Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Development and validation of a SHELL model-based risk assessment scale for anesthesia medication management: A cross-sectional study

Lile Sheng
Union Hospital, Tongji Medical College, Huazhong University of Science and Technology
Fang Tan
Union Hospital, Tongji Medical College, Huazhong University of Science and Technology
Tao Zhang
Union Hospital, Tongji Medical College, Huazhong University of Science and Technology
Wanli Xie
Union Hospital, Tongji Medical College, Huazhong University of Science and Technology
Meng Meng Shen
Zhongnan Hospital of Wuhan University
Zhenghua Zhao
whxhmz@foxmail.com

Union Hospital, Tongji Medical College, Huazhong University of Science and Technology

Research Article

Keywords: Anesthesia, Medication Safety, Risk Assessment, Scale Development, SHELL Model, Nursing

Posted Date: August 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9938871/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Managing anesthesia medications in operating rooms is high-risk, and traditional checklists often miss underlying system vulnerabilities. The SHELL model—which examines Software, Hardware, Environment, Liveware, and Liveware-Liveware interfaces—provides a systems-based approach to identify risks before they cause harm. To date, no validated SHELL-based tool exists specifically for anesthesia medication management.

Objective

To develop and validate the Anesthesia Medication Management Risk Assessment Scale (AMM-RAS) based on the SHELL model.

Design:

A cross-sectional instrument development and validation study.

Settings and participants:

A total of 771 anesthesiology nurses from 21 hospitals across China participated between October and November 2025.

Methods

The initial item pool was generated through systematic literature review, semi-structured interviews, and two rounds of Delphi expert consultation. The sample was randomly split into derivation (n = 385) and validation (n = 386) subsamples. Item analysis and exploratory factor analysis (EFA) were performed on the derivation sample, followed by confirmatory factor analysis (CFA), convergent/discriminant validity testing, and reliability testing on the validation sample.

Results

The final AMM-RAS comprised 25 items across five dimensions aligned with the SHELL model. EFA explained 74.511% of total variance. CFA demonstrated acceptable model fit ($\chi^2/df = 1.96$, RMSEA = 0.06, CFI = 0.93, TLI = 0.92, SRMR = 0.04). The scale demonstrated high internal consistency (overall Cronbach's α = 0.960; dimension α range: 0.886–0.971) and test-retest reliability (ICC = 0.87).

Conclusion

The AMM-RAS is a reliable and valid instrument for systematically assessing anesthesia medication management risks. It enables nurse managers to identify vulnerable system interfaces and implement targeted quality improvement interventions.

1. Introduction

1.1 Background

Anesthesia medications—including opioids, sedatives, muscle relaxants, and local anesthetics—are indispensable in surgery, analgesia, and resuscitation. However, the medication administration presents unique safety challenges due to the frequent use of controlled substances, complex oversight processes, and high-acuity clinical conditions (Wong et al., 2023; Zhang et al., 2024). Studies have identified persistent issues including drug diversion, misappropriation, documentation errors, and inadequate supervision in anesthesiology departments worldwide (deVries et al., 2025; Habchi et al., 2020).

In China, narcotics and Schedule I psychotropic substances are classified as specially controlled substances under the Drug Administration Law and the Regulations for the Administration of Narcotic Drugs and Psychotropic Substances. The anesthesiology department, as the unit with the highest concentration of these agents, bears particular responsibility for medication safety. Ineffective management not only threatens patient safety but also carries legal implications for healthcare workers and potential social harm from drug diversion (Le et al., 2025).

1.2 Knowledge gap

Current approaches to anesthesia medication safety rely heavily on compliance checklists and retrospective incident analysis. While these tools can identify explicit violations or individual errors, they struggle to detect systemic vulnerabilities embedded in the interactions between policies, equipment, environment, personnel, and teamwork (Wahr et al., 2017). For example, a well-designed policy may fail because it does not align with workflow realities; an advanced automated dispensing cabinet may be underutilized because staff lack adequate training or the physical environment creates distractions.

What remains missing is a theoretically grounded, prospectively applicable instrument that can assess the "interfaces" between system elements—where most latent risks reside. Without such a tool, nurse managers lack systematic guidance for diagnosing whether a safety gap stems from inadequate policies, unreliable equipment, disruptive environments, individual competence, or team communication breakdowns.

