

Early deep-RASS burden and in-hospital mortality in mechanically ventilated neurocritical patients: a retrospective landmark study

Jie Lu

Suzhou Ninth People's Hospital

Xiao Chen

Suzhou Ninth People's Hospital

Weikang Xing

xingweikang@alu.suda.edu.cn

Suzhou Ninth People's Hospital

Research Article

Keywords: Richmond Agitation–Sedation Scale, neurocritical care, mechanical ventilation, landmark analysis, in-hospital mortality.

Posted Date: October 11th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11009341/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

Serial Richmond Agitation–Sedation Scale (RASS) assessments capture evolving arousal, but neurological and treatment factors complicate their prognostic interpretation in neurocritical care. We evaluated the association between early deep-RASS burden and subsequent in-hospital mortality.

Methods

We conducted a retrospective 72-h landmark study using the Medical Information Mart for Intensive Care IV (MIMIC-IV) as the primary cohort and a semantically restricted eICU Collaborative Research Database (eICU-CRD) cohort for directional replication. Deep-RASS burden was the proportion of valid 4-h windows with $\text{RASS} \leq -3$ during the first 72 h; patients had to be alive and in the intensive care unit (ICU) at 72 h. MIMIC-IV used component-level multiple-imputation logistic regression and eICU used hospital-clustered generalised estimating equations. Sensitivity analyses examined landmark choice, measurement density, non-linearity, bootstrap stability, and hospital influence.

Results

The MIMIC-IV cohort included 2,252 patients and 541 deaths. Each 10-percentage-point increase in deep-RASS burden was associated with higher mortality odds (adjusted odds ratio [OR] 1.335, 95% confidence interval [CI] 1.289–1.382). Estimates were similar in sensitivity analyses; all 500 bootstrap estimates exceeded 1, with no clear evidence of non-linearity. The eICU cohort included 250 patients, 46 deaths, and 12 hospitals. The hospital-clustered association was directionally consistent (OR 1.269, 95% CI 1.212–1.329); leave-one-hospital-out ORs ranged from 1.238 to 1.302.

Conclusions

Early deep-RASS burden was associated with subsequent mortality in MIMIC-IV and was directionally consistent in the smaller, semantically restricted eICU cohort. It may support risk recognition but does not estimate the causal effect of modifying sedation therapy.

Background

Sedation and arousal assessment guide care for mechanically ventilated adults. The Richmond Agitation–Sedation Scale (RASS), ranging from + 4 (combative) to – 5 (unarousable), provides a reproducible bedside measure [1, 2]. Serial measurements capture changes that a single assessment cannot. This distinction is especially important in neurocritical care, where comfort and ventilator tolerance must be balanced against neurological examination and, when needed, control of intracranial

pressure or seizures [3]. Early deep-RASS burden may therefore summarise evolving arousal state, but it should not be interpreted as sedative exposure alone.

Evidence from general intensive care supports a prognostic role for early and longitudinal RASS patterns. Deep sedation during the first 48 h of mechanical ventilation has been associated with delayed extubation and higher hospital and 180-day mortality [4]. Persistent low-RASS trajectories and lower early RASS variability have also been linked to adverse outcomes [5, 6]. These observational findings do not show that changing sedation depth changes mortality. Consistent with that distinction, a systematic review of randomised trials found no clear mortality difference between lighter and deeper sedation strategies [7]. Whether the same RASS patterns carry the same meaning in neurocritical care remains uncertain.

Extrapolation to neurocritical care is difficult because a low RASS score may reflect the primary brain injury, impaired motor or language responses, metabolic disturbance, seizures, or treatment for intracranial hypertension, in addition to sedative and analgesic drugs [3]. Recent studies have shown substantial between-centre variation in sedation intensity and drug choice after acute brain injury, while adjusted mortality or functional differences between sedation strategies have been inconsistent [8–10]. Routine electronic health records add two methodological concerns: outcomes may occur during the RASS ascertainment window, and similarly named sedation fields may not represent the same scale across databases. A neurocritical study must therefore treat RASS as a mixed prognostic state marker while explicitly addressing temporal ordering and measurement provenance.

We evaluated early deep-RASS burden (the proportion of valid assessment windows with $\text{RASS} \leq -3$) as a continuous prognostic state marker in mechanically ventilated adults with neurocritical illness. A 72-h landmark preserved temporal ordering, and record-level RASS source auditing preceded eICU directional replication. We hypothesised that greater burden would be associated with subsequent in-hospital mortality without implying a causal effect of sedation therapy.

Methods

Study design and data sources

We conducted a retrospective prognostic association study using two deidentified critical care databases. The Medical Information Mart for Intensive Care IV (MIMIC-IV) version 3.1 was used for the primary analysis. MIMIC-IV contains electronic health record data from Beth Israel Deaconess Medical Center and is distributed through PhysioNet under a credentialed data-use agreement [11, 12]. The eICU Collaborative Research Database version 2.0 (eICU-CRD), a multicentre database derived from United States intensive care units, was used for directional replication [13, 14]. The analysis was reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) and REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statements [15, 16]. Both databases contain deidentified data released under their respective

institutional review and data-use frameworks. The investigators completed the required credentialing and complied with the applicable data-use agreements.

Study population

The MIMIC-IV cohort included adults (≥ 18 years) with acute ischaemic stroke, intracerebral haemorrhage, subarachnoid haemorrhage, traumatic brain injury, or status epilepticus, identified from prespecified International Classification of Diseases, Ninth and Tenth Revisions, Clinical Modification (ICD-9-CM/ICD-10-CM) code groups. For multiple eligible admissions, we retained the first neurologically eligible ICU stay of the first eligible hospitalisation. Invasive ventilation was identified by procedure item 225792. We excluded neuromuscular disease, poisoning or overdose, and pregnancy. ICU admission age was reconstructed from anchor_age, anchor_year, and admission year.

The eICU base cohort used the same disease groups, mapped from diagnosis-table ICD strings. Day-1 invasive ventilation was defined by ventday1 or oobintubday1 in apachePredVar and cross-checked against respiratoryCare. Analyses were at unit-stay level because no verified person-level identifier linked separate stays. Ages > 89 years were coded as 90 years. Replication was restricted to unit stays meeting the strict RASS definition.

