eISSN (Online): 2598-0580

Bioscientia Medicina: Journal of Biomedicine & Translational Research

Journal Homepage: www.bioscmed.com

Optimal Timing of Tracheostomy in Critically Ill Adults Requiring Prolonged Mechanical Ventilation: An Updated Meta-Analysis of Randomized Controlled Trials

Fauzul Nurul Azmi[1*], Novita Anggraeni[2], Johannas[2], Wan Novriza Wijaya[2]

[1]Specialized Residency Training Program, Anesthesiology and Intensive Therapy Study Program, Faculty of Medicine Universitas Riau, Pekanbaru, Indonesia

[2]Department of Anesthesiology and Intensive Therapy, Arifin Achmad Regional General Hospital, Pekanbaru, Indonesia

ARTICLE INFO

Keywords:
Critical illness
Mechanical ventilation
Meta-analysis
Timing
Tracheostomy
Ventilator-associated pneumonia

***Corresponding author:**
Fauzul Nurul Azmi

E-mail address:
dr.fauzul@gmail.com

The article has been reviewed and approved for publication in accordance with the policies of Bioscientia Medicina.

https://doi.org/10.37275/bsm.v10i8.1640

A B S T R A C T

Background: Tracheostomy is among the most frequently performed procedures in the intensive care unit (ICU), yet the optimal timing relative to the onset of invasive mechanical ventilation remains contested. This study aimed to provide an updated quantitative synthesis of randomized controlled trials (RCTs) comparing early versus late tracheostomy in critically ill adults, incorporating the two most recent landmark trials.

Methods: PubMed was systematically searched, supplemented by reference-list screening, for RCTs comparing early with late tracheostomy in mechanically ventilated adults. Study selection and data extraction were performed independently and in duplicate. Risk of bias was appraised with the Cochrane RoB 2 tool and the certainty of evidence with the GRADE framework. Dichotomous outcomes (all-cause mortality, ventilator-associated pneumonia [VAP]) were pooled as risk ratios (RR); continuous outcomes (duration of mechanical ventilation, ventilator-free days) as standardized mean differences (SMD, Hedges' g), using a DerSimonian–Laird random-effects model.

Results: Nine RCTs enrolling 2,500 critically ill adults were included. Early tracheostomy was not associated with reduced all-cause mortality (RR 0.88, 95% CI 0.70–1.09; p=0.24; I^2=58.9%; prediction interval 0.48–1.59; seven trials; moderate certainty). A non-significant trend towards fewer VAP episodes was observed (RR 0.67, 95% CI 0.42–1.05; p=0.08; I^2=71.0%; four trials; low certainty). Early tracheostomy showed non-significant tendencies towards a shorter duration of mechanical ventilation (SMD −1.38, 95% CI −3.44 to 0.68; I^2=97.2%) and more ventilator-free days (SMD 0.20, 95% CI −0.07 to 0.47). Leave-one-out and Hartung–Knapp–Sidik–Jonkman analyses confirmed the robustness of the neutral mortality finding.

Conclusion: In critically ill adults requiring prolonged mechanical ventilation, early tracheostomy did not significantly reduce mortality and conferred, at most, modest and uncertain benefits on VAP and ventilation-related resource use. Timing should remain an individualized clinical decision rather than a uniform protocol.

1. Introduction

Invasive mechanical ventilation is a cornerstone of contemporary critical care, but prolonged translaryngeal intubation carries a well-recognized burden of complications, including laryngotracheal injury, sinusitis, patient discomfort, heightened sedation requirements and an increased risk of ventilator-associated pneumonia (VAP). Tracheostomy was developed, in part, to mitigate these hazards by providing a more secure and comfortable airway, facilitating pulmonary toilet, reducing dead space and anatomical resistance, and enabling earlier liberation from sedation.[1] As a consequence, tracheostomy has become one of the most commonly performed surgical

procedures among ventilated patients in the ICU; approximately one in ten patients who require mechanical ventilation, and a far higher proportion of those who require it for more than one week, ultimately undergo tracheostomy.[1] The percutaneous dilatational technique has progressively supplanted the open surgical approach at the bedside owing to its speed, lower cost and comparable safety profile, and its widespread adoption has made the question of timing more pressing still, because the procedure can now be performed rapidly and early in the ICU course.[1]

Despite the ubiquity of the procedure, the question of when a tracheostomy should be performed has remained one of the most enduring controversies in intensive care medicine. Proponents of an early strategy argue that performing tracheostomy within the first few days of ventilation shortens the cumulative exposure to the endotracheal tube, thereby reducing airway trauma, sedation load and, potentially, the incidence of VAP, while accelerating weaning and ICU discharge. Advocates of a more conservative, late strategy counter that a substantial proportion of patients recover and are extubated before a tracheostomy would ever become necessary, so that an early policy inevitably exposes a number of patients—and a sizeable number, in practice—to an invasive procedure and its attendant bleeding, stomal infection and tracheal stenosis, from which they would not have benefited.

Over the past two decades, a series of RCTs has attempted to resolve this question, ranging from small single-center studies in selected populations to large pragmatic multicenter trials in general critical care. These trials have produced strikingly heterogeneous results: whereas some early single-center studies reported dramatic reductions in mortality and pneumonia with an early approach,[2] the larger and more rigorously conducted multicenter trials have generally failed to demonstrate a survival benefit.[3,4] This discordance has, in turn, generated a crowded landscape of systematic reviews and meta-analyses, several of which reached conflicting conclusions and at least one of which was retracted and subsequently republished with materially different, and attenuated, effect estimates.[5]

Two recent landmark trials have meaningfully changed the evidence base. A large randomized trial in patients with severe stroke compared an early tracheostomy strategy with a standard approach and assessed functional outcome at six months,[6] while a multicenter trial in patients with coronavirus disease 2019 (COVID-19) examined the timing of tracheotomy during the pandemic surge.[7] Both trials post-date the majority of earlier pooled analyses, and their inclusion is essential for any contemporary appraisal of the question. The continued emergence of new data, the persistent clinical equipoise and the methodological limitations of previous syntheses together justify an updated, RCT-restricted quantitative synthesis.

The novelty of this study lies in providing an updated, randomized-evidence-only meta-analysis of tracheostomy timing in critically ill adults that, for the first time, integrates the two most recent landmark trials—in severe stroke and in COVID-19—within a single coherent random-effects framework; that appraises dichotomous and continuous outcomes in parallel; that incorporates a pre-specified subgroup analysis according to the percutaneous dilatational technique; and that explicitly quantifies the residual uncertainty through a prediction interval, leave-one-out and alternative-variance sensitivity analyses, and a structured GRADE certainty assessment. The specific incremental contribution of this work, beyond the confirmation of earlier syntheses, is therefore the formal incorporation of the contemporary trials together with a transparent characterization of how much, and why, uncertainty persists. The aim of this study was to determine whether early tracheostomy, compared with late tracheostomy, reduces all-cause mortality and ventilator-associated pneumonia and shortens the duration of mechanical ventilation in critically ill adults requiring prolonged mechanical ventilation.