1.3 Theoretical framework and study aim

The SHELL model, originally developed by Elwyn L. Edwards (1972) for aviation safety and subsequently adopted by the World Health Organization for healthcare applications (WHO, 2011), provides an appropriate framework for addressing this gap. The model conceptualizes system safety as emerging from the interactions among five components:**S**oftware (policies, procedures, regulations),**H**ardware (equipment, facilities, technology),**E**nvironment (physical workspace, safety culture),**L**iveware (individual knowledge, skills, attitudes), and **L**iveware-**L**iveware (teamwork, communication, supervision). The model's core premise is that safety failures typically arise not from isolated component deficiencies but from poor interfaces between components (Hong et al., 2016).

This study aimed to develop and psychometrically validate the Anesthesia Medication Management Risk Assessment Scale (AMM-RAS) based on the SHELL model, thereby providing a systematic instrument for proactive risk identification in anesthesia nursing practice.

2. Methods

2.1 Study design

This was a cross-sectional instrument development and validation study, reported according to the COSMIN (COnsensus-based Standards for the selection of health Measurement INstruments) guidelines for patient-reported outcome measure development.

2.2 Stage 1: Item pool development

2.2.1 Systematic literature review

A systematic literature review was conducted following PRISMA guidelines (Page et al., 2021). We searched PubMed, Elsevier ScienceDirect, Ovid MEDLINE, CNKI, and Wanfang Data from database inception to May 2025. Search terms included combinations of "anesthesia medication," "opioids," "controlled substances," "medication management," "risk assessment," and "SHELL model." Additional searches were performed on the websites of the American Society of Anesthesiologists and the Chinese Society of Anesthesiology. Reference lists of included articles were manually screened.

After removing duplicates, 2,891 records underwent title and abstract screening. Full-text review of 36 articles resulted in the extraction of risk factors, safety practices, and management processes relevant to anesthesia medication management. Figure 1 presents the PRISMA 2020 flow diagram of the study identification and selection process. This information was synthesized with key Chinese regulations (Drug Administration Law, Regulations for Narcotic and Psychotropic Drug Administration) to generate an initial pool of 58 items across five SHELL-aligned dimensions.

**Consider, if feasible to do so, reporting the number of records identified from each database or register searched (rather than the total number across all databases/registers).*

***If automation tools were used, indicate how many records were excluded by a human and how many were excluded by automation tools.*

2.2.2 Semi-structured interviews

Purposive sampling was used to recruit 22 anesthesia nurses and nurse managers from tertiary hospitals for individual semi-structured interviews. The interview guide was developed based on the SHELL model dimensions, exploring: (1) perceptions of policy clarity and feasibility, (2) equipment usability issues, (3) environmental barriers to safe practice, (4) individual knowledge and competency gaps, and (5) teamwork and communication challenges in medication processes. Interviews were audio-recorded, transcribed verbatim, and analyzed using Colaizzi's seven-step method. Interview findings refined the item pool, resulting in 58 items retained.

2.3 Stage 2: Content validation

2.3.1 Expert panel

A panel of 17 experts was purposively selected from five specialty areas: clinical anesthesia, anesthesia nursing, nursing management, nursing quality control, and medication management. Inclusion criteria were: (a) working in a tertiary Grade A hospital with recognized national influence in anesthesiology; (b) holding a bachelor's degree or higher and an intermediate or higher professional title; (c) ≥ 10 years of experience in their specialty; (d) prior involvement in research projects. Experts had a mean of 16.53 (SD = 2.61) years of experience; four held doctoral degrees, nine held master's degrees, and four held bachelor's degrees.

2.3.2 Delphi process

Two rounds of electronic Delphi consultation were conducted from August to September 2025. Experts rated each item on a 5-point Likert scale (1 = not important to 5 = very important) and provided qualitative feedback on item clarity, relevance, and representativeness. Items were retained if they met the following criteria: importance score > 3.5 and coefficient of variation < 0.25. Items suggested for deletion, merging, or addition by ≥ 20% of experts were discussed by the research team for modification.

After round 1, 11 experts (64.7%) provided 17 modification suggestions. One item was deleted (coefficient of variation ≥ 0.25), 19 items were merged into seven items, and three items were added (e.g., "During shift handover, drug handover information is clear and without omission"). The scale was revised to 42 items. After round 2, Kendall's coefficient of concordance was 0.395 (χ^2=241.887, $p < 0.001$), indicating good expert agreement. Four items were deleted, two items were reworded, resulting in a 37-item preliminary scale.