Landmark and risk-set definition

RASS measurements from the first 72 h defined exposure. The primary risk set comprised patients alive and in the index ICU beyond the 72-h landmark; death or ICU discharge exactly at 72 h occurred before risk-set entry. At least three valid RASS windows were required. The outcome was in-hospital death after the landmark and before discharge.

A 48-h landmark analysis was prespecified as a sensitivity analysis. The primary 72-h window was not selected on the basis of the observed effect estimate.

RASS provenance and exposure construction

The Richmond Agitation–Sedation Scale (RASS) is a 10-level bedside scale ranging from + 4 (combative) to – 5 (unarousable) [1]. In MIMIC-IV, RASS was extracted from chartevents using item 228096. Values outside the valid range of – 5 to + 4 were discarded. Measurements from ICU admission up to, but not including, 72 h were assigned to 18 consecutive half-open 4-h windows. If more than one measurement occurred within a window, the latest value was retained; no interpolation was performed across missing windows.

Deep-RASS was defined a priori as $\text{RASS} \leq -3$. Burden was calculated as the number of valid windows containing $\text{RASS} \leq -3$ divided by the total number of valid RASS windows and was modelled per 10-

percentage-point increase. Valid-window count adjusted for measurement density; thresholds of at least six and nine windows were examined in sensitivity analyses.

For eICU-CRD, RASS was reconstructed from nurseCharting records during the first 4,320 min. The strict layer retained labels explicitly identifying RASS or SEDATION SCORE. A generic Sedation Scale pathway was rejected because retained values were never negative and thus incompatible with RASS. Strict-layer values were restricted to - 5 to + 4 and assigned to the same 4-h windows; where both layers contributed, the latest record was retained. The source audit is shown in Additional file 1: Figure S1.

Covariates

The MIMIC-IV adjustment set included age at ICU admission, sex, body mass index (BMI), Charlson comorbidity index without age points, number of valid RASS windows, any standardised vasoactive infusion during the first 24 h, neurological disease group, MIMIC anchor-year group, and six traceable 0–24-h organ-severity source variables: the worst ratio of arterial oxygen tension to inspired oxygen fraction ($\text{PaO}_2/\text{FiO}_2$), minimum platelet count, maximum bilirubin, maximum creatinine, minimum mean arterial pressure, and minimum Glasgow Coma Scale (GCS) motor score. BMI was reconstructed from first-day height and weight measurements. Charlson comorbidities were derived using the Quan ICD-9-CM/ICD-10-CM algorithm, with age points removed because age was modelled separately [17].

We did not calculate complete Sequential Organ Failure Assessment (SOFA) because available data did not support the urine-output renal component or all official first-day dependencies. Six organ-severity variables were retained separately; definitions and missingness are in Additional file 2: Table S1.

For eICU-CRD, the replication model adjusted for age, sex, Acute Physiology and Chronic Health Evaluation IV (APACHE IV) score, number of valid RASS windows, and neurological disease group. APACHE IV was used as the database-specific severity measure [18]. Because the intracerebral haemorrhage subgroup contained five patients and no deaths, acute ischaemic stroke and traumatic brain injury were retained as separate categories and intracerebral haemorrhage, subarachnoid haemorrhage, and status epilepticus were combined as other neurological to avoid complete separation.

Missing data

Missing MIMIC-IV covariates were handled at the component level using 60 multiply imputed datasets generated by fully conditional specification with posterior draws. The imputation model included the outcome, deep-RASS burden, all adjustment variables, and indicator variables for disease and anchor-year groups. Each imputation used fully conditional iterative regression with posterior sampling and a maximum of 30 iterations. BMI was bounded to 10–80 kg/m², the no-age Charlson index to 0–30, and organ-severity source variables to non-negative values after imputation. The primary logistic regression was fitted in every completed dataset, and coefficients and variances were combined using Rubin's rules.

This approach followed the chained-equations framework while preserving uncertainty between completed datasets [19].

In eICU-CRD, APACHE IV was missing for 15 of 250 patients (6.0%). Thirty posterior-draw iterative imputations used the outcome and all replication-model variables; generalised estimating equation (GEE) estimates were combined with Rubin's rules. A complete-case model was a sensitivity analysis.

Statistical analysis

Continuous variables are medians (interquartile ranges [IQRs]) and categorical variables counts (%). Characteristics were described across deep-RASS burden quartiles without causal interpretation.

Primary MIMIC-IV associations came from multivariable logistic regression and are reported as odds ratios (ORs) with 95% confidence intervals (CIs) per 10-percentage-point increase. A four-degree-of-freedom natural cubic spline was compared with the linear term in each imputed dataset using likelihood-ratio statistics; P values across imputations were summarized rather than selecting one.

Sensitivity models used a 48-h landmark and at least six or nine valid windows. Stability was assessed with 500 patient-level bootstrap resamples of one deterministic completed dataset, retaining duplicate patients; the Rubin multiple-imputation interval remained primary. Apparent Brier score and calibration intercept/slope used predictions averaged across 60 datasets and were treated as in-sample diagnostics.

The eICU replication analysis used a binomial generalised estimating equation (GEE) with an exchangeable working correlation structure and hospital as the clustering unit [20]. The imputed model was primary for replication, with APACHE IV complete-case analysis as a sensitivity analysis. To assess whether the association was dominated by one centre, the 30 imputed models were refitted after excluding each eligible hospital in turn; no hospital was removed from the primary analysis. Full model specifications and sensitivity-analysis definitions are provided in Additional file 2: Table S2.

Exploratory two- to six-class trajectory models were assessed without using mortality to select classes. Because the Bayesian information criterion improved through the upper boundary and persistent-deep patterns were unstable across class counts and specifications, trajectory classes were supplementary only.

Analyses followed statistical analysis plan (SAP) version 5.1 and used Python 3.12.7, DuckDB 1.5.5, pandas 2.3.3, NumPy 2.4.6, scikit-learn 1.9.0, and statsmodels 0.14.6. We report two-sided 95% CIs and interpret findings as prognostic associations. The analysis plan and code documentation are provided in Additional file 3.

Manuscript preparation

During manuscript preparation, the authors used ChatGPT (OpenAI) for language editing, manuscript organisation, and document formatting. The authors reviewed and edited the content and take full responsibility for the publication.

