MHJ

LUMORA JOURNALS

Multidisciplinary Healthcare Journal

Vol. 1, No. 3 (2025) · e-ISSN 3152-8775

Original research (retrospective secondary cohort analysis) · DOI: not registered

Clinical seizure presentation and 60-day mortality among adults selected for continuous EEG in a Saudi intensive care unit: a secondary cohort analysis

Fatema Alawad[1],*, Abdullah Naji[2], Abdullah Nimer[3], Abdullah Talib Mohammed Alkathiri[4], Abdulrahman Al Zhrani[5], Abdulrahman Mohammed Aldayel[6]

Submission date	Acceptance date	Publication date
20 March 2025	25 March 2025	31 March 2025

[1] College of Medicine, King Faisal University, Al-Ahsa, Saudi Arabia

[2] Oregon Health Science University Hospital, Portland, United States

[3] The University of Jordan, Amman, Jordan

[4] King Abdulahziz University

[5] King Faisal Specialist Hospital and Research Centre, Riyadh, Saudi Arabia

[6] Imam Mohammad Ibn Saud Islamic University, Saudi Arabia

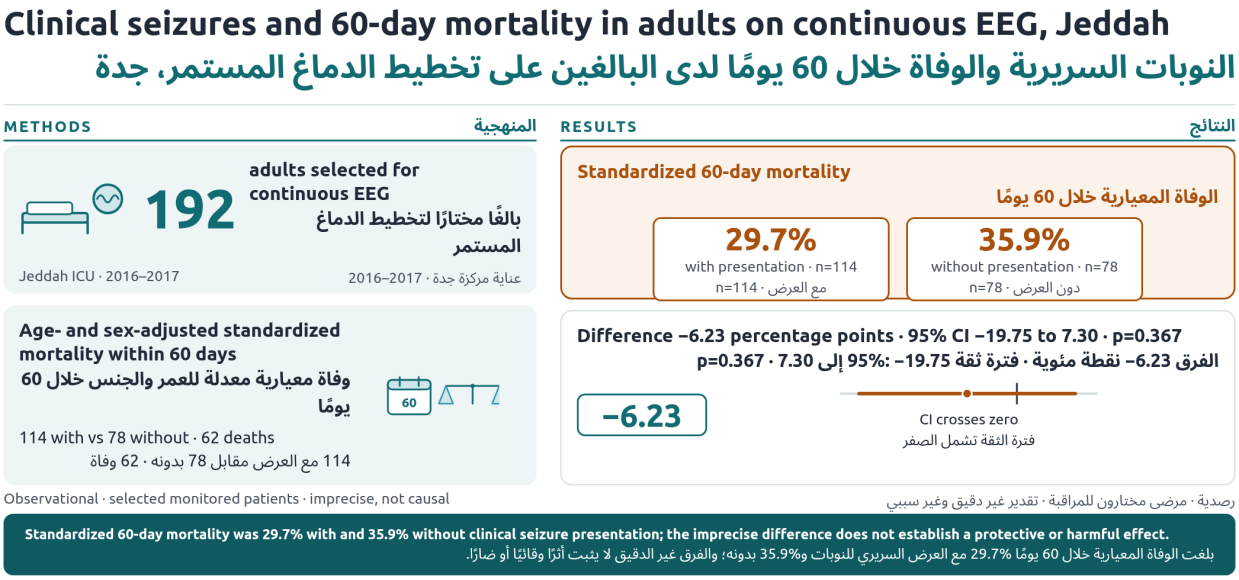
* Correspondence: Fatema Alawad, fatemalumora@gmail.com

English edition. An Arabic edition of this article is available.

Cite as: Alawad F, Naji A, Nimer A, Alkathiri ATM, Al Zhrani A, Aldayel AM. Clinical seizure presentation and 60-day mortality among adults selected for continuous EEG in a Saudi intensive care unit: a secondary cohort analysis. Multidisciplinary Healthcare Journal. 2025;1(3):MHJ-2026-000347. © The authors. Published open access under the Creative Commons Attribution 4.0 International licence (CC BY 4.0), https://creativecommons.org/licenses/by/4.0/.

Abstract

Graphical abstract



Clinical seizures at presentation and seizures subsequently detected on continuous electroencephalography (CEEG) describe different stages of an ICU patient's course. We examined whether recorded clinical seizure presentation was associated with mortality within 60 days of admission among adults selected for CEEG in Jeddah, Saudi Arabia. Public patient-level data from King Abdulaziz University Hospital during September 2016–August 2017 were reanalysed under a frozen plan. Age- and sex-adjusted logistic predictions were standardized over eligible adults. Among 192 eligible adult monitoring records, 114 had recorded clinical seizure presentation and 78 did not; 62 deaths were recorded. Unique-person and episode independence could not be verified from the supplied identifiers. Standardized mortality was 29.7% with and 35.9% without presentation, a difference of -6.23 percentage points (95% confidence interval, -19.75 to 7.30; p=0.367). A prespecified Firth model retaining all nominal diagnosis codes estimated a difference of -4.11 percentage points (bootstrap interval, -19.42 to 10.19). Diagnosis labels and totals did not fully reconcile with the source publication. These imprecise estimates do not establish a protective or harmful effect of clinical seizure presentation. Selection for monitoring, unmeasured illness severity, incomplete timing, and unresolved source coding limit interpretation. The study supplies a reproducible local reanalysis and a specification for better prospective data capture; it does not estimate the effectiveness of CEEG or antiseizure treatment.

Keywords: Saudi Arabia; neurocritical care; clinical seizures; continuous electroencephalography; mortality; retrospective cohort; marginal standardization.

1. Introduction

Clinical seizures are an important reason for neurological evaluation during critical illness, but a seizure observed at presentation and a nonconvulsive seizure detected later by EEG are not the same exposure. Their timing, ascertainment, and clinical context differ. A public dataset from a Saudi ICU records both clinical seizure presentation and findings obtained during continuous electroencephalography (CEEG) [1]. The original study emphasized monitoring yield and associations of electrographic patterns with outcomes.

The present analysis examines the earlier recorded clinical presentation, while retaining the limitations imposed by studying only patients who were subsequently selected for monitoring.

Consensus recommendations describe CEEG as a method for identifying nonconvulsive seizures, characterizing paroxysmal events, and assessing selected critically ill patients with altered consciousness [2,4]. Its successful implementation also depends on personnel, equipment, interpretation, and operational procedures [3]. These considerations influence which patients enter an EEG cohort and when monitoring begins. Consequently, a monitored population is not a random sample of all ICU admissions. The reasons for referral may differ between patients with overt clinical seizures and those whose main problem is persistent unexplained encephalopathy. A comparison between those groups must account for this selection context before it is interpreted prognostically.