2. Methods

2.1. Protocol and reporting

This systematic review and meta-analysis was conducted and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement.[8] The review

question was framed using the Population–Intervention–Comparator–Outcome (PICO) structure: in critically ill adults requiring prolonged invasive mechanical ventilation (P), does early tracheostomy (I), compared with late tracheostomy or prolonged intubation (C), improve mortality, VAP and ventilation-related outcomes (O). Only RCTs were considered eligible. The eligibility criteria, outcomes and analysis plan were defined a priori and fixed before data extraction commenced.

2.2. Search strategy

The PubMed/MEDLINE database was systematically searched from inception. The search combined controlled vocabulary and free-text terms for the intervention, the population and the study design. The core Boolean string was:

("tracheostomy"[MeSH Terms] OR "tracheotomy"[MeSH Terms] OR tracheostomy[tiab] OR tracheotomy[tiab]) AND (early[tiab] AND (late[tiab] OR delayed[tiab] OR prolonged[tiab] OR timing[tiab])) AND ("critical illness"[MeSH Terms] OR "critically ill"[tiab] OR "intensive care"[tiab] OR "mechanical ventilation"[tiab] OR "respiration, artificial"[MeSH Terms]) AND ("randomized controlled trial"[Publication Type] OR randomized[tiab] OR randomised[tiab]).

To mitigate the limitation inherent in relying on a single bibliographic database, the reference lists of all retrieved trials and of every previous systematic review and meta-analysis on the topic were screened manually to identify additional eligible studies; this backward-citation approach captured the full set of major randomized trials in the field, and no eligible RCT identified in any prior synthesis was missing from our set. No language restriction was applied at the screening stage, although all finally included trials were reported in English. We acknowledge that a multi-database search would be preferable for a definitive synthesis, and this point is revisited in the limitations.

2.3. Eligibility criteria

Studies were eligible if they: (1) were RCTs; (2) enrolled critically ill adults (≥18 years) receiving, or anticipated to require, prolonged invasive mechanical ventilation; (3) compared a clearly defined early tracheostomy strategy with a late tracheostomy or prolonged intubation strategy; and (4) reported at least one outcome of interest. Studies were excluded if they were non-randomized or observational; were systematic reviews, meta-analyses, narrative reviews, editorials or conference abstracts; compared tracheostomy techniques or airway devices rather than timing; or were conducted exclusively in paediatric or neonatal populations. Conference abstracts were excluded because they typically lack sufficient methodological detail for risk-of-bias appraisal and for verifiable data extraction, and because their results are frequently revised before full publication.

2.4. Study selection and data extraction

Records identified by the search were screened by title and abstract, and potentially eligible reports were assessed in full against the eligibility criteria. Study selection and data extraction were performed independently and in duplicate by the reviewers, with disagreements resolved by discussion and consensus. For each included trial, the following data were extracted into a standardized spreadsheet: first author, year of publication, country, clinical setting and population, definitions of the early and late strategies, the technique of tracheostomy, the number of participants randomized to each arm, the proportion of each arm that actually received a tracheostomy, the precise definition and timepoint of each outcome, and the outcome data themselves. Dichotomous outcomes were recorded as the number of events and the number of participants in each arm; continuous outcomes were recorded as the mean, standard deviation and number of participants in each arm. Where a value was not reported in the primary publication, it was sought from the published data tables of high-quality prior systematic reviews, and the provenance of every such value was documented; where a value remained unavailable, the corresponding study did not contribute to that particular pooled estimate. All extracted numerical values were cross-checked against the source records by a second reviewer.

2.5. Outcome definitions and harmonization

The primary outcome was all-cause mortality. We recognized at the outset that the timepoint of mortality ascertainment differed across trials—ranging from ICU or 28-day mortality in the general critical-care and head-injury trials, through 30-day mortality in the

largest pragmatic trial,[3] to 6-month mortality in the severe-stroke trial[6]—and that mortality was a secondary rather than a primary endpoint in the two most recent trials, which were designed around functional outcome and ventilation duration respectively.[6,7] The exact timepoint extracted from each trial is reported in Table 1, and the implications of pooling mortality measured over heterogeneous horizons are addressed explicitly in the discussion. Secondary outcomes comprised VAP, the duration of mechanical ventilation and ventilator-free days; the diagnostic criteria for VAP varied across trials and eras, a limitation that is given due weight in the interpretation.

2.6. Risk of bias and certainty of evidence

The methodological quality of each included trial was appraised with the Cochrane RoB 2 tool across its five domains: bias arising from the randomization process, bias due to deviations from intended interventions, bias due to missing outcome data, bias in measurement of the outcome and bias in selection of the reported result.[9] Each domain, and the overall judgement, was rated as low risk, some concerns, or high risk of bias. Because tracheostomy is an overt procedure, blinding of participants and clinicians was not feasible in any trial; the impact of this unavoidable lack of blinding was therefore considered domain by domain, with particular attention to the measurement of subjective outcomes such as VAP, for which unblinded ascertainment is a recognized source of detection bias. The certainty of the evidence for each outcome was additionally rated using the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) framework, considering risk of bias, inconsistency, indirectness, imprecision and the possibility of publication bias, and was summarized as high, moderate, low or very low.[10]

2.7. Statistical analysis

All analyses were performed using a DerSimonian–Laird random-effects model, chosen a priori because clinical and methodological diversity across populations, timing thresholds and eras was anticipated.[11] For dichotomous outcomes, effect sizes were expressed as risk ratios (RR) with 95% CIs; corresponding absolute risk differences were derived from the trial-level event data to convey the practical magnitude of any effect. For continuous outcomes, the standardized mean difference was computed as Hedges' g, which incorporates a correction for small-sample bias, with 95% CIs. Statistical heterogeneity was quantified using the I^2 statistic and the Cochran Q test, with I^2 values of approximately 25%, 50% and 75% interpreted as low, moderate and high heterogeneity, respectively.[12] For the primary outcome, a prediction interval was derived to convey the likely range of true effects in future settings, and this interval—rather than the confidence interval alone—was treated as the principal expression of uncertainty.

A pre-specified subgroup analysis compared trials using the percutaneous dilatational technique with those using a surgical or mixed technique; in view of the small number of trials per subgroup, this analysis was regarded as hypothesis-generating. The robustness of the primary outcome was examined through leave-one-out sensitivity analysis and by re-estimating the pooled effect with the Hartung–Knapp–Sidik–Jonkman (HKSJ) variance correction.[13] Because the continuous outcomes were informed by only two trials each, these analyses were pre-specified as exploratory. Small-study effects and potential publication bias were explored solely by qualitative visual inspection of a contour-enhanced funnel plot; no numerical asymmetry statistic was computed, because formal testing is unreliable when fewer than ten trials are available.[14] A two-sided p-value below 0.05 was considered statistically significant.