Results

Cohort assembly

Among 28,173 MIMIC-IV admissions with eligible neurological diagnoses, 12,133 first adult eligible ICU stays remained; 4,616 had invasive ventilation. After clinical exclusions, 4,217 remained. Of these, 2,252 were alive, in the index ICU at 72 h, and had ≥ 3 valid RASS windows (Fig. 1a); 541 (24.0%) died subsequently.

In eICU-CRD, 3,983 patients met the neurological diagnosis and day-1 ventilation criteria and 3,479 remained after clinical exclusions. The record-level audit identified 365 patients with at least three strict explicit RASS or SEDATION SCORE windows. Of these, 250 patients from 12 hospitals remained in the ICU at the 72-h landmark and formed the replication cohort; 46 (18.4%) subsequently died (Fig. 1b). APACHE IV was complete for 235 patients from 11 hospitals, including 42 deaths.

RASS burden and cohort characteristics

MIMIC-IV patients had median age 65 years (IQR 52–77), 1,326 (58.9%) were male, and median valid RASS windows were 15 (IQR 13–17). Median deep-RASS burden was 31.3% (IQR 7.7–78.6%; Additional file 1: Figure S2). Disease groups comprised 721 acute ischaemic stroke, 465 intracerebral haemorrhage, 314 subarachnoid haemorrhage, 675 traumatic brain injury, and 77 status epilepticus; Table 1 presents characteristics by burden quartile.

Table 1
Characteristics of the MIMIC-IV landmark cohort by deep-RASS burden quartile

Variable	Overall	Q1	Q2	Q3	Q4
Patients, n	2252	578	559	561	554
Age at ICU admission, years	65.0 (52.0–77.0)	65.5 (52.0–77.0)	64.0 (53.0–76.5)	66.0 (52.0–78.0)	63.0 (49.3–75.0)
Male sex, n (%)	1326 (58.9%)	332 (57.4%)	333 (59.6%)	326 (58.1%)	335 (60.5%)
BMI, kg/m ²	25.2 (19.1–30.0)	25.0 (17.6–30.0)	25.2 (18.6–29.8)	25.4 (19.8–30.0)	25.6 (19.3–30.4)
Charlson comorbidity index, no age points	3.0 (1.0–4.0)	2.0 (1.0–4.0)	3.0 (1.0–4.0)	3.0 (1.0–5.0)	3.0 (1.0–4.0)
Minimum GCS motor score, 0–24 h	4.0 (1.0–5.0)	5.0 (4.0–6.0)	4.0 (1.0–5.0)	3.0 (1.0–4.0)	1.0 (1.0–4.0)
Deep-RASS windows, %	31.3 (7.7–78.6)	0.0 (0.0–5.9)	18.8 (12.5–23.5)	53.3 (41.7–66.7)	100.0 (92.9–100.0)
Subsequent in-hospital death, n (%)	541 (24.0%)	55 (9.5%)	71 (12.7%)	141 (25.1%)	274 (49.5%)
Continuous variables are presented as median (interquartile range); categorical variables as n (%). Quartiles were defined by the proportion of valid 4-h RASS windows with RASS ≤ –3 during the first 72 h and are descriptive only. APACHE IV, Acute Physiology and Chronic Health Evaluation IV; BMI, body mass index; GCS, Glasgow Coma Scale; ICU, intensive care unit; RASS, Richmond Agitation–Sedation Scale; Q1–Q4, deep-RASS burden quartiles.					

In the eICU replication cohort, median age was 57 years (IQR 38–69), 144 patients (57.6%) were male, and the median number of strict RASS windows was 18 (IQR 14–18). Median deep-RASS burden was 22.9% (IQR 5.6–65.6%). The cohort included 73 patients with acute ischaemic stroke, 5 with intracerebral haemorrhage, 17 with subarachnoid haemorrhage, 136 with traumatic brain injury, and 19 with status epilepticus (Additional file 2: Table S3).

Primary MIMIC-IV association

In the component-level multiple-imputation model, each 10-percentage-point increase in deep-RASS burden was associated with higher odds of subsequent in-hospital death (adjusted OR 1.335, 95% CI 1.289–1.382; Table 2 and Fig. 2). Figure 3 shows the observed burden distribution and the adjusted linear exposure-response function. This estimate describes the association of a higher proportion of deep-RASS assessment windows with mortality among patients who survived and remained in the ICU to the 72-h landmark.

Table 2
Associations between deep-RASS burden and subsequent in-hospital mortality

Analysis	OR	95% CI lower	95% CI upper	Model or handling
MIMIC-IV, 72-h primary	1.335	1.289	1.382	60 component-level imputations; traceable organ-severity source variables
MIMIC-IV, 48-h landmark	1.318	1.278	1.358	60 component-level imputations
MIMIC-IV, ≥ 6 RASS windows	1.333	1.288	1.381	60 component-level imputations
MIMIC-IV, ≥ 9 RASS windows	1.330	1.283	1.378	60 component-level imputations
eICU, complete-case GEE	1.262	1.188	1.341	235 complete APACHE IV cases; 11 hospitals
eICU, APACHE IV MI GEE	1.269	1.212	1.329	30 imputations; 12 hospitals
The exposure was the proportion of valid 4-h RASS windows with $\text{RASS} \leq -3$ during the first 72 h; ORs are reported per 10-percentage-point increase. The MIMIC-IV model used component-level imputation of traceable 0–24-h organ-severity source variables and did not use a complete SOFA score. The eICU model used hospital-clustered GEE. All estimates are observational prognostic associations and do not represent a causal effect of sedation depth. OR, odds ratio; CI, confidence interval; GEE, generalized estimating equations; MI, multiple imputation; APACHE IV, Acute Physiology and Chronic Health Evaluation IV. SOFA, Sequential Organ Failure Assessment; RASS, Richmond Agitation–Sedation Scale; ICU, intensive care unit.				

MIMIC-IV sensitivity and model-stability analyses

At 48 h, the OR was 1.318 (95% CI 1.278–1.358; $n = 2,696$; 674 deaths; Fig. 2). Requiring ≥ 6 or ≥ 9 valid windows gave ORs of 1.333 (95% CI 1.288–1.381; $n = 2,231$; 536 deaths) and 1.330 (95% CI 1.283–1.378; $n = 2,136$; 508 deaths), respectively.