Earlier work on seizure detection identified clinical events before monitoring among characteristics associated with electrographic seizures [5]. Other investigators have described the acute outcomes of patients presenting with seizures and directly admitted from emergency care to an ICU [6]. These questions are adjacent to, but distinct from, asking whether recorded seizure presentation separates mortality risks in a cohort already referred for CEEG. An older monitored-seizure cohort also examined seizure duration and diagnostic delay in relation to mortality, illustrating that a binary presentation indicator omits clinically relevant dimensions [7]. None of these different exposures can be substituted for another without changing the question.

The distinction between prognostic association and monitoring effectiveness is particularly important. The CERTA randomized trial compared continuous with repeated routine EEG among critically ill adults with impaired consciousness and no recent seizure [8]. Subsequent analyses examined electrographic prognostic features and the post-cardiac-arrest subgroup [9,10]. These studies address monitoring strategy or EEG characteristics under defined conditions. Our observational dataset has no unmonitored comparator and no randomized treatment allocation. It cannot estimate the effect of introducing a CEEG service, extending monitoring, or treating a detected pattern. A mortality contrast between presentation groups should therefore not be presented as evidence for or against those interventions.

Outcome studies in patients with persistent unresponsiveness after sedation interruption, septic encephalopathy, and nonhypoxic encephalopathy further show that the clinical setting in which EEG is obtained matters [11–13]. Their populations, measures, and follow-up origins differ. Saudi observational evidence from another hospital has also described clinical correlates of EEG patterns [14], but it does not validate every exposure or outcome code in the present source. Locally relevant research benefits from acknowledging this heterogeneity rather than treating all EEG cohorts as measurements of a single universal relationship.

Our objective was to estimate the age- and sex-adjusted difference in recorded mortality within 60 days of admission between adults with and without clinical seizure presentation in the open Jeddah cohort. We chose standardized absolute risks to make the comparison interpretable within the observed adult case mix. Prespecified analyses assessed age functional form, inclusion of the source's younger records, limited and full recorded-diagnosis adjustment, and propensity overlap weighting. A parallel source audit examined whether the public data reproduced the publication's marginal totals. This combination allows the statistical estimate and the underlying data uncertainty to be assessed together, without claiming a newly recruited Saudi cohort or a validated clinical prediction tool.

2. Materials and Methods

2.1. Design, setting, and data provenance

This was a retrospective secondary analysis of the public dataset associated with Tayeb's report of CEEG at King Abdulaziz University Hospital, Jeddah [1]. The source covers September 2016 through August 2017 in an adult ICU with medical and surgical divisions. It describes clinician-directed monitoring for patients with altered consciousness and related concerns. Records continuing over several days, or occurring within 60 days of one another, were considered a single recording for the original study. The public file does not contain a patient linkage identifier, so the original deduplication procedure cannot be independently reconstructed.

The original SPSS file was retrieved from the Open Science Framework project linked by the source publication. Its public license is CC0; the article is licensed under Creative Commons Attribution. We retained the unmodified SPSS file, project and license metadata, numeric CSV transcription, value labels, source article text, and cryptographic fingerprints. No simulated observations or research-engine fixture records were used. A repeated SPSS-to-CSV conversion with value formatting disabled reproduced the numeric CSV byte for byte. This check establishes transcription consistency, not the correctness of the clinical labels or the original extraction from medical records.

The source reports institutional review board approval and waiver of additional consent for its retrospective anonymized study [1]. We attribute those statements to that study. The present secondary analysis was exempt from ethical review because the data are publicly available and de-identified. Reporting follows STROBE cohort items and applicable RECORD principles concerning source definition, selection, coding, and transparency [15,16]. No linkage, re-identification, participant contact, or access to private hospital records was undertaken.

2.2. Plan freeze and eligibility

The statistical plan was timestamped and hashed before cross-tabulation or modeling of the presentation–mortality relationship. Feasibility work had already established the variable names, marginal counts, missingness, younger ages, and diagnosis-label problems. Published mortality and EEG totals were also known. The freeze therefore documents the order of analytic decisions, not prospective registration or complete blinding to the source. The primary model and all reported sensitivities were specified at that point, and no new subgroup or outcome was selected after model results became available.

The primary cohort included records with age at least 18 years and valid codes for sex, clinical seizure presentation, and mortality. Records with missing or invalid required values would have been excluded explicitly. The original report referred to an adult ICU but did not provide a clear age-based entry threshold; an all-source-age sensitivity was therefore specified to show the effect of our adult research definition. Diagnosis coding did not determine eligibility. Unlabeled diagnosis categories remained in the cohort, avoiding an additional selection rule based on incomplete metadata.

The public data represent the source's monitoring records rather than a separately verified population of unique first-ever hospital admissions. We use the terms record or admission episode when discussing analysis units. Their independence is inherited from the source's consolidation rule. Repeat admissions outside that rule cannot be assessed because the dataset lacks identifiers and dates. This limitation is

relevant to variance estimation and to the interpretation of the cohort size; numerical rows should not be equated automatically with independently verified individuals.

2.3. Exposure, outcome, and temporal interpretation

The exposure was the source variable SeizurePresentation, coded one for yes and two for no. The article describes clinical seizures on presentation and before CEEG monitoring [1]. We used that recorded binary classification without inferring seizure type, duration, status epilepticus, treatment response, or underlying cause. The file does not contain the clinical description needed to validate individual events or distinguish all epileptic events from other paroxysmal phenomena. It also does not specify exactly how much time elapsed between presentation, ICU admission, the monitoring request, and CEEG initiation.

The endpoint was the source mortality code indicating death within 60 days after admission, compared with the code indicating alive. We did not create event dates from the hospital-stay categories. The file contains neither individual time to death nor a detailed follow-up or censoring history. Accordingly, the analysis concerns recorded fixed-horizon mortality status. It does not estimate a death hazard, a Kaplan–Meier survival curve, or a competing-risk cumulative incidence function. The completeness and method of postdischarge vital-status ascertainment cannot be independently established from the public file.

Admission is the source's reference point for the mortality horizon, whereas selection into this dataset depends on later CEEG. An episode must therefore reach the monitoring process to be represented. This may select patients according to survival, severity, clinical suspicion, and service availability. Although the recorded presentation precedes EEG ascertainment, that sequence alone does not remove selection bias. We interpret the result conditionally among records selected for CEEG and do not describe it as the effect of having a seizure on mortality among all ICU admissions.