2.4 Stage 3: Pilot testing

Cognitive interviews were conducted with 20 anesthesia nurses from a tertiary hospital in Wuhan in July 2025. Using the think-aloud method and verbal probing, nurses verbalized their understanding of each

item. No items required modification; all nurses reported the scale was clear and could be completed within 5 minutes.

2.5 Stage 4: Psychometric testing

2.5.1 Sample and setting

A cross-sectional survey was conducted from October to November 2025. Convenience sampling was used to recruit anesthesiology nurses from 21 hospitals across China. Inclusion criteria were: (a) currently working in anesthesia nursing; (b) ≥ 1 year of experience; (c) able to provide informed consent. Exclusion criteria were visiting nurses, standardized training nurses, and intern nurses.

Based on the subject-to-item ratio of at least 10:1 recommended for factor analysis, the 37-item scale required at least 370 participants. Accounting for an approximately 20% invalid response rate, a target of 462 was set. A total of 832 questionnaires were distributed; 771 were returned and valid (92.7% effective response rate).

2.5.2 Measures

Demographic questionnaire

Self-developed items collecting age, gender, education level, professional title, years of anesthesia experience, and work setting [anesthesia outpatient, Post-Anesthesia Care Unit (PACU), intraoperative monitoring, general duty].

Preliminary AMM-RAS

The 37-item scale used a 5-point Likert response format (1 = never to 5 = always), with higher scores indicating lower perceived risk (better safety practices).

2.5.3 Data collection

Trained research assistants distributed electronic questionnaires via the "Questionnaire Star" platform with assistance from head nurses. The first page presented study information and obtained informed consent. Quality control measures included: single response per IP address, mandatory completion of all items, and exclusion of questionnaires with completion time < 3 minutes or patterned responses (e.g., identical ratings across all items).

2.5.4 Split-sample strategy

The total sample (N = 771) was randomly split into two subsamples using SPSS 27.0: the derivation sample (n = 385) for exploratory factor analysis (EFA) and initial reliability testing, and the validation sample (n = 386) for confirmatory factor analysis (CFA) and final reliability testing. No significant differences in demographic characteristics were found between subsamples (all $p > 0.05$).

2.6 Data analysis

2.6.1 Item analysis (derivation sample)

Three methods were used: (a) critical ratio method—comparing top and bottom 27% of total scores using independent t-tests, retaining items with critical ratio > 3.0 and $p < 0.05$; (b) correlation coefficient method—retaining items with item-total correlation > 0.4; (c) Cronbach's α method—retaining items whose deletion did not increase the overall α coefficient (Piredda et al., 2024).

2.6.2 Exploratory factor analysis (derivation sample)

The Kaiser-Meyer-Olkin (KMO) measure and Bartlett's test of sphericity assessed sampling adequacy. Principal component analysis with varimax orthogonal rotation was used to extract factors with eigenvalues > 1. Items were deleted if they met any of the following criteria: factor loading < 0.40; cross-loading ≥ 0.40 on two or more factors; belonging to a factor with < 3 items; or lack of conceptual interpretability.

2.6.3 Confirmatory factor analysis (validation sample)

CFA using maximum likelihood estimation tested the factor structure derived from EFA. Model fit was assessed using the following indices: $\chi^2/df < 3$ acceptable, RMSEA < 0.08 acceptable, SRMR < 0.08 acceptable, CFI > 0.90 acceptable, TLI > 0.90 acceptable, and IFI > 0.90 acceptable.

2.6.4 Validity and reliability

Convergent validity was assessed using average variance extracted (AVE; acceptable > 0.50) and composite reliability (CR; acceptable > 0.70). Discriminant validity was confirmed when the square root of each dimension's AVE exceeded its correlations with other dimensions. Content validity was evaluated using item-level (I-CVI) and scale-level (S-CVI) content validity indices, with I-CVI > 0.78 and S-CVI > 0.90 indicating good content validity.

Internal consistency was assessed using Cronbach's α coefficient (> 0.80 acceptable). Test-retest reliability was assessed by re-administering the scale to 30 participants after a two-week interval, calculating the intraclass correlation coefficient (ICC; >0.70 acceptable).

Statistical analyses were performed using SPSS 27.0 and Amos 26.0, with two-tailed $p < 0.05$ considered statistically significant.