Across 60 datasets, the median likelihood-ratio P value for non-linearity was 0.111, with none < 0.05 ; this indicated no clear evidence of non-linearity. All 500 bootstrap models converged (median OR 1.339; 2.5th–97.5th percentile interval 1.301–1.392), and all estimates exceeded 1.

eICU directional replication and hospital influence

In the multiply imputed model, after adjustment for age, sex, APACHE IV, RASS measurement density, and disease group, each 10-percentage-point increase in strict-layer deep-RASS burden was associated with higher odds of subsequent in-hospital death (GEE OR 1.269, 95% CI 1.212–1.329; Table 2 and Fig. 2). The complete-case estimate was similar ($n = 235$; 42 deaths; OR 1.262, 95% CI 1.188–1.341).

Eleven hospitals were eligible for leave-one-hospital-out refitting. The pooled ORs after individual hospital exclusion ranged from 1.238 to 1.302, and all 11 estimates were greater than 1 (Additional file 1: Figure S3 and Additional file 2: Table S4). The positive direction did not disappear after exclusion of any single assessed hospital, but this analysis did not eliminate hospital-level confounding.

Auxiliary model diagnostics and trajectory assessment

The apparent MIMIC-IV model diagnostics were as follows (Brier score 0.1399; mean predicted and observed risk 0.2402; intercept approximately 0; slope 1.007; Additional file 1: Figure S4) and describe in-sample performance rather than independent validation.

Trajectory screening did not yield a stable primary class structure: model fit improved from two through six classes and persistent-deep patterns were specification-sensitive. Trajectories were therefore retained only as supplementary findings (Additional file 1: Figure S5).

Discussion

Greater deep-RASS burden during the first 72 h was associated with subsequent in-hospital mortality in MIMIC-IV, where adjusted death odds increased by 33.5% per 10-percentage-point increase. The estimate was similar at the 48-h landmark, with higher measurement-density thresholds, spline assessment, and bootstrap analyses. The eICU association remained directionally consistent after APACHE IV imputation and hospital exclusion. These results support deep-RASS burden as a prognostic marker among ventilated neurocritical patients alive and in the ICU at the landmark.

Our findings are consistent with, but add a distinct neurocritical-care perspective to, previous work in general critical care. Shehabi and colleagues linked deep sedation during the first 48 h of mechanical ventilation to delayed extubation and higher mortality [4]. Hyun et al. identified persistent low-RASS trajectories associated with adverse outcomes, and Liao et al. reported an association between lower early RASS variability and 28- and 90-day mortality [5, 6]. Effect estimates cannot be compared directly because the populations, exposure summaries, and follow-up periods differed. The present study focused on heterogeneous neurocritical illnesses, used a continuous deep-RASS burden, separated exposure ascertainment from subsequent mortality with a landmark risk set, and audited RASS semantics before cross-database replication. Contemporary ICU guidance emphasizes individualized sedation targets, whereas evidence on implementation strategies and outcome effects remains heterogeneous [21, 22].

In neurocritical care, deep-RASS burden is broader than sedation depth alone. A persistently low score may reflect brain injury, impaired responses, metabolic dysfunction, seizures, sedative and analgesic treatment, and therapies for intracranial hypertension [3]. Recent neurocritical studies support this interpretation: TRACK-TBI reported centre-level variation in sedation and intracranial pressure-directed treatment [8], the ENIO cohort documented heterogeneous sedative practice with neutral adjusted mortality comparisons [9], and a smaller intracranial haemorrhage study found no clear outcome

differences by sedation exposure [10]. Deep-RASS burden is useful because it captures this integrated bedside state. A high burden should prompt assessment of neurological deterioration, seizures, metabolic disturbance, treatment targets, and medication exposure, rather than an automatic label of excessive sedation. Recent TRACK-TBI work similarly evaluated early sedative exposure alongside clinical and functional outcomes, underscoring the observational nature of treatment-selection comparisons [23].

The landmark and semantic audit addressed different sources of bias: temporal ordering and measurement validity. For exposures accumulated over time, landmark or time-dependent approaches preserve the ordering of exposure and outcome [24]. The 72-h landmark therefore estimated prognosis among patients alive and still in the index ICU after RASS ascertainment. The eICU audit showed that similarly named electronic health record fields did not necessarily encode the same scale. Excluding a generic pathway with no negative values reduced sample size but improved construct validity, in line with RECORD principles for transparent derivation of routinely collected variables [16]. Replication in this stricter layer indicates that the association was not confined to the MIMIC implementation.

Exploratory trajectory analyses did not identify a stable class structure across candidate specifications [25]. We therefore retained trajectory patterns as descriptive findings and used continuous deep-RASS burden as the primary exposure, summarising cumulative low-arousal observations without relying on uncertain class assignment. This measure yields an effect estimate that is directly interpretable per 10-percentage-point increase in burden (Additional file 1: Figure S5).

Clinically, the finding is useful for risk recognition, not for imposing a universal sedation target. Repeated deep-RASS observations identify patients with higher subsequent mortality risk and can be calculated from routine bedside assessments. A high burden may justify earlier clinical review of neurological status, treatment indications, sedative and analgesic exposure, and the documented RASS target. Randomised evidence has not established a mortality advantage for lighter over deeper sedation strategies [7], and the present association does not determine whether changing sedation would improve outcome. Prospective studies should collect actual and target RASS, time-varying drug doses, intracranial pressure and seizure indications, and functional outcomes to distinguish modifiable sedation burden from low arousal caused by the underlying neurological condition.

Several design safeguards strengthen the interpretability of our findings. We reconstructed exposure timing from source-level records, separated the 72-h exposure-ascertainment period from subsequent outcome follow-up using a landmark estimand, handled covariate missingness at the component level with multiple imputation, conducted targeted sensitivity analyses, audited eICU RASS semantics at the record level, and accounted for within-hospital correlation in the replication analysis. These procedures are consistent with established guidance for routinely collected health data and with accepted approaches to landmark analysis, multiple imputation, and correlated outcomes [15, 16, 19, 20, 24, 26].

Residual confounding by neurological severity and treatment indication remains possible because detailed imaging, pupillary findings, electroencephalography, target RASS, and time-varying drug doses

were unavailable. This limitation is particularly relevant in neurocritical care, where low RASS values may reflect both treatment exposure and the underlying neurological condition [3, 8–10]. A complete urine-output-based SOFA score could not be constructed because the required source data were unavailable; although traceable organ-severity variables were adjusted separately, this should not be interpreted as equivalent to full SOFA adjustment [27].