2.4. Primary model and absolute-risk estimand

The primary logistic regression included an intercept, clinical seizure presentation, age per decade centered at 60 years, and recorded male sex. The outcome was recorded death within the source horizon. These variables were complete and clearly coded, but they form a limited adjustment set. In particular, the model does not claim to control illness severity or all diagnostic differences. The primary result was selected as a transparent descriptive adjusted association rather than a sufficient causal model.

We generated two model predictions for each adult record, retaining age and sex and assigning each presentation label in turn. Averaging each set over the adult cohort gave standardized mortality probabilities, and their difference compared clinical presentation with no presentation. Marginal standardization targets the empirical covariate distribution and differs from evaluating a single patient with mean or modal characteristics [17]. Setting a label in this calculation is a statistical device for describing an adjusted association; it is not a feasible intervention or a claim that all other aspects of a patient's course could remain unchanged if the presenting event differed.

The primary uncertainty interval used the delta method and an individual-score sandwich covariance from the logistic fit, conditional on the observed standardization covariates. A two-sided Wald p value was retained, but interpretation emphasizes effect size and precision. Crude mortality proportions were accompanied by Wilson intervals. We did not test baseline differences to decide which covariates to include. The exploratory nature of the secondary analysis was retained regardless of whether any p value crossed a conventional threshold.

2.5. Diagnosis audit and sensitivity models

The SPSS labels named epilepsy, metabolic or infectious disease, cerebrovascular disease, central nervous system infection, brain tumor, traumatic brain injury, and post-arrest status. Two observed codes had no value labels. Moreover, the labeled category counts did not fully reproduce the diagnostic totals reported in the article. We retained the numeric codes and explicitly reported their unknown meanings. No code was reassigned from its frequency, its relation to mortality, or an attempt to make published totals match. The source-audit table distinguishes verified agreement from unresolved discrepancies.

The plan specified an age-quadratic model, the all-source-age model, and addition of an indicator for the code labeled epilepsy. The latter is deliberately coarse and does not imply that all remaining diagnoses are clinically similar. We also fitted a model with all observed diagnosis codes as nominal categories. Because some categories were small, this sensitivity used Firth penalized logistic regression, which can reduce small-sample coefficient bias and address separation [18]. The penalty does not resolve label uncertainty, improve the original clinical diagnosis, or create within-category exposure overlap where it is absent.

For the full diagnosis model, we maximized the binomial log likelihood plus one-half of the log determinant of its information matrix. Standardized risks used the adult cohort's observed diagnosis distribution. Percentile intervals came from 500 successful nonparametric bootstrap samples, with the model re-estimated in each sample. Empty diagnosis dummy columns were omitted within a resample; no new clinical grouping was introduced. This sensitivity uses a different estimator and uncertainty calculation from the primary model and is interpreted accordingly, rather than treated as a mechanically superior version of the same analysis.

2.6. Propensity overlap and variables deliberately excluded

A propensity model estimated the probability of clinical seizure presentation from age and sex. Overlap weights were one minus that probability for records with presentation and the probability for records without it [21]. We estimated weighted mortality probabilities, their difference, effective sample sizes, and standardized differences in the modeled covariates. A separate bootstrap used 500 successful resamples and a fixed recorded random seed. The overlap target emphasizes records with comparable age and sex; it does not standardize over exactly the same population as the primary analysis.

The source's NCS, rhythmic or periodic patterns, monitoring duration, and hospital-stay categories were not primary adjustment variables. They occur or are ascertained after the presenting clinical state and may reflect both disease evolution and clinical decisions. Adjusting for them without a clear causal and temporal model could change the estimand or introduce bias. Conditioning on consequences of exposure and prognosis can itself induce an association [19]. The present data do not supply enough timing or treatment detail to distinguish mediators, confounders, and selection variables for those later measurements.

We did not impute any diagnosis label or clinical event. The missing hospital-stay values were not needed for the mortality endpoint. Complete coding of primary fields does not mean absence of measurement error, and missingness in a field unused by the model should not be confused with complete clinical information. Guidance on incomplete observational data emphasizes that handling missing values depends on the target and assumed mechanism [20]. Here, inventing labels or follow-up times would create information rather than recover it; the appropriate response was to preserve the uncertainty.

2.7. Source reconciliation and computational verification

The source audit compared the public record count, sex, presentation, NCS, rhythmic or periodic pattern, mortality, and diagnosis totals with the publication. Agreement of major totals was treated as evidence of correct file identification, not proof that every row was clinically accurate. Disagreement was retained in the results and release record. Published outcomes related to NCS were inspected as a source-reconciliation check, not promoted as another independent study or an additional primary discovery from the same dataset.

Numerical analysis used explicit Python likelihood and standardization code with NumPy, pandas, and SciPy. A separate implementation parsed the numeric CSV with the standard-library CSV reader, rebuilt the adult design matrix, and fitted ordinary logistic regression with Newton updates. It recomputed the primary standardized risks, risk difference, and sandwich uncertainty without importing the analyst's model functions. A numerical-gradient optimization of the penalized likelihood independently checked the Firth coefficients. Raw files, source metadata, model code, verification scripts, bootstrap records, and a machine-readable results lock accompany the manuscript.

The verification tolerance was exact agreement for counts, a small specified tolerance for coefficients, and a tighter tolerance for standardized risks and their difference. Model rank and final score magnitude were inspected. The bootstrap receipt records unsuccessful fits if any occurred; no resamples were omitted because their estimates were inconvenient. These checks address computational implementation. They do not certify source clinical coding, patient independence, institutional permission, causal identification, or the clinical usefulness of the resulting contrast.

3. Results

3.1. Source audit and cohort construction

The public file contained 202 records. The main published totals for clinical seizure presentation, NCS, rhythmic or periodic patterns, mortality, and female sex were reproduced. Diagnosis reconciliation was incomplete: the code labeled epilepsy contained 50 records, and the cerebrovascular code contained 53 records, while the article reported different corresponding totals. Codes with no source labels accounted for 11 records (Table 3). These observations were preserved without relabeling or excluding the affected records.

Applying the prespecified age threshold excluded 10 records and yielded 192 adults (Figure 1). Required presentation, sex, age, and mortality fields were complete under their valid-code definitions. Of the eligible adults, 114 had clinical seizure presentation and 78 did not. A total of 62 deaths was recorded. The younger source records were included only in the planned all-age sensitivity; their presence does not imply that the original ICU's admission policy was incorrectly implemented, because its research age threshold was not explicitly supplied.