3. Results

3.1 Participant characteristics

A total of 771 anesthesia nurses participated, with mean age 31.63 (SD = 5.90) years. The sample was predominantly female (96.6%). The majority held bachelor's degrees (82.1%) and were nurse

practitioners (47.6%) or supervising nurses (31.3%). Work settings included PACU (55.2%), intraoperative monitoring (23.5%), anesthesia outpatient (10.3%), and general duty (11.0%). Participant characteristics did not differ significantly between derivation and validation subsamples.

3.2 Item analysis

In the derivation sample (n = 385), all 37 items demonstrated critical ratio values > 3.0 ($p < 0.001$) and item-total correlations ranging from 0.615 to 0.816 ($p < 0.001$). The overall Cronbach's α was 0.974, and no single item deletion increased this coefficient. Consequently, all 37 items were retained for EFA.

3.3 Exploratory factor analysis

The EFA sample (n = 385) yielded a KMO of 0.959 and a significant Bartlett's test (χ^2=27,757.235, $p < 0.001$), confirming data suitability for factor analysis.

The first EFA extracted five factors with eigenvalues > 1, explaining 69.771% of variance. Ten items with cross-loadings ≥ 0.40 and two items with loadings < 0.40 were deleted. The second EFA on the remaining 25 items again extracted five factors (KMO = 0.943; Bartlett's χ^2=17,885.316, $p < 0.001$), explaining 74.511% of total variance. All factor loadings exceeded 0.40 with no cross-loadings. The five factors corresponded precisely to the SHELL model dimensions: Software (5 items), Hardware (7 items), Environment (3 items), Liveware (3 items), and Liveware-Liveware (7 items).

Table 1 presents the factor loadings for the final 25-item AMM-RAS.

Table 1
Exploratory factor analysis results: Factor loadings for the 25-item AMM-RAS (n = 385)

Item	Software	Hardware	Environment	Liveware	Liveware-Liveware
Software (S)					
S1: Policy awareness	**0.786**	0.147	0.196	0.091	0.095
S2: Emergency procedures	**0.823**	0.234	0.266	0.157	0.139
S3: Effective training	**0.775**	0.272	0.259	0.188	0.177
S4: Standardized records	**0.685**	0.297	0.250	0.242	0.267
S5: Smooth processes	**0.589**	0.316	0.297	0.234	0.241
Hardware (H)					
H1: Reliable facilities	0.175	**0.768**	0.209	0.059	0.063
H2: Strict management	0.243	**0.790**	0.213	0.160	0.176
H3: Clear labeling	0.287	**0.758**	0.213	0.214	0.220
H4: Standardized labels	0.328	**0.692**	0.254	0.231	0.222
H5: Controlled environment	0.291	**0.664**	0.291	0.288	0.245
H6: Effective security	0.317	**0.550**	0.298	0.288	0.245
H7: System proficiency	0.343	**0.427**	0.335	0.317	0.232
Environment (E)					
E1: Physical environment	0.201	0.173	**0.872**	0.144	0.069
E2: Safety culture	0.266	0.247	**0.827**	0.244	0.163
E3: Safety priority	0.303	0.281	**0.699**	0.259	0.270
Liveware (L)					
L1: Possesses knowledge	0.159	0.195	0.130	**0.856**	0.103
L2: Adheres to policies	0.280	0.225	0.240	**0.830**	0.162
L3: Sense of responsibility	0.378	0.256	0.314	**0.660**	0.192
Liveware-Liveware (LL)					
LL1: Preoperative count	0.163	0.091	0.128	0.151	**0.768**

Item	Software	Hardware	Environment	Liveware	Liveware-Liveware
LL2: High-risk verification	0.267	0.192	0.169	0.179	**0.792**
LL3: Issuance verification	0.275	0.237	0.177	0.201	**0.777**
LL4: Destruction supervision	0.271	0.266	0.202	0.189	**0.738**
LL5: Empty vial recovery	0.268	0.356	0.187	0.179	**0.677**
LL6: Clear handover	0.298	0.361	0.233	0.197	**0.581**
LL7: Reminder culture	0.314	0.310	0.296	0.220	**0.554**

**Note: Bold values indicate primary factor loadings (> 0.40) for each item.*

3.4 Confirmatory factor analysis

In the validation sample (n = 386), CFA tested the five-factor structure. The model demonstrated acceptable fit: $\chi^2/df = 1.96$, RMSEA = 0.06 (90% CI: 0.055–0.065), SRMR = 0.04, CFI = 0.93, TLI = 0.92, IFI = 0.93. All factor loadings were statistically significant ($p < 0.001$), supporting the structural validity of the five-dimension model.