The strict eICU semantic definition improved construct validity but reduced sample size and hospital coverage. With 12 hospitals, cluster-level inference and transportability remain limited. The estimand applies only to patients alive and in the ICU at 72 h, and the outcome was in-hospital mortality rather than long-term neurological function.

Future prospective studies should pair serial RASS assessments with treatment targets, time-varying drug doses, neurological examination findings, and physiological indications to distinguish potentially modifiable sedation burden from low arousal caused by the underlying neurological condition. A recent neurocritical-care review also emphasizes that sedation interruption should be individualized and balanced against patient-specific risks [28].

Conclusions

Early deep-RASS burden was associated with subsequent in-hospital mortality in ventilated neurocritical patients and was directionally consistent in eICU. The framework combines continuous burden, landmark risk-set definition, and source auditing for prognostic assessment of repeated RASS.

Abbreviations

APACHE IV

Acute Physiology and Chronic Health Evaluation IV

BIC

Bayesian information criterion

BMI

body mass index

CI

confidence interval

eICU-CRD

eICU Collaborative Research Database

FiO₂

inspired oxygen fraction

GCS

Glasgow Coma Scale

GEE

generalised estimating equation

ICD-9-CM/ICD-10-CM

International Classification of Diseases,Ninth/Tenth Revision,Clinical Modification

ICU

intensive care unit

IQR

interquartile range

MAP

mean arterial pressure

MI

multiple imputation

MIMIC-IV

Medical Information Mart for Intensive Care IV

OR

odds ratio

PaO₂

arterial oxygen tension

RASS

Richmond Agitation–Sedation Scale

RECORD

REporting of studies Conducted using Observational Routinely-collected health Data

SAP

statistical analysis plan

SOFA

Sequential Organ Failure Assessment

STROBE

Strengthening the Reporting of Observational Studies in Epidemiology.

Declarations

Ethics approval and consent to participate

This study used de-identified secondary records from MIMIC-IV and eICU-CRD, accessed under credentialed data-use agreements after the required training. The analysis was conducted in accordance with the principles of the Declaration of Helsinki and the applicable data-governance and data-use requirements of the source databases. MIMIC-IV data collection and sharing were approved by the Institutional Review Board at Beth Israel Deaconess Medical Center, which waived informed consent. The eICU-CRD data were provided under the governance arrangements described in the database documentation. No patients were contacted, no intervention was performed, and no new individual consent was obtained. No additional study-specific ethics approval was obtained for this secondary analysis.

Consent for publication

Not applicable. The manuscript contains no identifiable individual information.

Competing interests

The authors declare that they have no competing interests.

Additional files

Additional file 1 (.docx). Supplementary figures. Figures S1–S5 document the RASS source audit, measurement completeness, hospital influence, apparent calibration, and exploratory trajectory assessment.

Additional file 2 (.docx). Supplementary tables. Tables S1–S4 provide missingness and variable definitions, model specifications, eICU cohort characteristics, and hospital-exclusion analyses.

Additional file 3 (.zip). Non-sensitive reproducibility materials. Analysis scripts, SAP, data dictionary, software specifications, and aggregate quality-control documentation. The archive contains no patient-level records.

Funding

Supported by the Open Research Fund of Jiangsu Province Key Laboratory of Brain Disease Bioinformatics, Xuzhou Medical University (grant no. XZSYSKF2025016, to X.C.). The funder had no role beyond financial support.

Author Contribution

J.L.: Visualization, Writing – original draft. X.C.: Formal analysis, Funding acquisition, Writing – review and editing. W.X.: Conceptualization, Data curation, Methodology, Resources, Supervision, Validation, Writing – review and editing. All authors read and approved the final manuscript. J.L. and X.C. contributed equally to this work.

Acknowledgements

Not applicable.

Data Availability

MIMIC-IV version 3.1 (<https://doi.org/10.13026/kpb9-mt58>) and eICU-CRD version 2.0 (<https://doi.org/10.13026/C2WM1R>) are available through PhysioNet to credentialed users who complete the required training and accept the applicable data-use agreements. The authors are not permitted to redistribute patient-level records. The non-sensitive analysis scripts, SAP, data dictionary, software specifications, and aggregate quality-control documentation are available at <https://doi.org/10.17605/OSF.IO/VWQT9> and in Additional file 3. No restricted patient-level data are included in that package.

References

1. Sessler CN, Gosnell MS, Grap MJ, Brophy GM, O'Neal PV, Keane KA, et al. The Richmond Agitation–Sedation Scale: Validity and Reliability in Adult Intensive Care Unit Patients. *Am J Respir Crit Care Med*. 2002;166:1338–44. <https://doi.org/10.1164/rccm.2107138>.
2. Devlin JW, Skrobik Y, Gélinas C, Needham DM, Slooter AJC, Pandharipande PP, et al. Clinical Practice Guidelines for the Prevention and Management of Pain, Agitation/Sedation, Delirium, Immobility, and Sleep Disruption in Adult Patients in the ICU. *Crit Care Med*. 2018;46:e825–73. <https://doi.org/10.1097/ccm.0000000000003299>.
3. Prabhakar H, Tripathy S, Gupta N, Singhal V, Mahajan C, Kapoor I, et al. Consensus Statement on Analgo-sedation in Neurocritical Care and Review of Literature. *Indian J Crit Care Med*. 2021;25:126–33. <https://doi.org/10.5005/jp-journals-10071-23712>.
4. Shehabi Y, Chan L, Kadiman S, Alias A, Ismail WN, Tan MATI, et al. Sedation depth and long-term mortality in mechanically ventilated critically ill adults: a prospective longitudinal multicentre cohort study. *Intensive Care Med*. 2013;39:910–8. <https://doi.org/10.1007/s00134-013-2830-2>.
5. Hyun D, Ahn JH, Gil H-Y, Nam CM, Yun C, Lim C-M. Longitudinal trajectories of sedation level and clinical outcomes in patients who are mechanically ventilated based on a group-based trajectory model: a prospective, multicentre, longitudinal and observational study in Korea. *BMJ Open*. 2023;13:e072628. <https://doi.org/10.1136/bmjopen-2023-072628>.
6. Liao S-F, Huang C-J, Lee R-P, Chen T-Y, Wang H-C. Sedation level variability as an indicator for mortality in mechanically ventilated critically ill patients: a propensity score-weighted cohort study. *Front Med*. 2026;13. <https://doi.org/10.3389/fmed.2026.1755288>.
7. Ceric A, Holgersson J, May TL, Skrifvars MB, Hästbacka J, Saxena M, et al. Effect of level of sedation on outcomes in critically ill adult patients: a systematic review of clinical trials with meta-analysis and trial sequential analysis. *eClinicalMedicine*. 2024;71:102569. <https://doi.org/10.1016/j.eclinm.2024.102569>.
8. Dolmans RGF, Barber J, Foreman B, Temkin NR, Okonkwo DO, Robertson CS, et al. Sedation Intensity in Patients with Moderate to Severe Traumatic Brain Injury in the Intensive Care Unit: A TRACK-TBI Cohort Study. *Neurocrit Care*. 2025;42:551–61. <https://doi.org/10.1007/s12028-024-02054-7>.
9. Feng SN, Laws LH, Diaz-Cruz C, Cinotti R, Schultz MJ, Asehnoune K, et al. Sedation Practices in Mechanically Ventilated Neurocritical Care Patients from 19 Countries: An International Cohort