3. Results

3.1. Study selection

The database search yielded 2,401 records. After title and abstract screening, 11 randomized records on tracheostomy timing were assessed in full, of which two were excluded as off-topic trials comparing tracheostomy techniques or airway devices rather than timing. Nine RCTs, enrolling 2,500 critically ill adults, met all eligibility criteria and were included in the quantitative synthesis. The complete selection process is presented in the PRISMA flow diagram in Figure 1; as Figure 1 shows, the totals reconcile exactly with the

nine included trials and the patient counts reported below.

3.2. Characteristics of included studies

The characteristics of the nine included trials are detailed in Table 1.[2,3,4,6,7,15-18] As Table 1 shows, the trials were heterogeneous with respect to population and setting, spanning general medical and surgical ICUs, isolated severe head injury, post-cardiac surgery, severe stroke and COVID-19. Definitions of the early strategy ranged from within 48 hours to within eight days of intubation, whereas the late or control strategy ranged from day 10 to day 16, or comprised prolonged intubation with tracheostomy only if still indicated. Importantly, the comparison embodied in these trials is not strictly one procedure against another but one management strategy against another: in several trials a substantial proportion of patients allocated to the late or control arm never underwent tracheostomy at all, because they were extubated or died beforehand.

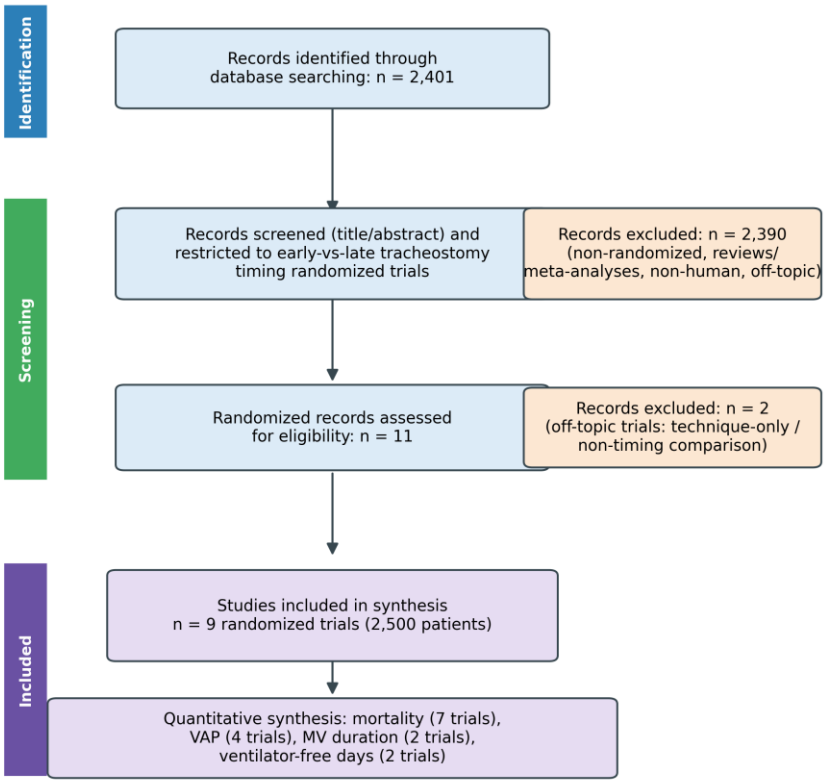

Figure 1. PRISMA flow diagram of study identification, screening and inclusion.

As recorded in Table 1, the proportion of the control arm that actually received a tracheostomy was 44.9% in the largest pragmatic trial, 56.7% in the largest pneumonia-focused trial, and 67.0% in the severe-stroke trial, underscoring that an early policy commits to the procedure a number of patients who would otherwise have avoided it. Four trials used the percutaneous dilatational technique exclusively or predominantly, and the remainder used a surgical or mixed approach. Sample sizes ranged from 62 to 909 participants.

3.3. Risk of bias

The risk-of-bias appraisal is displayed in Figure 2. As Figure 2 illustrates, the two largest and most recent multicenter trials (TracMan and SETPOINT2) were judged to be at low overall risk of bias, reflecting registered protocols, intention-to-treat analyses and objective primary outcomes. Three further multicenter trials (Blot, Terragni and Trouillet) were judged to be at low risk in most domains but raised some concerns in the domain of deviations from intended interventions or selective reporting, principally because the open nature of the intervention could influence the

unblinded ascertainment of softer outcomes such as VAP. The older single-center trials (Bouderka, Rumbak and Zheng), although randomized, raised some concerns across several domains owing to the absence of trial registration, limited allocation-concealment detail and single-center conduct. No domain was rated as conferring an unequivocally high risk of bias. A central and unavoidable consideration, evident across the outcome-measurement domain in Figure 2, was that because the intervention is an overt surgical procedure no trial could blind clinicians or outcome assessors; this is of particular consequence for VAP, the diagnosis of which is partly subjective, and the possibility that the apparent reduction in VAP is at least partly an artifact of detection bias cannot be excluded.

Table 1. Characteristics of the nine included randomized controlled trials.

Study (year)	Country	Population / setting	Early	Late / control	N (E/L)	Control received trach.	Mortality timepoint	Technique
Bouderka (2004)	Morocco	Severe head injury	Day 5	Prolonged intubation	31/31	NR	ICU	Surgical
Rumbak (2004)	USA	Medical ICU	≤48 h	Day 14–16	60/60	100% (by design)	ICU	Percutaneous
Blot (2008)	France	Mixed ICU	≤4 d	Prolonged intubation	61/62	NR	28-day	Open/percut.
Terragni (2010)	Italy	Adult ICU	6–8 d	13–15 d	209/210	56.7% (119/210)	28-day	Mixed
Trouillet (2011)	France	Post-cardiac surgery	Immediate	Trach day 15	109/107	NR	28/60/90-day	Percutaneous
Zheng (2012)	China	Surgical ICU	Day 3	Day 15	58/61	NR	60-day	Percutaneous
Young/TracMan (2013)	UK	General ICU	≤4 d	≥10 d	455/454	44.9%	30-day	Mixed
Bösel/SETPOINT2 (2022)	USA/Germany	Severe stroke	≤5 d	Standard ≥d10	188/194	67.0%	6-month	Predom. percut.
Eeg-Olofsson/TTCOV19 (2022)	Sweden	COVID-19	≤7 d	≥10 d	72/78	68% (overall)	ICU	Mixed

Notes: NR = not reported; E = early; L = late.

Figure 2. Cochrane Risk of Bias 2.0 (RoB 2) traffic-light summary for the nine included trials.