3.5 Convergent and discriminant validity

As shown in Table 2, AVE values ranged from 0.636 to 0.762 (all > 0.50) and CR values from 0.905 to 0.931 (all > 0.90), indicating good convergent validity. For each dimension, the square root of AVE exceeded the correlations with other dimensions, supporting discriminant validity (Erkayiran and Aslan, 2025).

Table 2
Convergent and discriminant validity of the AMM-RAS dimensions (n = 386)

Dimension	AVE	CR	1	2	3	4	5
1. Software	0.636	0.924	**0.797**				
2. Hardware	0.662	0.931	0.612	**0.814**			
3. Environment	0.762	0.907	0.548	0.589	**0.873**		
4. Liveware	0.682	0.914	0.519	0.554	0.508	**0.826**	
5. Liveware-Liveware	0.762	0.905	0.635	0.671	0.542	0.533	**0.873**

Note: AVE = average variance extracted; CR = composite reliability. Diagonal bold values are square roots of AVE; off-diagonal values are correlation coefficients between dimensions.

3.6 Reliability

Table 3 presents the internal consistency and test-retest reliability results. The overall Cronbach's α was 0.960, with dimension α coefficients ranging from 0.886 to 0.971, all exceeding the 0.80 threshold. Two-week test-retest ICC for the total scale was 0.87, with dimension ICCs ranging from 0.886 to 0.929 (all $p < 0.001$).

Table 3
Reliability of the AMM-RAS (n = 386 for Cronbach's α; n = 30 for test-retest)

Dimension	Number of items	Cronbach's α	Test-retest ICC (95% CI)
Software	5	0.886	0.886 (0.802–0.938)
Hardware	7	0.926	0.887 (0.804–0.939)
Environment	3	0.931	0.896 (0.816–0.945)
Liveware	3	0.941	0.917 (0.847–0.958)
Liveware-Liveware	7	0.971	0.929 (0.867–0.964)
Total scale	**25**	**0.960**	**0.870 (0.779–0.928)**
Note: ICC = intraclass correlation coefficient; CI = confidence interval. All p < 0.001.			

4. Discussion

4.1 Summary of main findings

We developed the AMM-RAS, a 25-item SHELL-based tool, and tested its psychometric properties in a large sample of anesthesia nurses. The results were strong: a clean five-factor structure that explained nearly 75% of the variance, excellent internal consistency (α = 0.960), and good two-week stability (ICC = 0.87). Each of the five factors lined up exactly with the five SHELL components, which supports the tool's theoretical grounding.

4.2 Theoretical contribution

The theoretical contribution lies in applying the SHELL model to anesthesia medication management—not just checking individual elements, but assessing how well they work together.The core innovation resides in its reframing of system safety evaluation from isolated element checks to interface quality assessment.

Traditional risk assessment tools in perioperative medication safety, such as checklists from the World Health Organization Surgical Safety Checklist or institutional audit forms, typically focus on verifying the presence or absence of specific elements (e.g., "Is the medication labeled?" "Was double-checking performed?"). While useful for compliance monitoring, these approaches cannot detect whether the *interaction* between a policy and a practitioner's workflow is effective, or whether environmental factors undermine equipment usability (Wahr et al., 2017).

The SHELL model's application in this scale addresses this limitation. By operationalizing each dimension and its interfaces, the AMM-RAS enables identification of specific breakdown points. For example, a low score on the Hardware dimension might reflect unreliable equipment, but a low Liveware-Liveware score with adequate Hardware suggests the problem lies in team communication rather than technology. This diagnostic granularity has been identified as a critical gap in existing safety measurement approaches (Carayon et al., 2014).

4.3 Comparison with existing instruments

To our knowledge, no validated tool specifically targets anesthesia medication management risks.The Safety Attitudes Questionnaire (Sexton et al., 2006) assesses safety climate broadly but does not address medication-specific processes. The Medication Safety in the Operating Room recommendations provide evidence-based practices but not a quantifiable assessment tool (Wahr et al., 2017). The AMM-RAS fills this gap by offering a theory-driven, validated scale that produces both a total risk score and dimension-specific profiles.