- Study. *Neurocrit Care*. 2025;43:232–42. <https://doi.org/10.1007/s12028-024-02200-1>.
10. Kim HJ, Kwon Y, Seok J, Roh H, Byun J, Yoon W, et al. Impact of sedation levels on outcomes in neurocritical care patients with intracranial hemorrhage: a retrospective cohort study. *Neurosurg Rev*. 2025;48. <https://doi.org/10.1007/s10143-025-03507-z>.
 11. Johnson AEW, Bulgarelli L, Shen L, Gayles A, Shammout A, Horng S et al. MIMIC-IV, a freely accessible electronic health record dataset. *Scientific Data* 2023;10. <https://doi.org/10.1038/s41597-022-01899-x>
 12. Johnson A, Bulgarelli L, Pollard T, Gow B, Moody B, Horng S et al. MIMIC-IV. *PhysioNet* 2024. <https://doi.org/10.13026/kpb9-mt58>
 13. Pollard TJ, Johnson AEW, Raffa JD, Celi LA, Mark RG, Badawi O. The eICU Collaborative Research Database, a freely available multi-center database for critical care research. *Scientific Data* 2018;5. <https://doi.org/10.1038/sdata.2018.178>
 14. Pollard T, Johnson A, Raffa J, Celi LA, Badawi O, Mark R. eICU Collaborative Research Database. *PhysioNet*. 2019. <https://doi.org/10.13026/C2WM1R>
 15. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for Reporting Observational Studies. *PLoS Med*. 2007;4:e296. <https://doi.org/10.1371/journal.pmed.0040296>.
 16. Benchimol EI, Smeeth L, Guttman A, Harron K, Moher D, Petersen I, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLoS Med*. 2015;12:e1001885. <https://doi.org/10.1371/journal.pmed.1001885>.
 17. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi J-C, et al. Coding Algorithms for Defining Comorbidities in ICD-9-CM and ICD-10 Administrative Data. *Med Care*. 2005;43:1130–9. <https://doi.org/10.1097/01.mlr.0000182534.19832.83>.
 18. Zimmerman JE, Kramer AA, McNair DS, Malila FM. Acute Physiology and Chronic Health Evaluation (APACHE) IV: Hospital mortality assessment for today's critically ill patients. *Crit Care Med*. 2006;34:1297–310. <https://doi.org/10.1097/01.ccm.0000215112.84523.f0>.
 19. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. *J Stat Softw*. 2011;45. <https://doi.org/10.18637/jss.v045.i03>.
 20. LIANG K-Y, ZEGER SL. Longitudinal data analysis using generalized linear models. *Biometrika*. 1986;73:13–22. <https://doi.org/10.1093/biomet/73.1.13>.
 21. Lewis K, Balas MC, Stollings JL, McNett M, Girard TD, Chanques G, et al. A Focused Update to the Clinical Practice Guidelines for the Prevention and Management of Pain, Anxiety, Agitation/Sedation, Delirium, Immobility, and Sleep Disruption in Adult Patients in the ICU. *Crit Care Med*. 2025;53:e711–27. <https://doi.org/10.1097/CCM.0000000000006574>.
 22. Hume NE, Zervas I, Wong A, Klein-Fedyshin M, Smithburger PL, Buckley MS, et al. Clinical Impact of the Implementation Strategies Used to Apply the 2013 Pain, Agitation/Sedation, Delirium or 2018 Pain, Agitation/Sedation, Delirium, Immobility, Sleep Disruption Guideline Recommendations: A

- Systematic Review and Meta-Analysis. *Crit Care Med*. 2024;52:626–36. <https://doi.org/10.1097/CCM.0000000000006178>.
23. Liu SY, Kelly-Hedrick M, Temkin N, Barber J, Komisarow J, Hatfield J, et al. Association of Early Dexmedetomidine Utilization With Clinical and Functional Outcomes Following Moderate-Severe Traumatic Brain Injury: A Transforming Clinical Research and Knowledge in Traumatic Brain Injury Study. *Crit Care Med*. 2024;52:607–17. <https://doi.org/10.1097/CCM.0000000000006106>.
24. Zabor EC, Assel M. On the Need for Landmark Analysis or Time-dependent Covariates. *J Urol*. 2023;209:1060–2. <https://doi.org/10.1097/ju.0000000000003459>.
25. Nagin DS, Jones BL, Elmer J. Recent Advances in Group-Based Trajectory Modeling for Clinical Research. *Ann Rev Clin Psychol*. 2024;20:285–305. <https://doi.org/10.1146/annurev-clinpsy-081122-012416>.
26. Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*. 2009;338:b2393. <https://doi.org/10.1136/bmj.b2393>.
27. Vincent J-L, Moreno R, Takala J, Willatts S, De Mendonça A, Bruining H, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. *Intensive Care Med*. 1996;22:707–10. <https://doi.org/10.1007/BF01709751>.
28. Reyes-Bello JS, Moscote-Salazar LR, Janjua T. Sedation Vacations in Neurocritical Care: Friend or Foe? *Curr Neurol Neurosci Rep*. 2024;24:671–80. <https://doi.org/10.1007/s11910-024-01383-6>.

