment in radiation oncology.

CONTACT

Yunze Yang
University of Miami
1475 NW 12th Ave, Miami, FL 33136
yxy971@med.miami.edu
6028288639

INTRODUCTION

Incident Learning Systems (ILS), such as RO-ILS, provide a rich source of empirical safety information by documenting errors, near misses, and unsafe conditions. However, most reports are recorded as free-text narratives, making large-scale quantitative analysis difficult and resource intensive. Recent advances in Large Language Models (LLMs) provide an opportunity to automatically transform narrative incident reports into structured, standardized datasets suitable for statistical analysis. The purpose of this study was to develop and evaluate an LLM-powered framework for automated root cause analysis of radiation oncology incident reports and to perform quantitative risk assessment using structured incident-learning data.

METHODS AND MATERIALS

A total of 254 institutional RO-ILS reports collected between October 2022 and June 2025 were analyzed. Reports included errors, near misses, and unsafe conditions. The UK National Patient Safety Radiotherapy Event Taxonomy was used to classify radiotherapy pathway steps and contributory factors. Event severity was scored from 1 to 10 using the AAPM TG-100 framework. Gemini 2.5 Pro analyzed each incident narrative and generated structured fields for pathway category and subcode, contributory factor category and subcategory, treatment modality, responsible party, and event severity. Multiple pathway and contributory-factor assignments were allowed when clinically appropriate. Descriptive statistics summarized event patterns. Ordinal logistic regression identified predictors of higher severity. Association rule mining identified recurrent relationships among pathways, contributory factors, and responsible parties. Chi-square tests and ANOVA evaluated categorical associations and severity differences. Statistical significance was defined as $p \leq 0.05$.

RESULTS

LLM-Powered Structured UM RO-ILS Pathway and Contributing Factor Database (Taxonomy Mapping)

Events (n=252)

Event ID	Date	Location	Event Description	Brief Description
E001	10/7/2022	MAIN	patient arrived as scheduled for f/u @3pm followed by IV line insertion. Due to pt's MD out of clinic doing Y90, pt (who as an MD) was delayed for approx 2 hours for consent and then was a hard stick for the IV line. Patient was finally taken to CT sim, followed by V sim, and finished @7:06pm. **When scheduling CT sims, we are not aware of MD's schedule out fo the department. Initially we had 2 other similar simulation patients like this one on the schedule in the afternoon on the same day, but the patients were cancelled, or it would have been a big challenge for the schedule.	Significant patient delay for CT simulation due to lack of coordination between clinic scheduling and physician's external procedural schedule.
E002	10/17/2022	MAIN	Inpatient and receiving radiation treatment. Patient was discharged and showed up at Main department before we could call the patient with a time to come in. When checking they had their schedule from west radiation @9:30. Patient was delayed 2.5 hours while replan, QA and physics check was being done. Patient had 5 fractions remaining.	Inpatient was discharged and showed up for treatment before a replan and QA could be completed, causing a 2.5-hour delay.

Contributing Factors

Contributory factor taxonomy 1	CF 1 Reason	CF 1 Responsibility	Contributory factor taxonomy 2	CF 2 Reason	CF 2 Responsibility
CF1d Communication (inaccuracy or omission of verbal, written and so on)	Omission in sharing MD's schedule with the simulation scheduling team led to the delay.	Scheduler/Admin, Department Leadership, Physicist, Dosimetrist, Physician	CF2d Process design (impractical and inefficient processes and so on)	The scheduling process is inefficient as it does not integrate MD availability checks.	Department Leadership, Scheduler/Admin
CF1d Communication (inaccuracy or omission of verbal, written and so on)	Failure to communicate new appointment time to the patient after discharge from inpatient to outpatient.	Scheduler/Admin, Department Leadership, Physician, Nurse, Dosimetrist, Physicist	CF2a No procedures or protocols (not in place or unavailable and so on)	The description implies a complete lack of a standardized protocol for managing the transition of patients from inpatient to outpatient status, which is a more severe failure than just an inefficient process.	Department Leadership

Pathway Failures

Pathway subcode taxonomy 1	Pathway 1 Reason	Pathway subcode taxonomy 2	Pathway 2 Reason
(4j) Consent process and documentation	The event involved a delay in obtaining consent due to the MD's unavailability, directly relating to issues in consent process and documentation timing.	(6a) Bookings made according to protocol	The description notes, "When scheduling CT sims, we are not aware of MD's schedule out fo the department," indicating a failure in the booking protocol, which should account for the availability of essential personnel.
(5i) Recording of patient's specific requirements (includes communication, handover, documentation of patient specific information and so on)	The core issue was a breakdown in the handover of the patient from inpatient services to outpatient radiotherapy, resulting in the patient receiving incorrect scheduling information. This code specifically addresses handover failures.	(6c) Recording of booked appointments (including requested changes following initial booking)	The event was triggered by a discrepancy between two different schedules ("schedule from west radiation @9:30" vs. the Main department's unconfirmed time). This points to a failure in the system for recording and synchronizing appointments, especially after a change in patient status (discharge), which falls under the scope of recording booked appointments and requested changes.