3.2. Case mix and crude mortality

The mean age was 52.2 years among records with clinical seizure presentation and 58.7 years among those without it. Recorded sex and diagnosis distributions also differed (Table 1; Figure 3). One small unlabeled diagnosis category contained only records with presentation, leaving no observed comparison within that category. The full diagnosis sensitivity therefore necessarily relies on model structure for part of

its standardized contrast. Retaining that category is more transparent than inventing its clinical meaning, but it does not solve the limited support problem.

There were 32 recorded deaths among the 114 presentation records and 30 among the 78 records without presentation. Crude mortality proportions were 28.1% and 38.5%, respectively; the crude difference was -10.39 percentage points. These proportions describe the selected monitoring cohort. They do not measure seizure-attributable mortality or the probability of death among all patients admitted to the hospital with a clinical seizure.

3.3. Standardized primary result

The age- and sex-adjusted model converged with a full-rank design matrix. Standardized mortality was 29.7% under the clinical-presentation label and 35.9% under the no-presentation label. The adjusted risk difference was -6.23 percentage points (95% confidence interval, -19.75 to 7.30; p=0.367). The central estimate was lower in the presentation group, but the interval included substantial reductions and increases in mortality. It does not demonstrate that clinical seizures are protective, that presentation carries no prognostic information, or that the two groups have equivalent outcomes.

Adjustment reduced the magnitude of the crude difference, consistent with measured age and sex differences contributing to the unadjusted comparison. Because both standardized predictions were averaged over the same adult covariate distribution, the result is not a comparison of two different average patients. However, the standardization does not address unmeasured severity, medication exposure, functional status, or detailed etiology. Those omissions remain material even when the model converges and numerical verification succeeds.

3.4. Prespecified sensitivity analyses

Adding age curvature, including all source ages, or adding the recorded epilepsy indicator yielded similarly imprecise contrasts (Table 2). The full diagnosis Firth model estimated a standardized difference of -4.11 percentage points, with a bootstrap interval from -19.42 to 10.19. All 500 planned bootstrap estimates were successfully obtained. This estimate was closer to the null than the primary estimate, but its interval remained broad. The sensitivity supports uncertainty about the association; it does not provide evidence that the coding discrepancy is harmless.

The overlap-weighted difference was -6.38 percentage points (bootstrap interval, -19.87 to 9.22), based on 500 successful resamples. Estimated presentation propensities ranged from 0.399 to 0.765. Effective sample sizes were 106.8 for the presentation group and 75.7 for the no-presentation group. Absolute standardized differences for age and male sex decreased from 0.332 and 0.322 to numerical near-zero after weighting. The expected balance of modeled covariates does not imply balance of diagnoses or illness severity, which were not in this propensity model.

Figure 2 displays the complete set of planned absolute contrasts. None supplies a precise separation of mortality between the recorded presentation groups. The results were not used to create post hoc age thresholds, diagnostic subgroups, or alternative primary endpoints. The computational verification reproduced primary counts and standardized estimates and agreed with the independently optimized Firth coefficients within tolerance. Numerical reproducibility does not resolve the source-label and clinical-interpretation limitations.

4. Discussion

4.1. Meaning of the main estimate

In this Saudi cohort selected for CEEG, adults with recorded clinical seizure presentation had a lower central estimate of mortality than those without it after age and sex adjustment. The confidence interval included effects in both directions; the estimate therefore does not establish a protective or harmful association. Full recorded-diagnosis adjustment using a penalized model did not change that overall conclusion. The proper interpretation is therefore an uncertain adjusted comparison within a selected historical cohort, not a reassurance that clinical seizure presentation is benign and not an argument that visible seizures improve survival.

This distinction matters clinically because an overt seizure can prompt rapid recognition, consultation, and monitoring, while a patient without an overt event may be referred because of severe or persistent unexplained unconsciousness. These are plausible pathways through which referral could shape the observed comparison; they were not measured sufficiently to test in this dataset. The negative central risk difference could reflect such selection, the distribution of underlying diseases, treatment decisions, measurement error, or chance. The data cannot apportion the observed difference among those explanations.

The interval is also important in its own right. Failure to obtain a narrow interval excluding no difference does not establish equivalence or absence of prognostic information. The sample supports a range of associations that would have different clinical implications. Reporting only the p value would hide that uncertainty, while emphasizing the negative point estimate without its interval would invite an unjustified protective interpretation. The absolute-risk presentation is intended to make both the scale and uncertainty visible.

4.2. Clinical presentation is distinct from electrographic exposure

The original source examined NCS and rhythmic or periodic patterns observed during monitoring [1]. Our exposure was the recorded clinical event before monitoring. A patient may have a witnessed clinical seizure and no subsequently detected NCS, or have NCS despite no documented clinical seizure at presentation. Detection also depends on the interval sampled and the course of treatment. Evidence that prior clinical events are associated with electrographic seizure detection [5] does not imply that either exposure has the same relationship with mortality.

The cohort of patients initially presenting with seizures and admitted directly to ICU described by Tobochnik and colleagues addresses another selected pathway [6]. Such a study can clarify outcomes within seizure-presenting patients but does not provide a no-presentation comparator identical to ours. Similarly, the earlier work on NCS duration and diagnostic delay evaluated features within a monitored seizure population [7]. These differences illustrate why clinical interpretation should follow the exposure definition, entry criteria, and time origin rather than the common word “seizure” in each study title.

Our dataset cannot identify whether a presenting event was isolated, prolonged, recurrent, or accompanied by a specific acute brain lesion. It also lacks consistent measures of consciousness, antiseizure medication timing, sedative dosing, and treatment limitation. Each of these may help distinguish seizure severity from the severity of the illness producing it. A binary source field is useful for an auditable descriptive

comparison, but it is an incomplete representation of the clinical problem and should not be elevated into a standalone prognostic instrument.

4.3. Monitoring strategy and outcome prediction are separate questions

Consensus statements support defined roles for EEG assessment and describe the requirements for reliable monitoring [2–4]. They do not imply that every mortality association in a monitored cohort estimates the benefit of that monitoring. The CERTA trial explicitly assigned monitoring strategies in a population without recent seizures [8]. Its analyses of EEG features and post-arrest patients address further, distinct questions [9,10]. None of these designs is recreated by comparing the two presentation codes in the present file.