Figures

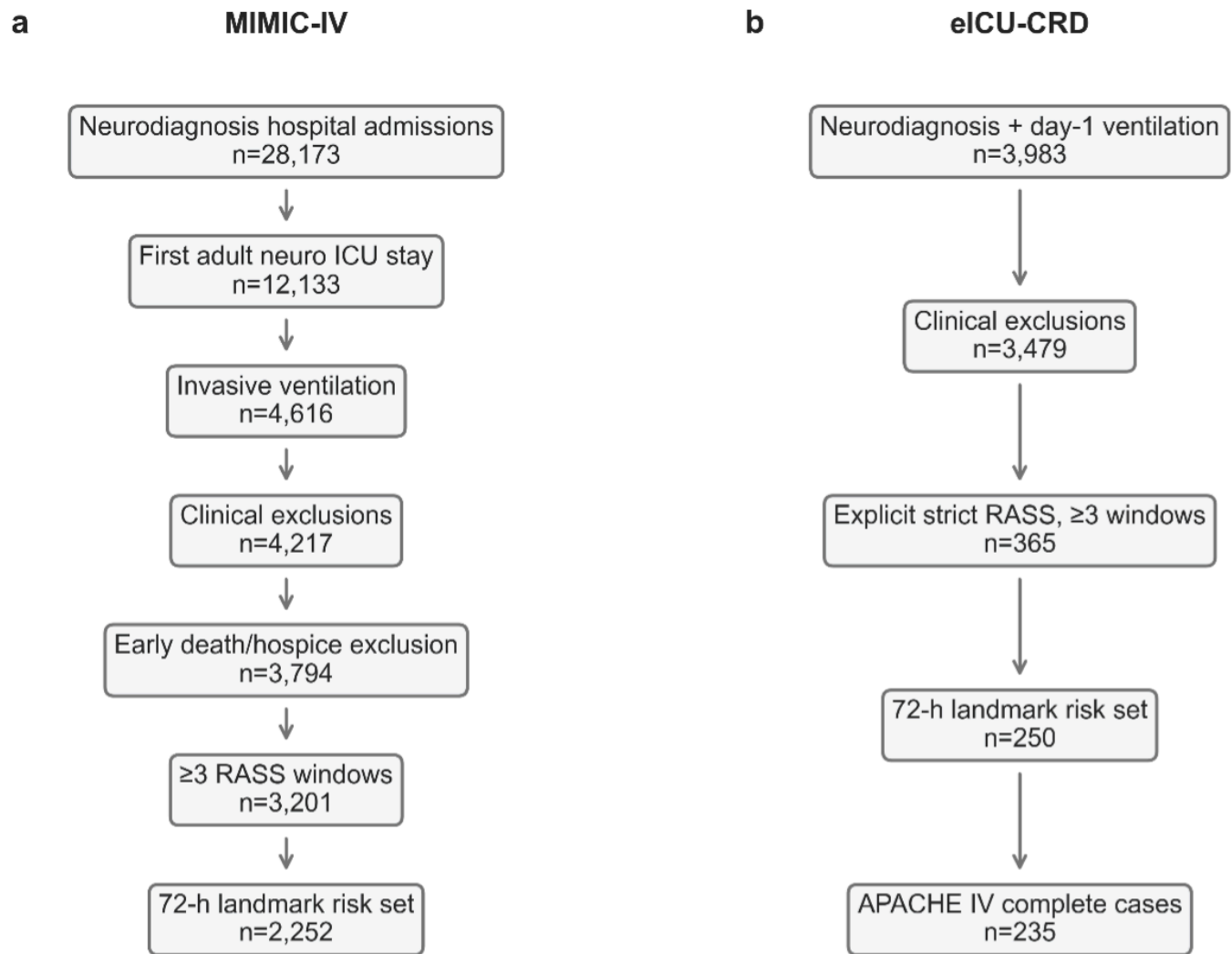


Figure 1

Cohort assembly and 72-h landmark restriction. a, MIMIC-IV cohort assembly from neurological-diagnosis hospital admissions through the first adult neurologically eligible ICU stay, invasive mechanical ventilation, clinical exclusions, and at least three RASS windows to the 72-h landmark risk set. b, eICU-CRD cohort assembly from the neurological diagnosis plus day-1 ventilation cohort (n=3,983), through clinical exclusions (n=3,479), to the strict layer with at least three records explicitly labelled RASS or SEDATION SCORE (n=365), the 72-h risk set (n=250), and APACHE IV complete cases (n=235). The generic Sedation Scale pathway was excluded because it contained no negative values. Screening counts refer to admissions or ICU stays, as applicable; eICU replication was conducted at the unit-stay level. ICU, intensive care unit; RASS, Richmond Agitation–Sedation Scale; APACHE IV, Acute Physiology and Chronic Health Evaluation IV.