Severity

Severity term	S value	Severity	Rationale
Inconvenience-patient	2–3	3	The primary impact of the event was significant inconvenience and schedule disruption for the patient, requiring an unexpected return visit, which directly aligns with the TG-100 definition of 'Inconvenience-patient'. Per TG-100, this corresponds to a score of 2-3; given the severity of the delay, a score of 3 is assigned.
Inconvenience-patient	2–3	3	The event caused a prolonged wait for the patient, who may have needed to reschedule their travel, constituting a significant inconvenience. Per TG-100, this corresponds to a score of 2-3; given the 2.5-hour delay, a score of 3 is assigned.

Figure 1. LLM-Automated Root Cause Analysis and Data Structuring Pipeline.

Analysis of 254 incident reports identified high-frequency failure categories, significant severity predictors, and recurrent systemic workflow-cause relationships.

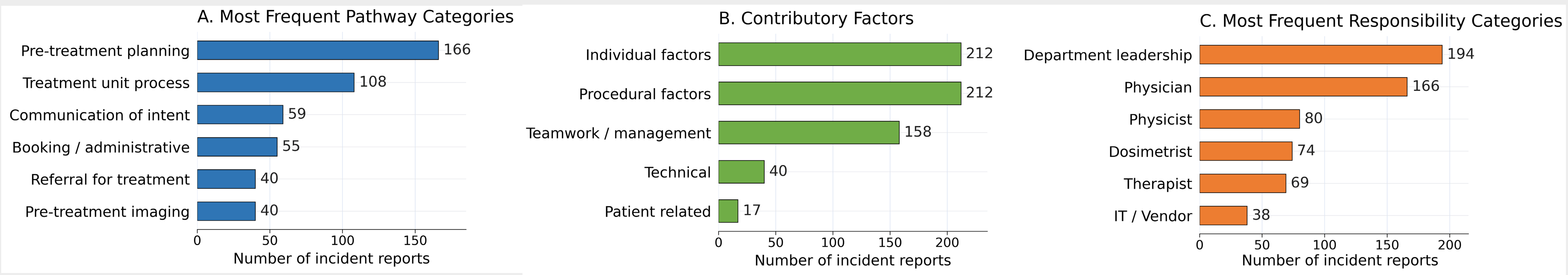

Figure 2. Most frequent pathway, contributory-factor, and responsibility categories. Reports could receive multiple assignments.

Systemic Vulnerabilities Identification of High-Frequency and High-Risk Contributory Factors

Table 1. Identification of high-frequency contributory factors in clinical failures and related severity.

CF Subcategory	Event Count	Mean Severity
Process design (impractical and inefficient processes and so on)	142	5
Slips and lapses (skill-based, involuntary automaticity and so on)	111	5.79
Inadequate leadership (inadequate supervision, congruence or consistency and so on)	106	5.32
Communication (inaccuracy or omission of verbal, written and so on)	90	5.23
Adherence to procedures or protocols	88	6.93
Unclear responsibilities and lines of accountability (across the radiotherapy pathway)	54	5.85
Inadequate procedures or protocols	37	7.43
Inadequate risk assessment (inadequate change management and so on)	32	8.22
Inadequate staffing (insufficient staffing levels or skill mix necessary to meet the demands of a service and so on)	31	4.48
Failure to recognise hazard (knowledge-based and so on)	30	8.20

Table 2. Identification of high-risk predictors to severity via Ordinal Logistic Regression (OLR).