The same distinction applies to studies of EEG in unresponsive ventilated patients, sepsis, and nonhypoxic encephalopathy [11–13]. Background reactivity, level of consciousness, underlying brain injury, and timing relative to sedation can be central to their interpretation. The public Jeddah dataset has limited information on those features. We cannot use the present result to infer that a patient without a clinical seizure needs less monitoring, or that one with a seizure needs a particular monitoring duration. Such decisions require information and evidence beyond this analysis.

Nor does the analysis compare antiseizure treatment strategies. Treatment is not assigned or characterized in sufficient detail, and the presenting event may influence subsequent therapy. A putative treatment effect cannot be recovered simply by adding post-presentation EEG findings to a mortality regression. Without an adequately specified temporal model, that adjustment could obscure the pathway of interest or condition on a common consequence of disease and management [19]. We therefore kept those later variables out of the primary model rather than treating every available field as a confounder.

4.4. What the sensitivity analyses add

The age-quadratic and all-age analyses show that the central result was not substantially altered by those particular specification choices. The epilepsy-indicator analysis introduces one clearly labeled but coarse diagnostic distinction. It cannot separate all etiologies or establish that the epilepsy code itself is fully accurate. These analyses are useful checks of measured features, but their agreement does not provide independent replication because they share the same source records, measurement procedures, and unobserved clinical information.

The full nominal diagnosis model retains more of the recorded case mix at the cost of additional parameters and dependence on uncertain coding. Firth estimation was selected to address small-cell estimation behavior [18]; it does not repair misclassification or remove unmeasured confounding. Because one small code had only presentation records, some standardized predictions extrapolate across a cell without observed within-code comparison. Penalization yields finite estimates in such settings, but finite coefficients should not be mistaken for strong empirical support for every counterfactual label assignment.

Overlap weighting targets a related age- and sex-comparable population [21]. Its near-exact mean balance is a mathematical property of the included covariates and fitted weights, not a comprehensive certificate of comparability. Diagnosis and severity may remain imbalanced. The agreement between overlap and regression estimates therefore suggests stability across two approaches to those limited measured covariates, while leaving the substantive confounding problem largely intact. This is why the overlap analysis remains a sensitivity rather than a replacement for the stated primary result.

4.5. Source quality, observation time, and selection

The data audit identified both reassuring agreement and important unresolved uncertainty. Major source totals reproduced, and the CSV was independently regenerated from the raw SPSS file. These checks support the identity and faithful transcription of the public dataset. However, diagnosis totals and labels did not fully reconcile with the publication. It would have been possible to merge or relabel codes to produce a more familiar clinical table, but that would add assumptions unsupported by the source. We instead retained the unknown categories and avoided disease-specific claims.

The age restriction also changes the target population. The original setting is described as an adult ICU, while the public file contains younger ages. This may reflect the institution's admission convention rather than an extraction error. Our primary research definition explicitly uses age at least 18, and the source-age sensitivity retains the original records. Distinguishing a unit's name from a stated research eligibility rule prevents an unnecessary assertion that the source is wrong solely because its youngest patient is below an externally selected age threshold.

The temporal limitations are more fundamental. Mortality is indexed to admission, but monitoring takes place later and may be delayed by clinical or practical factors described in the source [1]. No exact monitoring or event times permit delayed-entry analysis, landmark construction, or assessment of individual risk time. Hospital-stay categories cannot supply those missing dates. A survival analysis with artificial times would create unjustified precision. The fixed-endpoint analysis is the most direct use of the recorded outcome, but its validity still depends on adequate ascertainment of vital status and on recognizing selection into the monitored cohort.

The data also lack identifiers needed to verify that source-defined recording episodes correspond to independent patients. The source consolidated closely spaced recordings, but episodes farther apart might not be distinguishable in the public file. Our sandwich and bootstrap calculations treat rows as independent units. If repeated individuals remain, their dependence could affect uncertainty. No cluster correction can be justified without a reliable cluster key. This is a limitation to be documented, not a reason to fabricate identifiers or assume a deduplication procedure beyond the source description.

4.6. Implications for Saudi research and service evaluation

The setting makes the study geographically relevant to Saudi neurocritical-care research, but it does not provide a national estimate. It reflects one historical monitoring program and only records selected for that program. The separate Saudi report on clinical EEG correlates demonstrates that local observational data exist in another setting [14]; its different design and population should not be treated as direct validation of the present mortality contrast. A credible multicenter extension would require harmonized entry criteria, measurements, and follow-up rather than simply pooling available EEG reports.

For a prospective Saudi study, the first design decision should be the target population: all ICU admissions, all patients with suspected seizures, or all patients receiving CEEG. Those choices answer different questions and produce different selection mechanisms. A presentation-based prognosis study should define an index time before the outcome risk period and record the clinical event in enough detail to distinguish seizure types and duration. A service-effectiveness study would additionally need a defensible comparator monitoring strategy and careful documentation of how allocation occurs. These objectives should not be combined under a vague claim of EEG utility.

Useful data elements would include admission, event, request, and monitoring timestamps; standardized mental-status and severity assessments; a validated diagnosis dictionary; prior epilepsy and acute brain injury information; and the timing of sedatives, antiseizure medication, and treatment limitations. Monitoring duration should be recorded as an observed process rather than treated automatically as a baseline exposure. Patient identifiers can remain protected while enabling an authorized data custodian to construct a reliable episode and repeat-admission key. Clear data governance would allow these improvements without exposing identifiable records in the public analysis.

Outcome ascertainment should distinguish ICU death, hospital death, and mortality after discharge, with a documented method of follow-up to the chosen horizon. Functional and patient-centered outcomes may add information that survival alone cannot capture. Finally, source validation should be completed before fitting large models: label dictionaries should reconcile with category totals, a sample of records should be checked against the originating clinical source where authorized, and disagreements should be resolved by accountable investigators. These are proposed improvements for future work, not procedures completed in this public-data reanalysis.

5. Conclusions

In the historical Saudi ICU monitoring records, the adjusted mortality point estimate was lower with recorded clinical seizure presentation, but the confidence interval was compatible with differences in either direction, and selection into monitoring limits generalizability. The analysis does not establish protection, harm, or equivalence, and it does not estimate the effectiveness of monitoring or treatment. Its reproducible computations are accompanied by unresolved diagnosis coding, selection, timing, and clinical-information limitations. Prospective cohorts with validated diagnosis codes, distinct-person linkage, severity and event timing are required before this observational contrast can inform prognostic or service decisions.