Our findings align with previous SHELL model applications in healthcare. Kang et al. (2025) applied the model to multidrug-resistant organism infection control in ICUs, similarly demonstrating that the model's strength resides in interface analysis rather than component assessment. The present study extends this evidence by operationalizing the model for a new clinical context—anesthesia medication management—and providing a validated quantitative instrument for routine use.

4.4 Methodological considerations

This study has several methodological strengths.The split-sample strategy ensured independence between EFA and CFA, reducing the risk of capitalizing on chance (Baris et al., 2023). The sample size exceeded minimum requirements, with 771 participants providing stable factor analytic results. The two-round Delphi process achieved strong expert consensus (Kendall's $W = 0.395$, $p < 0.001$) and an authority coefficient of 0.91, indicating robust content validity.

The Cronbach's α of 0.960, while exceeding the conventional 0.80 threshold, approaches the level at which some methodologists suggest potential item redundancy (> 0.95). However, our item analysis showed that deleting any single item did not meaningfully increase α, and the clear five-factor structure with satisfactory discriminant validity argues against problematic redundancy. Instead, this pattern may reflect the inherently interconnected nature of safe practices in the tightly coupled anesthesia medication system, where policies, equipment, environment, individual behavior, and teamwork are necessarily interrelated.

4.5 Implications for nursing practice and management

What can nurse managers actually do with this scale? Here are a few practical uses.

Systematic risk diagnosis. Instead of guessing or waiting for an incident to happen, managers can give the scale to their staff and see which of the five SHELL dimensions scores lowest. For example, you

might find that Hardware and Liveware are fine, but Environment—especially safety culture—is weak. That tells you to work on non-punitive reporting, not buy more equipment.

Targeted improvement. The dimensional scores tell you where to put your effort. Low on Liveware-Liveware? Focus on team communication, handover protocols, or structured double-checks. Low on Software? Clarify policies or offer better training.

Tracking progress over time. Re-administer the scale after an intervention to see if scores improve. This turns safety management from reactive problem-solving into ongoing, data-driven improvement.

Complementing incident reports. The AMM-RAS doesn't replace your existing reporting system. Incident reports tell you what went wrong after the fact. This scale helps you spot latent risks before anything happens—a key step toward a real safety culture, not just a blame culture (Reason, 2000).

4.6 Limitations

This study has several limitations.First, the sample predominantly comprised nurses from tertiary hospitals in China, limiting generalizability to primary care settings or other countries with different regulatory frameworks. Second, the scale relies on self-report, which may be susceptible to social desirability bias despite assurances of anonymity. Third, the cross-sectional design does not permit assessment of predictive validity—the scale's ability to forecast actual medication errors or diversion incidents—which requires longitudinal follow-up. Fourth, the study did not assess inter-rater reliability, which would be relevant if the scale is used for observational assessments rather than self-report.

4.7 Future research directions

Future studies could test the AMM-RAS in other settings—primary hospitals, international contexts—with appropriate cross-cultural adaptation. Longitudinal studies are needed to see whether lower scores predict more medication errors down the line. It would also be useful to determine the minimal clinically important difference, so managers know how much score change actually matters. Finally, an electronic version with automated scoring and visualization could make the tool easier to use in daily practice and help benchmark across institutions.

5. Conclusions

The AMM-RAS is a valid and reliable 25-item instrument for assessing anesthesia medication management risks based on the SHELL model. The scale's five-factor structure aligns with theoretical dimensions, demonstrating good model fit, strong internal consistency, and temporal stability. For nursing practice, the AMM-RAS offers a systematic approach to diagnosing specific system vulnerabilities at the interfaces of policies, equipment, environment, individual competence, and teamwork. This enables nurse managers to move beyond compliance-oriented checklists toward targeted, evidence-based quality improvement interventions. The scale represents a practical tool for advancing anesthesia medication safety from retrospective error analysis to proactive risk identification.

Declarations

Ethical considerations

Not applicable.

Competing interests

The authors declare no competing interests.

Funding

There was no funding support for this study.