CF Subcategory	Event Count	P-value
Violation (deliberate action, acting outside scope and so on)	25	<0.001
Inadequate leadership (inadequate supervision, congruence or consistency and so on)	106	<0.001
Inadequate risk assessment (inadequate change management and so on)	32	<0.001
Failure to recognise hazard (knowledge-based and so on)	30	<0.001
Process design (impractical and inefficient processes and so on)	142	<0.001
Inadequate training (inadequate or lack of training and so on)	21	0.004
Equipment or IT network failure (including immobilisation and accessories)	22	0.011
Inadequate procedures or protocols	37	0.012

Co-occurrence Analysis of Failure Pathways (Failure Patterns)

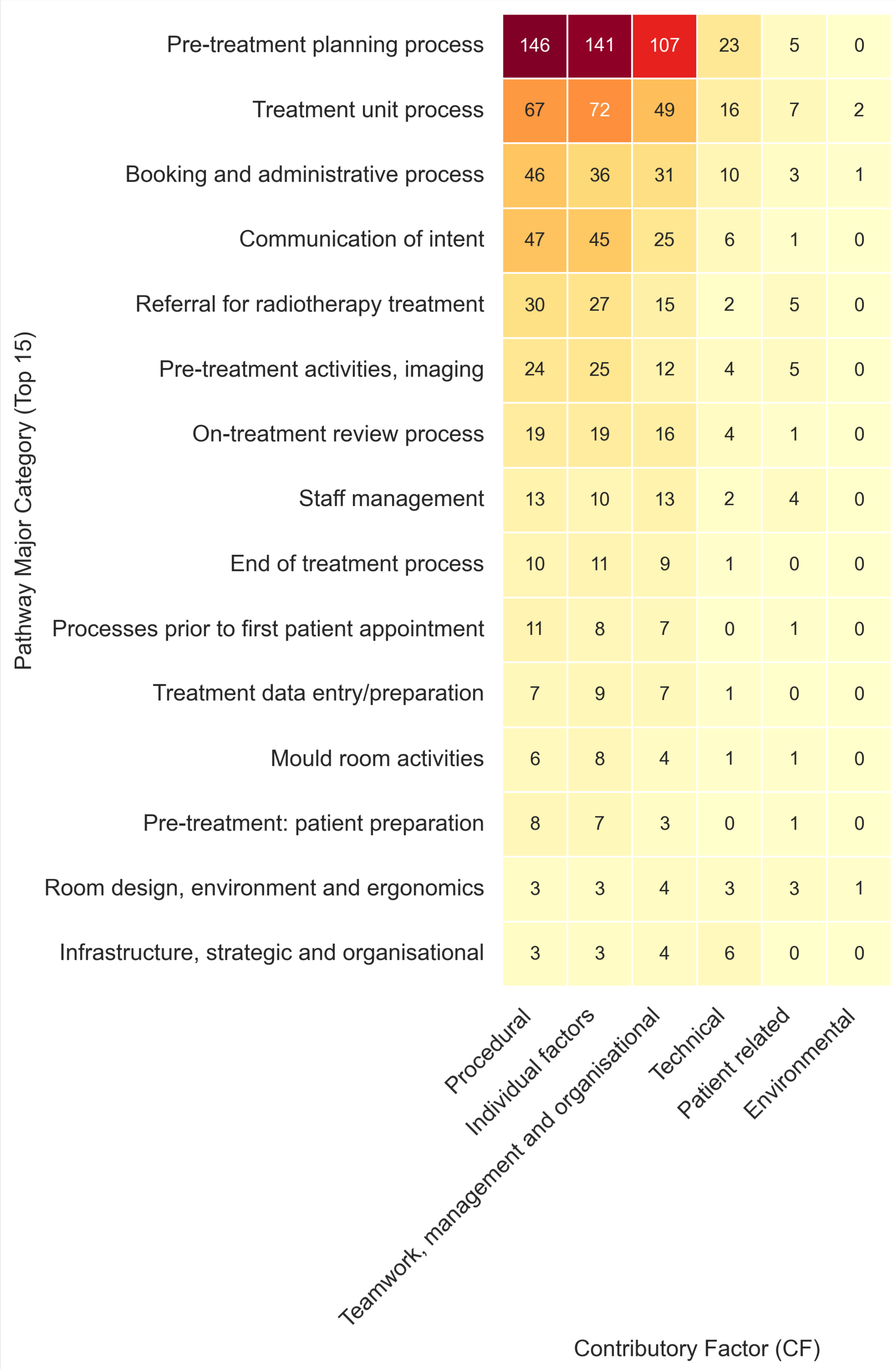

Figure 3. Heatmaps of Co-occurrence Frequencies for Pairwise Combinations. (Left) Heatmap between Pathway Categories (Top 15)

CONCLUSIONS

We demonstrate a LLM-powered approach that moves beyond subjective, prospective risk assessment to an objective, evidence-based, and quantitative model. By systematically converting unstructured narrative data into a rich, structured dataset, this approach enables the identification of the key drivers of event severity and the mapping of complex, systemic failure pathways. This methodology paves the way for a new generation of intelligent safety systems to improve the quality and safety of care.

REFERENCES

1. Wang Y et al. RCA of radiation oncology incidents using LLMs. arXiv. 2025.