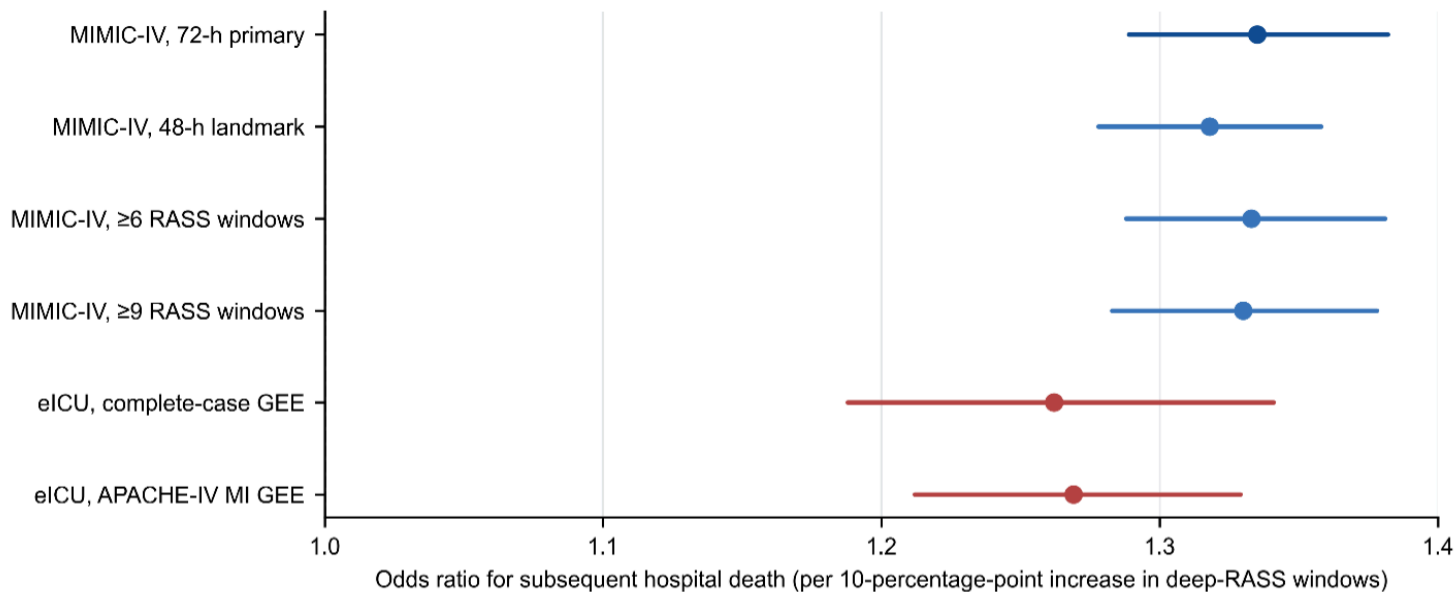


Figure 2

Associations between deep-RASS burden and subsequent in-hospital mortality. Forest plot of MIMIC-IV primary, 48-h landmark, and RASS measurement-density analyses, together with strict-layer eICU replication. Points denote odds ratios and horizontal lines 95% confidence intervals; the left axis marks OR=1. The exposure was the proportion of valid 4-h windows with RASS ≤ -3 , reported per 10-percentage-point increase. MIMIC-IV models used 60 component-level imputations. eICU estimates show complete-case and APACHE IV-imputed hospital-clustered GEE models. OR, odds ratio; GEE, generalised estimating equation; APACHE IV, Acute Physiology and Chronic Health Evaluation IV; RASS, Richmond Agitation–Sedation Scale.

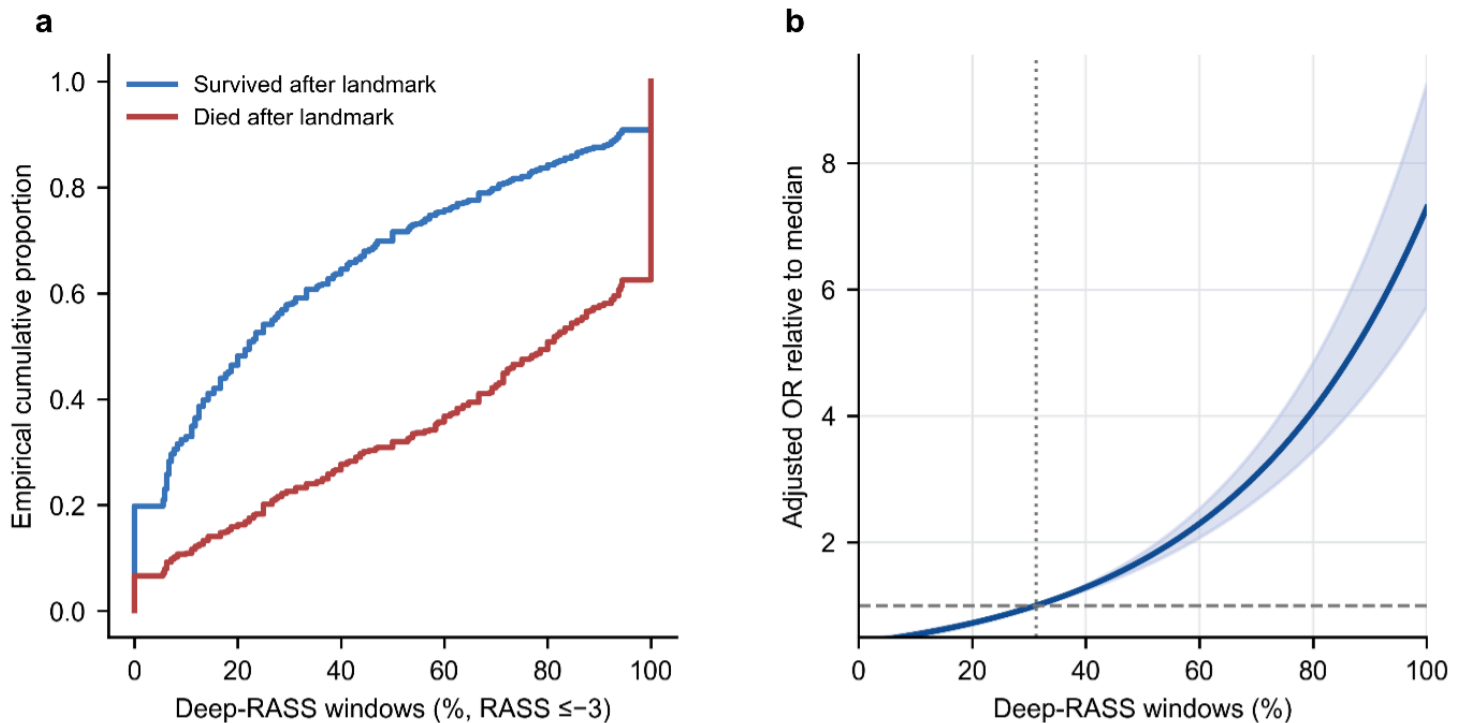


Figure 3

Deep-RASS burden distribution and adjusted exposure-response. a, Distribution of the proportion of deep-RASS windows during the first 72 h, stratified by survival status after the landmark. b, Adjusted odds ratios and 95% confidence intervals from the linear MIMIC-IV primary model, shown relative to the median deep-RASS burden. The horizontal dashed line denotes OR=1 and the vertical dotted line the median reference. This figure displays exposure distribution and model form; it does not estimate a causal effect of sedation. OR, odds ratio; RASS, Richmond Agitation–Sedation Scale.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile2.docx](#)
- [Additionalfile3.zip](#)
- [Additionalfile1.docx](#)