Author Contribution

Lile Sheng: Writing – review & editing, Writing – original draft, Methodology, Investigation.Fang Tan:Formal analysis, Data curation, Conceptualization. Tao Zhang: Resources, Methodology, Investigation, Formal analysis, Data curation. Wanli Xie: Writing – original draft, Methodology, Investigation, Data curation. Mengmeng Shen: Writing – original draft, Methodology, Investigation, Data curation, Conceptualization. Zhenghua Zhao: Writing – original draft, Validation, Methodology, Investigation, Data curation, Conceptualization.

Acknowledgement

We thank all the researchers who contributed to the study; and to the clinical nurses who participated in this study.

References

1. Baris VK, Intepeler SS, Unal A. 2023. Development and psychometric validation of the Sickness Presenteeism Scale-Nurse. Int J Nurs Pract 29 (5), e13168.
2. Carayon P, Wetterneck TB, Rivera-Rodriguez AJ, Hundt AS, Hoonakker P, Holden R, Gurses AP. Human factors systems approach to healthcare quality and patient safety. Appl Ergon. 2014;45(1):14–25.
3. deVries M, McGillis Hall L, Dainty K, Fan M, Tscheng D, Hamilton M, Trbovich P. A conceptual framework for identifying and managing system vulnerabilities for diversion of controlled substances in healthcare. Res Social Adm Pharm. 2025;21(4):228–38.
4. Erkayiran O, Aslan R. Development and psychometric validation of the Nurse-Patient Therapeutic Interaction Scale (NuPTIS). Nurse Educ Pract. 2025;87:104478.

5. Habchi KM, Li MT, Mallard CA, Baker M, Ortega R. The anesthesiologist's armamentarium: From recreation to medication and back. J Anesth Hist. 2020;6(1):17–26.
6. Hong S, Lee K, Seol E, Young S. Safety perceptions of training pilots based on training institution and experience. J Air Transp Manag. 2016;55:213–21.
7. Kang X, Zhang P, Xu Q, Feng Z, Yin B. Innovative solutions for multidrug-resistant organisms' infections in intensive care unit: A joint efficacy evaluation of multidisciplinary team and SHEL model. Curr Ther Res. 2025;102:100766.
8. Le C, Mellett J, Barreto J, Lopez B. Utilization of an electronic health record dashboard report to monitor unreconciled controlled substance dispenses. J Am Pharm Assoc. 2025;65(3):102968.
9. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hrobjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, McDonald S, McGuinness LA, Stewart LA, Thomas J, Tricco AC, Welch VA, Whiting P, Moher D. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ. 2021;372:n71.
10. Piredda M, Sguanci M, De Maria M, Petrucci G, Usai M, Fiorini J, De Marinis MG. 2024. Nurses' evidence-based knowledge and self-efficacy in venous access device insertion and management: Development and validation of a questionnaire. Nurs Open 11 (7), e2177.
11. Reason J. Human error: Models and management. BMJ. 2000;320(7237):768–70.
12. Sexton JB, Helmreich RL, Neilands TB, Rowan K, Vella K, Boyden J, Roberts PR, Thomas EJ. The Safety Attitudes Questionnaire: Psychometric properties, benchmarking data, and emerging research. BMC Health Serv Res. 2006;6:44.
13. Shao R, Tao TT. Governing drugs by law to ensure public medication safety: Interpretation of key amendments in the Drug Administration Law of the People's Republic of China. Adverse Drug React J. 2019;21(5):377–80.
14. Sheng LL, Lan X, Tan F, Zhao ZH, Chang WL, Xia HF, Chen QH, Du CY, Li CY. Development of sensitive quality indicators for nursing care in the post-anesthesia care unit. Chin J Nurs. 2020;55(6):805–10.
15. Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, Bouter LM, de Vet HCW. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34–42.
16. Wahr JA, Abernathy JH, III, Lazarra EH, Keebler JR, Wall MH, Lynch I, Wolfe R, Cooper RL. Medication safety in the operating room: Literature and expert-based recommendations. Br J Anaesth. 2017;118(1):32–43.
17. Wong MJ, Wang Y, Blake L, Ke JXC. Preventing controlled substance diversion in perioperative settings: A narrative review. Can J Anaesth. 2023;70(12):1989–2001.
18. World Health Organization. The WHO Patient Safety Curriculum Guide: Multi-Professional Edition. Geneva: WHO; 2011.
19. Zhang J, Li J, Li XX, Li ZQ, Du HM, Mi WD, Zhao J, Guo XY, Zhou Y, Yang YH. Current status and analysis of anesthetic and psychotropic substances management in anesthesiology departments of

secondary and above level hospitals in Beijing. J Anesth Saf Qual Control. 2024;6(4):191–5.