3.4. Primary outcome: all-cause mortality

Seven trials, comprising 2,242 patients, reported extractable all-cause mortality data. The mortality timepoint differed across these trials, as detailed in Table 1, ranging from ICU and 28-day mortality to 30-day and, in the stroke trial, 6-month mortality; this heterogeneity of horizon is an important caveat to the pooled estimate. As shown in the forest plot in Figure 3, early tracheostomy was not associated with a statistically significant reduction in all-cause mortality compared with late tracheostomy (RR 0.88, 95% CI 0.70–1.09; p=0.24). Expressed in absolute terms, the largest pragmatic trial reported 30-day mortality of 30.8% versus 31.5% (absolute risk difference −0.7%, 95% CI −5.4% to 6.7%), illustrating that any true effect on survival, if present, is small.[3] Moderate statistical heterogeneity was present (I^2=58.9%; Cochran Q=14.61, df=6, p=0.02), and, as Figure 3 also displays, the prediction interval (0.48–1.59) was wide, indicating that the true effect in an individual future setting could plausibly favor either strategy.

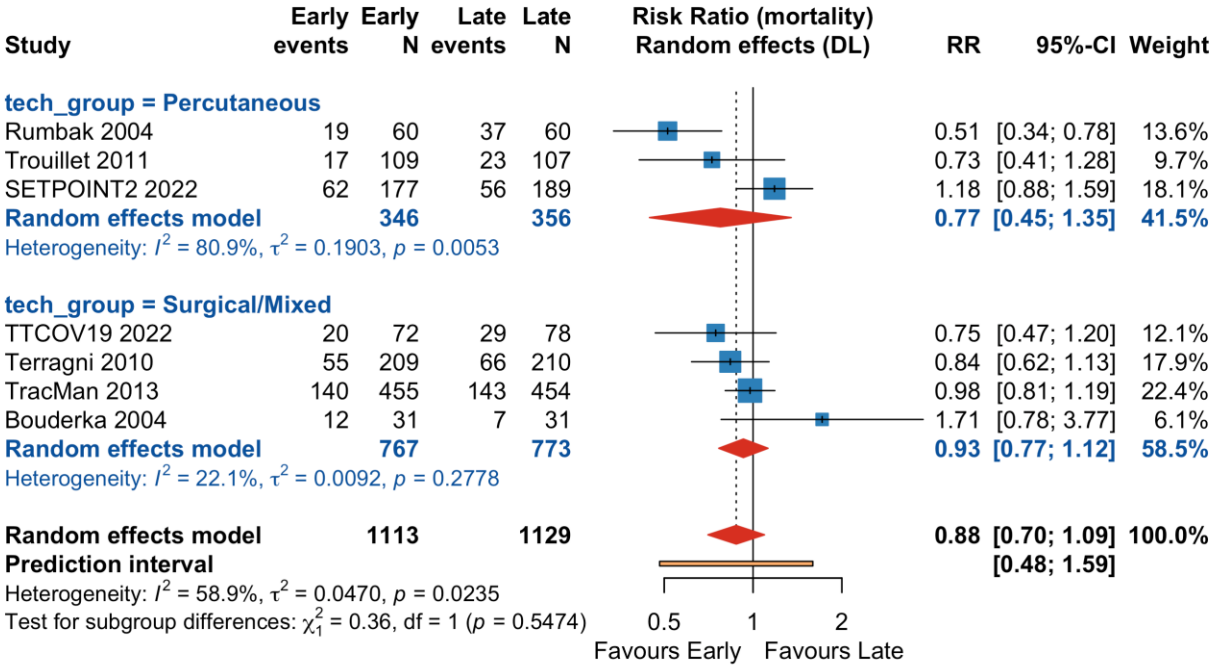

Figure 3. Forest plot of all-cause mortality (random-effects, DerSimonian–Laird estimator), stratified by tracheostomy technique, with prediction interval.

In the pre-specified subgroup analysis shown in Figure 3, trials employing the percutaneous dilatational technique yielded a point estimate that numerically favored early tracheostomy (RR 0.77, 95% CI 0.45–1.35) but with very high heterogeneity (I^2=80.9%), whereas trials using a surgical or mixed technique produced a near-null estimate (RR 0.93, 95% CI 0.77–1.12; I^2=22.1%). The test for subgroup differences was not statistically significant (p=0.55). Because the percutaneous subgroup contained only three trials and was dominated, in both its point estimate and its heterogeneity, by a single small single-center study,[2] this subgroup comparison is at high risk of being spurious and should be regarded as hypothesis-generating only; we explicitly caution against inferring that the technique modifies the effect of timing.

3.5. Ventilator-associated pneumonia

Four trials, comprising 874 patients, reported VAP as a dichotomous outcome. As shown in the forest plot in Figure 4, early tracheostomy was associated with a non-significant trend towards a lower risk of VAP (RR 0.67, 95% CI 0.42–1.05; p=0.08), with substantial heterogeneity (I^2=71.0%). The direction of effect was consistent with a protective signal, but, as Figure 4 makes clear, the confidence interval crossed

unity and the result was strongly influenced by an early single-center trial reporting a particularly large effect.[2] Given the unblinded ascertainment of VAP in every trial and the evolution of its diagnostic criteria over the two decades spanned by the dataset, this signal should be interpreted as hypothesis-generating at best, and not as established evidence of benefit.

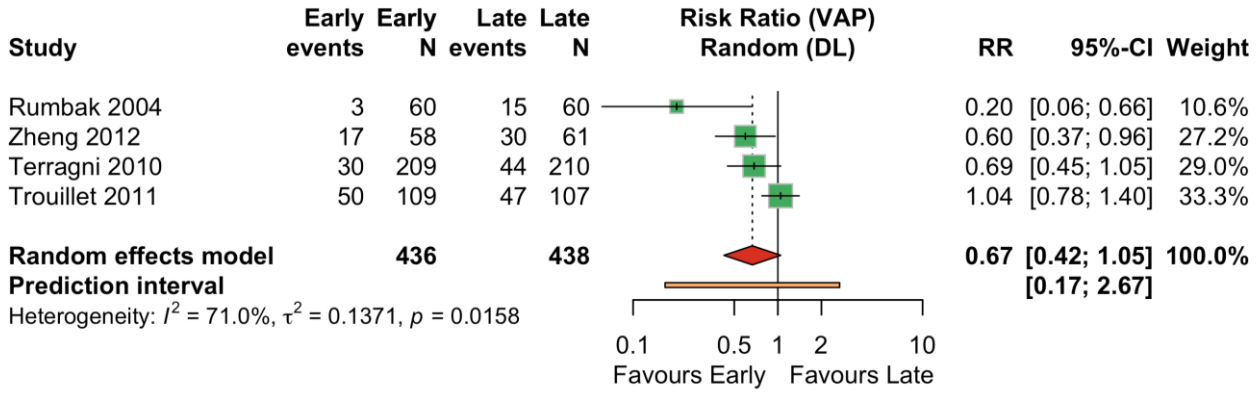

Figure 4. Forest plot of ventilator-associated pneumonia (random-effects, DerSimonian–Laird estimator).