Tables

Table 1. Recorded characteristics and mortality by clinical seizure presentation among eligible adults

Characteristic	No clinical seizure presentation	Clinical seizure presentation
Records	78	114
Age, years, mean (SD)	58.7 (18.9)	52.2 (19.9)
Age, years, median [IQR]	62 [44, 73]	56 [32, 67]
Male	39 (50.0%)	39 (34.2%)
Epilepsy (source label)	30 (38.5%)	16 (14.0%)
Code 4: unlabeled	0 (0.0%)	5 (4.4%)
Metabolic/infectious (source label)	21 (26.9%)	19 (16.7%)
Cerebrovascular disease (source label)	12 (15.4%)	36 (31.6%)
Code 9: unlabeled	1 (1.3%)	5 (4.4%)
CNS infection (source label)	4 (5.1%)	19 (16.7%)

Characteristic	No clinical seizure presentation	Clinical seizure presentation
Brain tumor (source label)	2 (2.6%)	6 (5.3%)
Traumatic brain injury (source label)	2 (2.6%)	4 (3.5%)
Post-arrest (source label)	6 (7.7%)	4 (3.5%)
Deaths within 60 days	30 (38.5%)	32 (28.1%)
Mortality, % (Wilson 95% CI)	38.5 (28.4–49.6)	28.1 (20.6–36.9)

Values are counts (percentages), mean (standard deviation), median [interquartile range], or mortality proportion with Wilson 95% confidence interval. Diagnosis labels are copied from the public SPSS metadata and are not independently clinically validated; codes 4 and 9 have no labels. CNS, central nervous system. Interpretation: presentation groups differ in age, sex, and recorded diagnosis mix; diagnosis uncertainties limit clinical interpretation of the category differences. Source: RESULTS_LOCK.json, group_summary and reporting.

Table 2. Standardized mortality contrasts under the primary and prespecified sensitivity analyses

Analysis	N	Risk without presentation, %	Risk with presentation, %	Difference, percentage points (95% interval)
Primary age/sex	192	35.9	29.7	-6.23 (-19.75, 7.30)
Age quadratic	192	35.9	29.7	-6.22 (-19.70, 7.25)
All source ages	202	34.0	28.3	-5.70 (-18.59, 7.20)
Recorded epilepsy indicator	192	35.7	29.8	-5.98 (-20.17, 8.22)
All diagnosis codes, Firth	192	35.4	31.3	-4.11 (-19.42, 10.19)
Age/sex overlap weighting	192	36.8	30.4	-6.38 (-19.87, 9.22)

Differences compare clinical seizure presentation minus no presentation. Ordinary logistic-model intervals use the sandwich/delta method conditional on standardization covariates; Firth and overlap intervals are nonparametric bootstrap percentile intervals. The all-age model changes the population, and overlap weighting changes the target distribution. Full diagnosis adjustment retains nominal source codes without assigning meanings to unlabeled categories. Interpretation: all contrasts are imprecise; a negative central estimate is not evidence that clinical seizures are protective. Source: RESULTS_LOCK.json, primary, sensitivity, overlap.

Table 3. Reconciliation of major published source totals with the public dataset

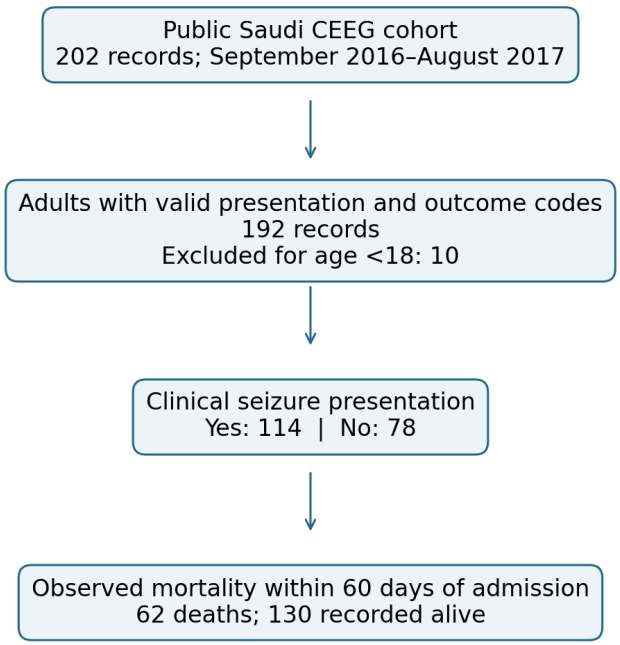
Source item	Published value	Public data value	Reconciliation
Source records	202	202	Matches
Clinical seizure presentation	120	120	Matches
Nonconvulsive seizures	52	52	Matches
Rhythmic/periodic patterns	138	138	Matches
Deaths within 60 days	62	62	Matches
Female patients	116	116	Matches
Epilepsy	53	50	Does not match source-labeled code 1

Source item	Published value	Public data value	Reconciliation
Cerebrovascular diagnoses	61	53	Published ischemic + hemorrhagic categories; public label does not distinguish them
Unlabeled diagnosis codes	Not separately described	11	Codes 4 and 9 retained as nominal categories

Published counts refer to the full source cohort, before the current adult restriction. The source separates ischemic stroke and intracranial hemorrhage, whereas the public metadata labels one code as cerebrovascular disease and leaves two other codes unlabeled. Their meanings were not inferred from counts. Interpretation: matching major totals supports file identity, but unresolved diagnostic discrepancies remain a scientific-release issue. Source: Tayeb, version 3 [1]; public SPSS value labels and RESULTS_LOCK.json source audit.