Figures

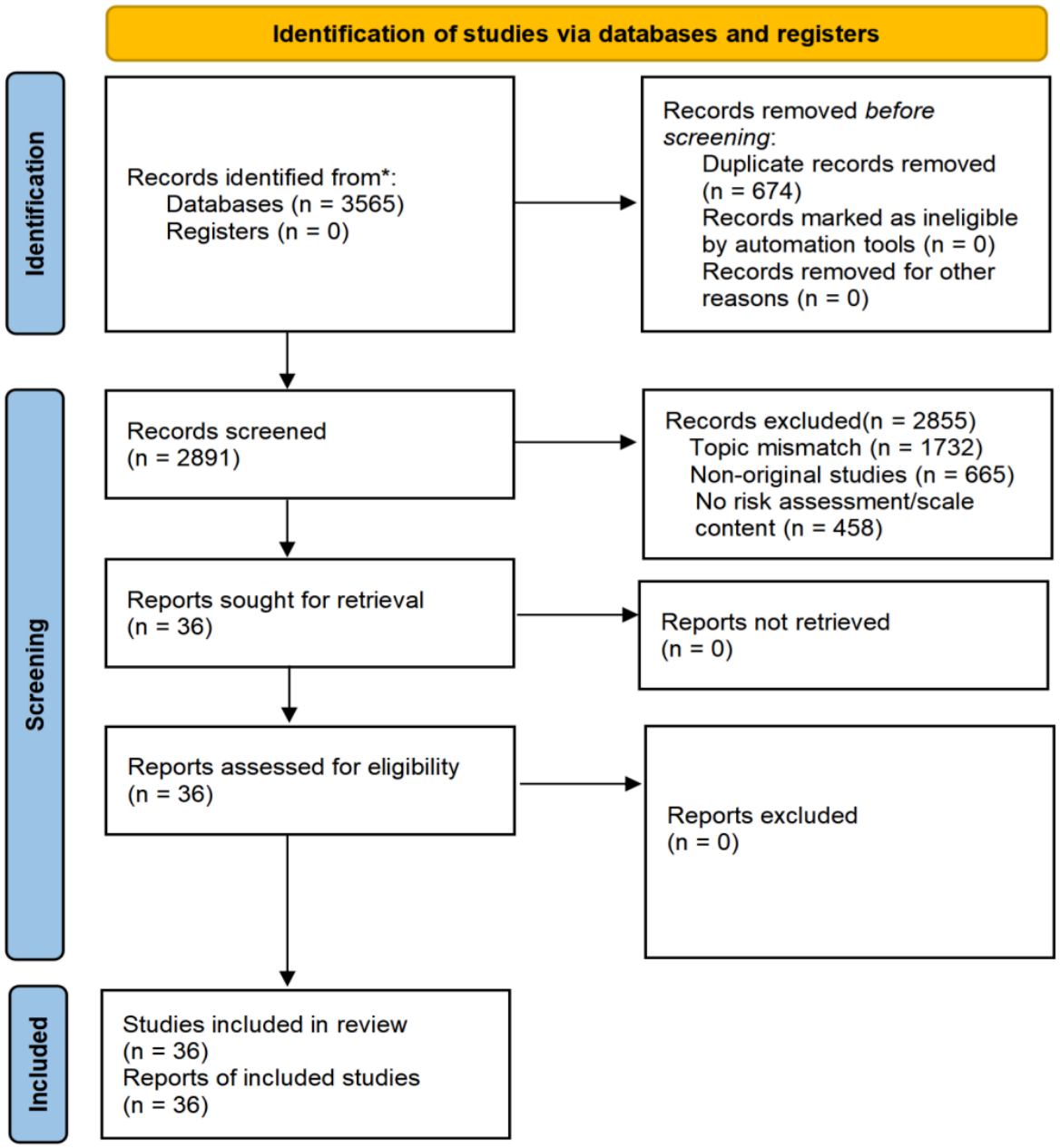


Figure 1

PRISMA 2020 flow diagram for new systematic reviews which included searches of databases and registers only

**Consider, if feasible to do so, reporting the number of records identified from each database or register searched (rather than the total number across all databases/registers).*

***If automation tools were used, indicate how many records were excluded by a human and how many were excluded by automation tools.*