3.6. Continuous outcomes (exploratory)

Owing to incomplete reporting of means and standard deviations across the dataset, only two trials provided usable data for each continuous outcome; these analyses were therefore pre-specified and are reported as exploratory, and are displayed together in Figure 5. For the duration of mechanical ventilation, the pooled standardized mean difference favored early tracheostomy (SMD −1.38, 95% CI −3.44 to 0.68, Hedges' g) but did not reach statistical significance and was accompanied by extreme heterogeneity (I^2=97.2%), reflecting the very different magnitudes reported by the contributing trials; as Figure 5 conveys, a single pooled estimate is of doubtful meaning at this level of inconsistency, and the result is presented for completeness rather than as a reliable summary. For ventilator-free days, the two contributing trials yielded a small, non-significant difference in favor of early tracheostomy (SMD 0.20, 95% CI −0.07 to 0.47; I^2=31.2%).

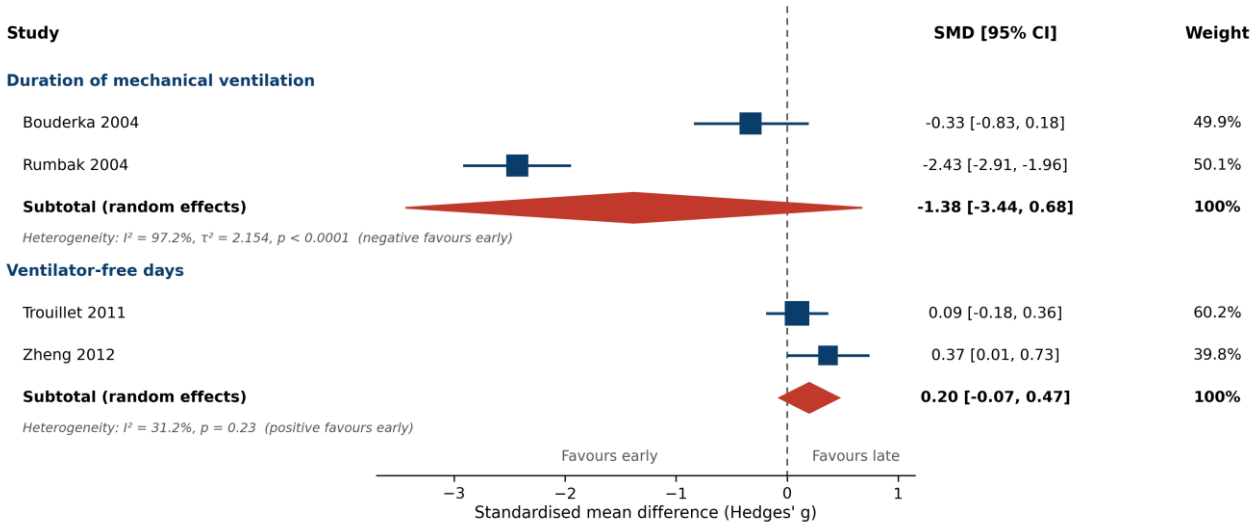

Figure 5. Forest plot of continuous outcomes expressed as standardized mean difference (Hedges' g); both analyses are exploratory (two trials each).

3.7. Sensitivity analyses

The neutral mortality result proved robust to leave-one-out analysis: the pooled risk ratio ranged from 0.82 to 0.96 across all iterations, and in no case did the 95% confidence interval exclude unity. The single most influential trial was the small single-center study by Rumbak;[2] when this trial was omitted, the pooled estimate moved further towards the null (RR 0.96, 95% CI 0.81–1.13) and heterogeneity fell markedly (I^2 from 58.9% to 26.4%), identifying that trial as the principal source of heterogeneity. Re-estimation of the primary outcome using the HKSJ variance correction widened the confidence interval (RR 0.88, 95% CI 0.65–1.19) without altering the conclusion of no significant effect.[13]

3.8. Publication bias

Visual inspection of the funnel plot for the primary mortality outcome, presented in Figure 6, revealed broadly symmetrical scatter of the trials about the pooled estimate, with the larger trials clustering near the summary line and the smaller trials dispersed more widely. Because only seven trials contributed to the primary outcome, formal statistical testing for funnel-plot asymmetry was not performed, as such tests are unreliable below ten studies;[14] Figure 6 was therefore interpreted purely qualitatively, and no gross evidence of small-study effects or selective non-publication was apparent.

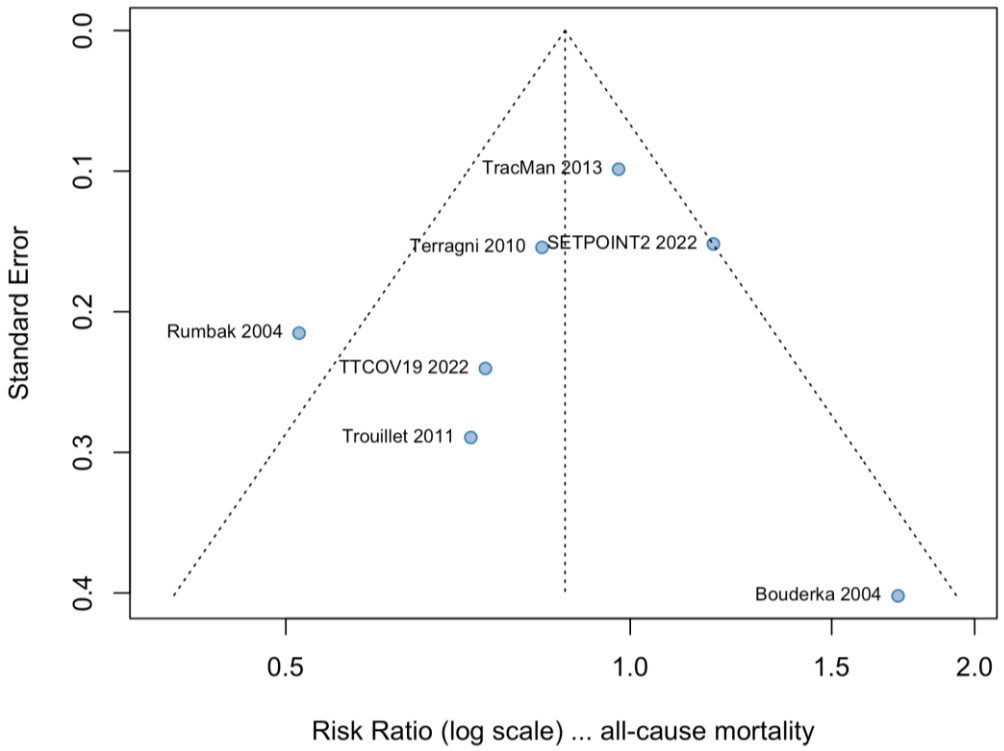

Figure 6. Contour-enhanced funnel plot for the primary outcome (all-cause mortality), interpreted qualitatively.