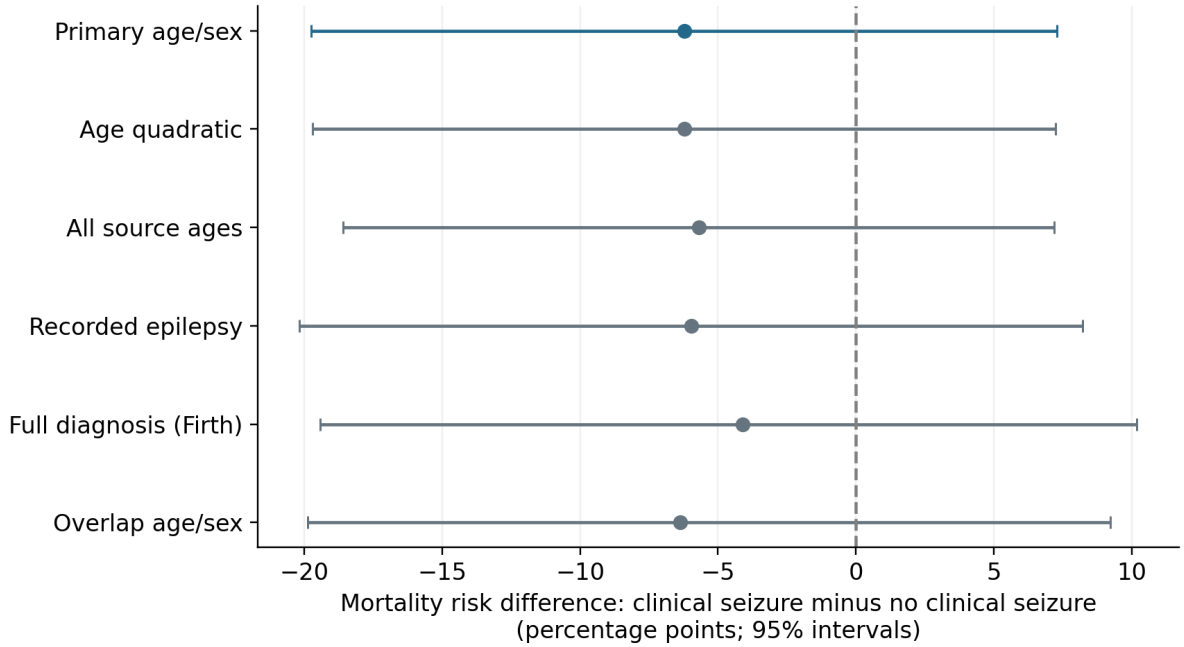
Figures and legends

Figure 1. Construction of the adult CEEG analysis cohort



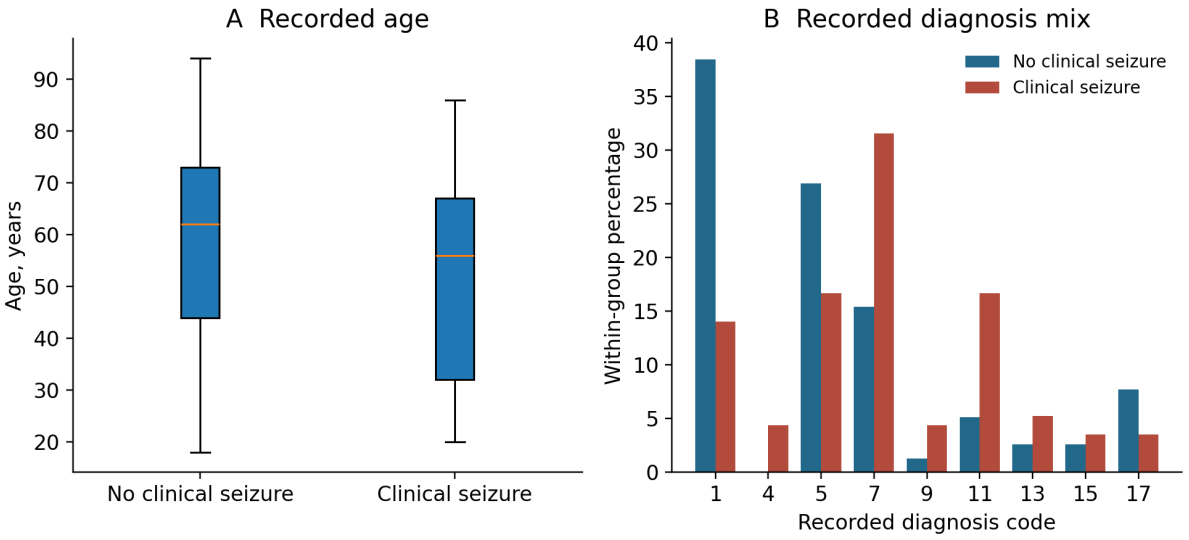
The source's recording-episode definition was inherited; unique patient identities cannot be independently verified. The current analysis applies an explicit adult age threshold and valid-code requirements. Interpretation: the target is adults already selected for CEEG, not every ICU admission. Source: RESULTS_LOCK.json, source_audit and primary.

Figure 2. Mortality probability differences across the primary and sensitivity analyses



Points are clinical-presentation-minus-no-presentation differences; bars are 95% intervals. Firth and overlap estimates use bootstrap intervals, while ordinary logistic estimates use sandwich/delta intervals. The vertical line denotes no difference. Interpretation: the estimates remain compatible with materially different directions of association. Source: RESULTS_LOCK.json, primary, sensitivity and overlap.

Figure 3. Age and recorded diagnosis distributions by presentation group



Panel A displays recorded age distributions. Panel B gives percentages within each presentation group for the exact source diagnosis codes. Codes 1, 5, 7, 11, 13, 15, and 17 are labeled epilepsy, metabolic/infectious, cerebrovascular, CNS infection, brain tumor, traumatic brain injury, and post-arrest, respectively; codes 4 and 9 are unlabeled. Interpretation: measurable case-mix differences and limited within-code overlap caution against causal interpretation. Source: locked cohort and source value-label dictionary.

Declarations

Ethics and consent: The source author reported King Abdulaziz University Hospital institutional review board approval and a waiver of consent for the original retrospective anonymised study [1]. The present secondary analysis was exempt from ethical review because the data are publicly available (CC0) and de-identified.

Funding: This research received no external funding.

Competing interests: The authors declare no competing interests.

Data and code availability: Source data: Open Science Framework, https://doi.org/10.17605/OSF.IO/Q56J3, associated with https://doi.org/10.12688/f1000research.19237.3. The frozen plan, scripts, bootstrap estimates, results and derived tables and figures are available from the corresponding author on request. Registration: the analysis plan was frozen after feasibility inspection and before presentation–mortality modelling; it was not prospectively registered.

Use of artificial intelligence: Artificial intelligence was not used in the preparation of this manuscript.

