SAEM 2026 · GUN VIOLENCE / AI

Topic Modeling Narrative Reports of United States Firearm Incidents in 2024–2025

R. Andrew Taylor, Andrew Muck, Thomas Hartka, Matthew Trowbridge, George Glass, Pavel Chernyavskiy, Jonathan Swap, Farah Turkistani, Rupesh Silwal

Department of Emergency Medicine, University of Virginia · UVA Health

50,154

U.S. FIREARM INCIDENTS ANALYZED (JAN 2024 – DEC 2025)

Across 51 states and 5,946 cities · **48.6% fatal**

117

TOPICS DERIVED (BERTOPIC + GPT-4O)

30 with >1% prevalence · 67.3% coverage

aOR 6.01

ARGUMENT-RELATED SHOOTINGS · FATAL OUTCOME

95% CI 5.14–7.02 · 83.0% fatal · 5.10% prevalence

aOR 13.70

INTIMATE PARTNER VIOLENCE IN HOMES

95% CI 11.4–16.5

BACKGROUND

Firearm injury context is missing from operational data.

Firearm injury imposes substantial emergency department burden, yet incident context is rarely available for timely operations and prevention planning.

Aim. Use topic modeling to derive an empirically grounded taxonomy of firearm-incident contexts from a national incident database and quantify associations with ED-relevant proxies of acuity, timing, and setting.

METHODS

BERTopic + GPT-4o on the EveryShot database.

Source. EveryShot database (derived from media reports), Jan 1, 2024 – Dec 16, 2025.

Inclusion. ≥1 wounded or ≥1 killed.

Exclusion. Brandishing.

Modeling. BERTopic with OpenAI text-embedding-3-large; GPT-4o for topic labels.

Topic selection. Coherence/diversity metrics + clinician interpretability; analyses focused on topics with >1% prevalence.

Validity. Convergent comparison with structured database fields (intent, incident type).

Statistics. Multivariable logistic regression for fatal outcome, nighttime occurrence, location categories.

RESULTS

Interpretable taxonomy with marked heterogeneity.

- **50,154** Incidents · 51 states · 5,946 cities
- **24,355 (48.6%)** Fatal incidents
- **117 / 30** Topics identified · with >1% prevalence (67.3% coverage)
- **aOR 1.63** Officer-involved · 6.97% prev · 57.1% fatal · CI 1.47–1.82
- **aOR 6.01** Argument-related · 5.10% prev · 83.0% fatal · CI 5.14–7.02
- **aOR 0.70** Gas station · 1.76% prev · 39.4% fatal · CI 0.56–0.87
- **aOR 0.89** Unintentional child self-inflicted · 1.17% prev · 29.1% fatal · CI 0.64–1.25
- **98.6%** Topic-intent alignment · road rage
- **97.0% / 96.2%** Escalating altercations / intimate partner violence alignment
- **aOR 1.65** Apartment-complex shootings · night · CI 1.3–2.1
- **aOR 13.70** IPV clustering in homes · CI 11.4–16.5

CONCLUSION

A clinically interpretable taxonomy of firearm-incident contexts for ED operations and prevention.

Topic modeling of a national incident-level firearm database produced a clinically interpretable taxonomy of firearm-incident contexts and revealed marked heterogeneity in timing, setting, and severity.

ASK ME ABOUT

- BERTopic + GPT-4o pipeline
- EveryShot database
- Topic-intent validation
- Operations + prevention applications