3.9. Certainty of evidence

The GRADE certainty assessment is summarized in Table 2. As Table 2 indicates, the certainty of evidence was rated as moderate for all-cause mortality (downgraded one level for inconsistency), low for VAP (downgraded for inconsistency and for the risk of detection bias inherent in an unblinded outcome), and very low for the continuous outcomes (downgraded for serious inconsistency and imprecision arising from only two contributing trials each). These ratings indicate that, while the absence of a mortality benefit can be asserted with reasonable confidence, the secondary findings should be regarded as uncertain.

Table 2. GRADE summary of findings.

Outcome	Trials (patients)	Pooled effect (95% CI)	I^2	Certainty	Principal reason for downgrading
All-cause mortality	7 (2,242)	RR 0.88 (0.70–1.09)	58.9%	Moderate	Inconsistency
Ventilator-associated pneumonia	4 (874)	RR 0.67 (0.42–1.05)	71.0%	Low	Inconsistency; detection bias (unblinded)
Duration of mechanical ventilation	2	SMD −1.38 (−3.44–0.68)	97.2%	Very low	Serious inconsistency; imprecision
Ventilator-free days	2	SMD 0.20 (−0.07–0.47)	31.2%	Very low	Imprecision (few trials)

4. Discussion

This updated meta-analysis of nine RCTs, enrolling 2,500 critically ill adults, found that early tracheostomy did not significantly reduce all-cause mortality compared with late tracheostomy, with a pooled risk ratio of 0.88 and a confidence interval that comfortably crossed unity. A non-significant trend towards a lower incidence of VAP and small, non-significant tendencies towards a shorter duration of mechanical ventilation and more ventilator-free days were observed, but none reached statistical significance, and the continuous outcomes were limited by extreme heterogeneity and by the small number of contributing trials. The wide prediction interval for the primary outcome (0.48–1.59) is, in our view, the single most informative expression of the current state of knowledge: it conveys that, although a survival benefit has not been demonstrated on average, the true effect in any particular future population could fall on either side of the null. On the basis of randomized evidence, early tracheostomy cannot be expected to deliver a survival advantage in the general population of critically ill ventilated adults, and the certainty of this conclusion is moderate.

The neutral mortality result is consistent with the trajectory of the highest-quality individual trials. The two largest and methodologically strongest trials—a pragmatic multicenter trial in general critical care and a multicenter trial in severe stroke—each found no significant difference in their respective primary endpoints, and both were judged here to be at low risk of bias.[3,6] By contrast, the most striking benefits were reported by an early, small, single-center trial;[2] our leave-one-out analysis identified precisely this trial as the dominant source of heterogeneity, since its removal moved the pooled estimate towards the null and reduced I^2 from approximately 59% to 26%. This pattern, in which early small studies suggest large benefits that are not reproduced in subsequent large trials, is a recurrent theme across critical-care interventions and reinforces the primacy of adequately powered, multicenter randomized evidence.

4.1. Comparison with previous meta-analyses

The present results align with the broad direction of recent quantitative syntheses while extending them. Earlier Cochrane reviews concluded that the evidence was, at best, only suggestive of a mortality benefit,[19] and a widely cited analysis in a leading respiratory journal—notably one whose initial version was retracted and republished with an attenuated, non-significant mortality estimate—ultimately reported no significant survival difference but a reduction in VAP.[5] Subsequent meta-analyses, including a trial-sequential analysis and earlier pooled syntheses, have consistently emphasized that the apparent benefits of early tracheostomy on softer outcomes are fragile and that the cumulative evidence remains insufficient to confirm a mortality effect.[20,21,22] An umbrella review of randomized evidence reached a similar conclusion, reporting a reduction in VAP but no significant mortality difference, with trial-sequential analysis indicating that the available information size was inadequate for a definitive answer.[23] An earlier systematic review that also incorporated non-

randomized data favored early tracheostomy performed within seven days of intubation, a finding that should be read in the light of its less stringent design.[24] The specific contribution of the present synthesis, beyond confirmation, is the formal incorporation of the two most recent landmark trials within a single coherent framework, together with a transparent GRADE-anchored characterization of how much uncertainty persists for each outcome; in this respect the manuscript is intended not as a mere update but as a consolidation that clarifies what is, and what is not, established.

With respect to VAP, our non-significant protective trend (RR 0.67) is directionally concordant with previous analyses that reported modest reductions with an early strategy.[25] However, VAP is an outcome that is particularly vulnerable to ascertainment bias in unblinded trials and to inconsistent diagnostic criteria across studies and eras, which may inflate the apparent benefit. The wide confidence interval, the substantial heterogeneity, and the low GRADE certainty together counsel firmly against interpreting the VAP signal as definitive.

4.2. Heterogeneity and the limits of a single estimate

Heterogeneity was a defining feature of this evidence base. The included trials differed in the populations studied, in the precise thresholds used to define early and late strategies, in the tracheostomy technique, in the calendar era and corresponding standards of sedation and ventilator management, and in the outcomes prioritized. The timing thresholds themselves varied appreciably—from within 48 hours to within eight days for the early arm, and from day 10 to day 16 for the late arm—so that pooling assumes a broadly similar biological effect of timing across these windows, an assumption that may not hold and that the limited number of trials precludes testing formally. Such clinical and methodological diversity justified the a priori use of a random-effects model and the derivation of a prediction interval, which proved wide and underscored that a single summary estimate cannot capture the full range of plausible effects. The pre-specified subgroup analysis by technique did not formally explain the heterogeneity. These observations imply that the optimal timing of tracheostomy may genuinely differ across patient populations, and that pooling all critically ill patients into a single estimate, while informative, inevitably obscures potentially important effect modification.

A related and clinically fundamental source of diversity concerns the comparator. As quantified in Table 1, in several trials a large fraction of patients allocated to the late or control arm never underwent tracheostomy, because they were extubated or died first; where reported, only 45% to 67% of control-arm patients actually received the procedure.[3,4,6] The comparison is therefore best understood as one between a strategy of early tracheostomy and a strategy of watchful waiting with selective late tracheostomy, rather than between two procedures. Because clinicians predict the need for prolonged ventilation poorly, an early policy necessarily commits a substantial number of patients to an invasive procedure from which they would not have benefited. This predictive uncertainty is arguably the single most important determinant of the clinical value of an early policy, and it weighs against the routine adoption of early tracheostomy in unselected patients.