References

1. Haythum O Tayeb. The yield of continuous EEG monitoring in the intensive care unit at a tertiary care hospital in Saudi Arabia: A retrospective study. F1000Research. 2019;8:663. doi:10.12688/f1000research.19237.3. PMID:31737255. https://doi.org/10.12688/f1000research.19237.3

2. Susan T Herman, Nicholas S Abend, Thomas P Bleck, et al. Consensus statement on continuous EEG in critically ill adults and children, part I: indications. Journal of clinical neurophysiology : official publication of the American Electroencephalographic Society. 2015 Apr;32(2):87-95. doi:10.1097/WNP.0000000000000166. PMID:25626778. https://doi.org/10.1097/WNP.0000000000000166

3. Susan T Herman, Nicholas S Abend, Thomas P Bleck, et al. Consensus statement on continuous EEG in critically ill adults and children, part II: personnel, technical specifications, and clinical practice. Journal of clinical neurophysiology : official publication of the American Electroencephalographic Society. 2015 Apr;32(2):96-108. doi:10.1097/WNP.0000000000000165. PMID:25626777. https://doi.org/10.1097/WNP.0000000000000165

4. Jan Claassen, Fabio S Taccone, Peter Horn, et al. Recommendations on the use of EEG monitoring in critically ill patients: consensus statement from the neurointensive care section of the ESICM. Intensive care medicine. 2013 Aug;39(8):1337-51. doi:10.1007/s00134-013-2938-4. PMID:23653183. https://doi.org/10.1007/s00134-013-2938-4

5. J Claassen, S A Mayer, R G Kowalski, et al. Detection of electrographic seizures with continuous EEG monitoring in critically ill patients. Neurology. 2004 May 25;62(10):1743-8. doi:10.1212/01.wnl.0000125184.88797.62. PMID:15159471. https://doi.org/10.1212/01.wnl.0000125184.88797.62

6. Steven Tobochnik, Camilo Gutierrez, Mercedes P Jacobson. Characteristics and acute outcomes of ICU patients with initial presentation of seizure. Seizure. 2015 Mar;26:94-7. doi:10.1016/j.seizure.2015.01.016. PMID:25799909. https://doi.org/10.1016/j.seizure.2015.01.016

7. G B Young, K G Jordan, G S Doig. An assessment of nonconvulsive seizures in the intensive care unit using continuous EEG monitoring: an investigation of variables associated with mortality. Neurology. 1996 Jul;47(1):83-9. doi:10.1212/wnl.47.1.83. PMID:8710130. https://doi.org/10.1212/wnl.47.1.83

8. Andrea O Rossetti, Kaspar Schindler, Raoul Sutter, et al. Continuous vs Routine Electroencephalogram in Critically Ill Adults With Altered Consciousness and No Recent Seizure: A Multicenter Randomized Clinical Trial. JAMA neurology. 2020 Oct 01;77(10):1225-1232. doi:10.1001/jamaneurol.2020.2264. PMID:32716479. https://doi.org/10.1001/jamaneurol.2020.2264

9. Isabelle Beuchat, Andrea O Rossetti, Jan Novy, et al. Continuous Versus Routine Standardized Electroencephalogram for Outcome Prediction in Critically Ill Adults: Analysis From a Randomized Trial. Critical care medicine. 2022 Feb 01;50(2):329-334. doi:10.1097/CCM.0000000000005311. PMID:34582427. https://doi.org/10.1097/CCM.0000000000005311

10. Valentina Urbano, Vincent Alvarez, Kaspar Schindler, et al. Continuous versus routine EEG in patients after cardiac arrest: Analysis of a randomized controlled trial (CERTA). Resuscitation. 2022 Jul;176:68-73. doi:10.1016/j.resuscitation.2022.05.017. PMID:35654226. https://doi.org/10.1016/j.resuscitation.2022.05.017

11. Camille Legouy, Laura Girard-Stein, Ruben Wanono, et al. Association of Standard Electroencephalography Findings With Mortality and Command Following in Mechanically Ventilated Patients Remaining Unresponsive After Sedation Interruption. Critical care medicine. 2021 Apr 01;49(4):e423-e432. doi:10.1097/CCM.0000000000004874. PMID:33591021. https://doi.org/10.1097/CCM.0000000000004874

12. Eric Azabou, Eric Magalhaes, Antoine Braconnier, et al. Early Standard Electroencephalogram Abnormalities Predict Mortality in Septic Intensive Care Unit Patients. PloS one. 2015;10(10):e0139969. doi:10.1371/journal.pone.0139969. PMID:26447697. https://doi.org/10.1371/journal.pone.0139969

13. Raoul Sutter, Peter W Kaplan. Clinical, Electroencephalographic, and Neuroradiological Outcome Predictors in Acute Nonhypoxic Encephalopathy: A Nine-Year Cohort Study. Clinical EEG and neuroscience. 2016 Jan;47(1):61-8. doi:10.1177/1550059415579768. PMID:25828484. https://doi.org/10.1177/1550059415579768

14. Saima Nazish, Azra Zafar, Erum Shariff, et al. Clinical Correlates of Electroencephalographic Patterns in Critically Ill Patients. Clinical EEG and neuroscience. 2021 Jul;52(4):287-295. doi:10.1177/1550059420966844. PMID:33104405. https://doi.org/10.1177/1550059420966844

15. Erik von Elm, Douglas G Altman, Matthias Egger, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS medicine. 2007 Oct 16;4(10):e296. doi:10.1371/journal.pmed.0040296. PMID:17941714. https://doi.org/10.1371/journal.pmed.0040296

16. Eric I Benchimol, Liam Smeeth, Astrid Guttmann, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. PLoS medicine. 2015 Oct;12(10):e1001885. doi:10.1371/journal.pmed.1001885. PMID:26440803. https://doi.org/10.1371/journal.pmed.1001885

17. Clemma J Muller, Richard F MacLehose. Estimating predicted probabilities from logistic regression: different methods correspond to different target populations. International journal of epidemiology. 2014 Jun;43(3):962-70. doi:10.1093/ije/dyu029. PMID:24603316. https://doi.org/10.1093/ije/dyu029

18. Georg Heinze, Michael Schemper. A solution to the problem of separation in logistic regression. Statistics in medicine. 2002 Aug 30;21(16):2409-19. doi:10.1002/sim.1047. PMID:12210625. https://doi.org/10.1002/sim.1047

19. Stephen R Cole, Robert W Platt, Enrique F Schisterman, et al. Illustrating bias due to conditioning on a collider. International journal of epidemiology. 2010 Apr;39(2):417-20. doi:10.1093/ije/dyp334. PMID:19926667. https://doi.org/10.1093/ije/dyp334

20. Rachael A Hughes, Jon Heron, Jonathan A C Sterne, et al. Accounting for missing data in statistical analyses: multiple imputation is not always the answer. International journal of epidemiology. 2019 Aug 01;48(4):1294-1304. doi:10.1093/ije/dyz032. PMID:30879056. https://doi.org/10.1093/ije/dyz032

21. Li, F.; Morgan, K.L.; Zaslavsky, A.M. Balancing Covariates via Propensity Score Weighting. Journal of the American Statistical Association 2018, 113, 390–400. https://doi.org/10.1080/01621459.2016.1260466

Multidisciplinary Healthcare Journal · Vol. 1, No. 3 (2025) · CC BY 4.0 · https://creativecommons.org/licenses/by/4.0/