4.3. Population-specific perspective and patient-centered outcomes

Because the trials pooled clinically distinct entities—general critical illness, head injury, cardiac surgery, stroke and COVID-19—the global estimate may have limited direct applicability to any individual patient, who is always treated within a specific diagnostic context rather than as the average of these populations. We therefore caution against treating the single summary estimate as the principal take-home message, and we emphasize instead a population-specific reading: the largest neurological and general-ICU trials, which are the most directly informative for their respective populations, were neutral for their primary endpoints.[3,6] Indeed, a dedicated synthesis in acutely brain-injured patients suggested a possible long-term mortality benefit that was attenuated on sensitivity analysis, illustrating that effect estimates can differ by population.[26] At the same time, the outcomes on which an early strategy appears to exert its most reproducible effect are not mortality or even

pneumonia, but patient-centered outcomes such as the duration and depth of sedation, patient comfort, and the ease of nursing care; these were reported consistently in favor of early tracheostomy across several trials, even where harder outcomes were unaffected.[2,5,17] For selected patients in whom prolonged ventilation is confidently anticipated, these comfort-related benefits may legitimately influence the decision, and they deserve greater prominence in clinical reasoning than a narrow focus on survival would allow.

It is also instructive to consider the internal coherence of the outcome pattern. If early tracheostomy genuinely shortened mechanical ventilation and increased ventilator-free days, a corresponding signal might be expected in resource-related outcomes and, plausibly, in pneumonia. The observation that the continuous outcomes trend, weakly and inconsistently, in the direction of benefit while mortality remains unaffected is compatible with a modest true effect on the duration of ventilation and its sequelae that is insufficient to translate into a survival difference—an interpretation that is biologically plausible and consistent with the wider literature.[21,25]

4.4. Temporal context and generalizability

The included trials span nearly two decades, a period over which standards of sedation practice, ventilator management, infection-control bundles and the very definition of VAP have all evolved substantially. The pooled estimates therefore combine trials conducted under materially different background standards of care, a source of clinical heterogeneity distinct from the statistical heterogeneity captured by I^2. A sensitivity analysis restricted to the more contemporary trials was not feasible given their small number, but this temporal confounding should temper confidence in the pooled secondary outcomes in particular. With respect to generalizability, most of the included trials were conducted in well-resourced centers, and the applicability of the findings to resource-limited settings—where the costs of the procedure, the availability of bronchoscopic guidance, the operator experience required for safe percutaneous tracheostomy, and the consequences of complications may all differ—cannot be assumed.[1] In such settings, the absence of a demonstrated survival benefit, combined with the procedural risks and resource implications of an early policy, may argue still more strongly for a conservative, individualized approach; this consideration is directly relevant to much of the readership of this journal.

4.5. Strengths and limitations

The strengths of this study include its restriction to randomized evidence, the inclusion of the two most recent landmark trials, the dual appraisal of dichotomous and continuous outcomes within a coherent random-effects framework, the duplicate and independent study selection and data extraction, the application of the contemporary RoB 2 tool and the GRADE framework, and the robustness checks afforded by leave-one-out and HKSJ sensitivity analyses. Several limitations must, nonetheless, be acknowledged. First, substantial clinical and statistical heterogeneity, with I^2 reaching 97% for the duration of mechanical ventilation, limits the confidence with which any pooled estimate—particularly for the continuous outcomes—can be interpreted; for that outcome a single summary is of doubtful meaning, and we present it only for completeness. Second, the continuous outcomes rested on only two trials each, because several trials did not report extractable means and standard deviations, so these analyses are exploratory and hypothesis-generating rather than conclusive. Third, the unavoidable absence of blinding in trials of an overt surgical procedure introduces a risk of ascertainment bias for subjective outcomes such as VAP and may inflate the apparent benefit of an early strategy. Fourth, the included trials were markedly heterogeneous in population, timing definitions, technique and era, and they pooled distinct clinical entities, so that the summary estimates may not apply uniformly to any single patient group; moreover, mortality was a secondary outcome in the stroke and COVID-19 trials and was measured over heterogeneous time horizons across the dataset, ranging from ICU to six-month mortality. Fifth, the search was based principally on a single database, supplemented by reference-list screening; although this captured all the major randomized trials, a formal

multi-database and grey-literature search would be preferable for a definitive synthesis. These limitations collectively define the boundaries within which the present conclusions should be read.

4.6. Implications for practice and research

For the practising intensivist, the principal implication of this synthesis is that the decision of when to perform a tracheostomy should not be driven by an expectation of improved survival. Because early tracheostomy does not reduce mortality and offers, at most, modest and uncertain benefits on pneumonia and ventilation-related resource use, an indiscriminate early policy would expose a substantial number of patients—those who would have been successfully extubated within the first one to two weeks—to an unnecessary invasive procedure. The timing of tracheostomy is therefore best individualized, taking account of the underlying diagnosis, the trajectory of recovery, the anticipated duration of ventilation, the burden of sedation and patient comfort, rather than applied as a uniform early protocol. For research, the field would be best served by adequately powered, population-specific randomized trials—particularly in neurological and post-surgical cohorts in whom prolonged ventilation can be predicted with greater confidence—that adopt standardized definitions of early and late timing, harmonized and objectively ascertained outcome measures, and patient-centered endpoints such as sedation exposure and long-term functional and quality-of-life outcomes.[26] Embedding such trials within contemporary sedation and infection-control practice would address the temporal confounding that limits the present synthesis and would, at last, allow the residual uncertainty surrounding the secondary benefits of early tracheostomy to be resolved.

5. Conclusion

In this updated meta-analysis of nine randomized controlled trials enrolling 2,500 critically ill adults, early tracheostomy did not significantly reduce all-cause mortality compared with a late tracheostomy strategy, with a pooled risk ratio of 0.88 (95% CI 0.70–1.09) and a wide prediction interval (0.48–1.59) that encompassed both benefit and harm; the certainty of this neutral finding was moderate. Early tracheostomy was associated with a non-significant trend towards a lower incidence of ventilator-associated pneumonia and with small, non-significant tendencies towards a shorter duration of mechanical ventilation and more ventilator-free days; however, these secondary findings were of low to very low certainty, constrained by substantial heterogeneity, by the limited number of trials reporting extractable continuous data, and by the susceptibility of unblinded outcomes to ascertainment bias. Sensitivity analyses, including leave-one-out and Hartung–Knapp–Sidik–Jonkman approaches, confirmed that the neutral mortality result was robust and was not dependent on any single trial, while identifying an early small single-center study as the principal source of heterogeneity. The totality of the randomized evidence, now strengthened by the inclusion of the two most recent landmark trials in severe stroke and in COVID-19, indicates that the timing of tracheostomy does not materially influence survival in the general population of critically ill ventilated adults. Because clinicians predict the need for prolonged ventilation poorly, and because a substantial fraction of conservatively managed patients never require the procedure at all, these findings argue against the adoption of a uniform early tracheostomy policy aimed at improving survival and support an individualized approach in which timing is determined by the underlying diagnosis, the anticipated duration of mechanical ventilation, the recovery trajectory, and considerations of sedation burden and patient comfort. The most reproducible advantages of an early strategy lie in reduced sedation and improved comfort rather than in survival, and these may reasonably inform the decision in selected patients in whom prolonged ventilation is confidently anticipated. Until adequately powered, population-specific trials with harmonized outcomes are available, the timing of tracheostomy should remain a matter of individualized clinical judgement rather than a fixed protocol.

Declarations

Funding: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Conflicts of interest: The authors declare that they have no competing interests.

Author contributions: F.N.A. conceived the study, designed the search strategy, performed study selection and data extraction, conducted the statistical analysis and drafted the manuscript. N.A., J. and W.N.W. contributed to study selection, data extraction, risk-of-bias appraisal, interpretation of the results and critical revision of the manuscript. All authors read and approved the final version.

Data availability: All data analyzed in this study were extracted from previously published randomized controlled trials, which are publicly available and cited in the reference list.

6. References

1. Cheung NH, Napolitano LM. Tracheostomy: epidemiology, indications, timing, technique, and outcomes. Respir Care. 2014; 59(6): 895–915. doi: 10.4187/respcare.02971.
2. Rumbak MJ, Newton M, Truncale T, et al. A prospective, randomized study comparing early percutaneous dilational tracheotomy to prolonged translaryngeal intubation (delayed tracheotomy) in critically ill medical patients. Crit Care Med. 2004; 32(8): 1689–94. doi: 10.1097/01.ccm.0000134835.05161.b6.
3. Young D, Harrison DA, Cuthbertson BH, et al. Effect of early vs late tracheostomy placement on survival in patients receiving mechanical ventilation: the TracMan randomized trial. JAMA. 2013; 309(20): 2121–9. doi: 10.1001/jama.2013.5154.
4. Terragni PP, Antonelli M, Fumagalli R, et al. Early vs late tracheotomy for prevention of pneumonia in mechanically ventilated adult ICU patients: a randomized controlled trial. JAMA. 2010; 303(15): 1483–9. doi: 10.1001/jama.2010.447.
5. Siempos II, Ntaidou TK, Filippidis FT, et al. Effect of early versus late or no tracheostomy on mortality and pneumonia of critically ill patients receiving mechanical ventilation: a systematic review and meta-analysis. Lancet Respir Med. 2015; 3(2): 150–8. doi: 10.1016/S2213-2600(15)00007-7.
6. Bösel J, Niesen WD, Salih F, et al. Effect of early vs standard approach to tracheostomy on functional outcome at 6 months among patients with severe stroke receiving mechanical ventilation: the SETPOINT2 randomized clinical trial. JAMA. 2022; 327(19): 1899–909. doi: 10.1001/jama.2022.4798.
7. Eeg-Olofsson M, Pauli N, Hafsten L, et al. TTCOV19: timing of tracheotomy in SARS-CoV-2-infected patients: a multicentre, single-blinded, randomized, controlled trial. Crit Care. 2022; 26(1): 142. doi: 10.1186/s13054-022-04005-0.
8. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021; 372: n71. doi: 10.1136/bmj.n71.
9. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019; 366: l4898. doi: 10.1136/bmj.l4898.
10. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008; 336(7650): 924–6. doi: 10.1136/bmj.39489.470347.AD.
11. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986; 7(3): 177–88. doi: 10.1016/0197-2456(86)90046-2.
12. Higgins JPT, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. BMJ. 2003; 327(7414): 557–60. doi: 10.1136/bmj.327.7414.557.
13. IntHout J, Ioannidis JPA, Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. BMC Med Res Methodol. 2014; 14: 25. doi: 10.1186/1471-2288-14-25.
14. Egger M, Davey Smith G, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997; 315(7109): 629–34. doi: 10.1136/bmj.315.7109.629.

15. Bouderka MA, Fakhir B, Bouaggad A, et al. Early tracheostomy versus prolonged endotracheal intubation in severe head injury. J Trauma. 2004; 57(2): 251–4.
doi: 10.1097/01.ta.0000087646.68382.9a.
16. Blot F, Similowski T, Trouillet JL, et al. Early tracheotomy versus prolonged endotracheal intubation in unselected severely ill ICU patients. Intensive Care Med. 2008; 34(10): 1779–87.
doi: 10.1007/s00134-008-1195-4.
17. Trouillet JL, Luyt CE, Guiguet M, et al. Early percutaneous tracheotomy versus prolonged intubation of mechanically ventilated patients after cardiac surgery: a randomized trial. Ann Intern Med. 2011; 154(6): 373–83.
doi:10.7326/0003-4819-154-6-201103150-00002.
18. Zheng Y, Sui F, Chen XK, et al. Early versus late percutaneous dilational tracheostomy in critically ill patients anticipated requiring prolonged mechanical ventilation. Chin Med J (Engl). 2012; 125(11): 1925–30.
19. Andriolo BNG, Andriolo RB, Saconato H, et al. Early versus late tracheostomy for critically ill patients. Cochrane Database Syst Rev. 2015; 1: CD007271.
doi: 10.1002/14651858.CD007271.pub3.
20. Wang R, Pan C, Wang X, et al. The impact of tracheotomy timing in critically ill patients undergoing mechanical ventilation: a meta-analysis of randomized controlled clinical trials with trial sequential analysis. Heart Lung. 2019; 48(1): 46–54.
doi: 10.1016/j.hrtlng.2018.09.005.
21. Hosokawa K, Nishimura M, Egi M, et al. Timing of tracheotomy in ICU patients: a systematic review of randomized controlled trials. Crit Care. 2015; 19(1): 424.
doi: 10.1186/s13054-015-1138-8.
22. Deng H, Fang Q, Chen K, et al. Early versus late tracheotomy in ICU patients: a meta-analysis of randomized controlled trials. Medicine (Baltimore). 2021; 100(3): e24329.
doi: 10.1097/MD.0000000000024329.
23. Boni A, Tonietto TA, Rihl MF, et al. Early versus late tracheostomy in critically ill patients: an umbrella review of systematic reviews of randomised clinical trials with meta-analyses and trial sequential analysis. BMJ Open Respir Res. 2025; 12(1): e002434.
doi: 10.1136/bmjresp-2024-002434.
24. Adly A, Youssef TA, El-Begermy MM, et al. Timing of tracheostomy in patients with prolonged endotracheal intubation: a systematic review. Eur Arch Otorhinolaryngol. 2018; 275(3): 679–90. doi: 10.1007/s00405-017-4838-7.
25. Chorath K, Hoang A, Rajasekaran K, et al. Association of early vs late tracheostomy placement with pneumonia and ventilator days in critically ill patients: a meta-analysis. JAMA Otolaryngol Head Neck Surg. 2021; 147(5): 450–9. doi: 10.1001/jamaoto.2021.0025.
26. McCredie VA, Alali AS, Scales DC, et alet al. Effect of early versus late tracheostomy or prolonged intubation in critically ill patients with acute brain injury: a systematic review and meta-analysis. Neurocrit Care. 2017; 26(1): 14–25.
doi: 10.1007/s12028-016-0297-z.