Supplement 3: additional analyses

Mortality Effect of Albumin Fluid Resuscitation in Adults with Septic Shock: A Systematic Review and Dual Frequentist–Bayesian Meta-Analysis of Randomised Trials

Authors:

Henrique Mendes, MD, MRCEM[1,2]

Henry Kinkei Wen, MD[1,2]

Harry de Souza, MD[1,2]

Yuanyuan Peng, MD[1,2]

Kush Deshpande, MD, FCICM[1,2]

Siebe G. Blok, MD, MSc[3]

Pieter R. Tuinman, MD, PhD[3,4]

Ary Serpa Neto, MD, PhD, FCICM[5,6,7,8]

Marcus J. Schultz, MD, PhD[9,10,11,12]

Manu L. N. G. Malbrain, MD, PhD[13,14]

Giacomo Grasselli, MD[15,16]

Frank M.P. van Haren, MD, PhD, FCICM[1,2,17,18]

[1]St George Hospital, Sydney, Australia [2]University of New South Wales, Sydney, Australia [3]Department of Intensive Care Medicine, Amsterdam UMC, Amsterdam, The Netherlands [4]Amsterdam Cardiovascular Sciences Research Institute, Amsterdam UMC, Amsterdam, The Netherlands [5]Australian and New Zealand Intensive Care Research Centre (ANZIC-RC), School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia [6]Department of Intensive Care, Austin Hospital, Melbourne, Australia [7]Department of Critical Care, Melbourne Medical School, University of Melbourne, Austin Hospital, Melbourne, Australia [8]Department of Critical Care Medicine, Hospital Israelita Albert Einstein, Sao Paulo, Brazil [9]Medical University Wien, Vienna, Austria [10]Cantonal Hospital St. Gallen, HOCH Health Ostschweiz, St. Gallen, Switzerland [11]University of Oxford, Oxford, UK [12]Mahidol University, Bangkok, Thailand [13]First Department of Anaesthesiology and Intensive Therapy, Medical University of Lublin, Jaczewskiego Lublin, Poland [14]International Fluid Academy, Lovenjoel, Belgium [15]Department of Emergency, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy [16]Department of Pathophysiology and Transplantation, University of Milan, Milan, Italy [17]Australian National University, Canberra, Australia [18]The George Institute for Global Health, Sydney, Australia

Table of Contents

- S1: Risk of Bias Assessment: Domain-Level Justifications (RoB 2.0)

Risk of bias for all-cause mortality was assessed using RoB 2 under the effect-of-assignment framework. Two trials were judged at high overall risk of bias: Rackow et al. (1983), owing to a predetermined non-concealed allocation schedule, and EARSS, primarily because only conference-abstract reporting was available, precluding full verification of prespecification and reported analyses. The remaining five trials (SAFE, ALBIOS, ICARUS-ED, Cusack et al., and ARISS) were judged as having some concerns overall. Across studies, bias from missing outcome data and measurement of the outcome was generally low, reflecting near-complete follow-up in the peer-reviewed trials and the objective nature of all-cause mortality. The main reasons for some-concerns judgements were the use of unpublished subgroup mortality data obtained through personal communication (SAFE), post hoc subgroup analysis together with control-arm albumin contamination (ALBIOS), post-registration addition of mortality-related endpoints in an underpowered feasibility trial (ICARUS-ED), and substantial control-arm albumin contamination together with premature trial termination (ARISS). All assessments are for the outcome of all-cause mortality, analysed under the effect-of-assignment framework. Two reviewers independently assessed risk of bias; disagreements were resolved by consensus.

Rackow 1983

Citation: Rackow et al., Critical Care Medicine, 1983

Outcome assessed: Study-period / in-hospital mortality (septic shock subgroup: albumin vs saline)

Analysis framework: Effect of assignment to intervention

Domain	Judgement	Support for judgement
Domain 1: Bias arising from the randomisation process	High	The publication describes a 'predetermined randomisation schedule,' which implies a fixed, non-concealed allocation list. No detail is provided on sequence generation methodology, and no allocation concealment mechanism is described. A predetermined schedule without concealment allows foreknowledge of upcoming assignments and is classified as high risk. The very small septic subgroup (n = 11) further limits the capacity of any randomisation method to balance prognostic factors.
Domain 2: Bias due to deviations from intended interventions	Low	The outcome assessed is all-cause mortality, which is objective and not susceptible to bias from knowledge of treatment assignment. For a hard endpoint in an

		acute resuscitation trial with protocolised haemodynamic targets, deviations from intended interventions are unlikely to have differentially affected mortality.
Domain 3: Bias due to missing outcome data	Low	Hospital mortality was reported for all analysed septic shock patients, with no loss to follow-up described. Small sample size increases imprecision, not bias from missing outcome data.
Domain 4: Bias in measurement of the outcome	Low	All-cause mortality is an objective endpoint. Ascertainment is unlikely to differ between arms regardless of blinding status.
Domain 5: Bias in selection of the reported result	Some concerns	No trial registration or published protocol is available (predates registration requirements). The primary outcome of the original trial was haemodynamic variables, not mortality; however, mortality was tabulated by individual patient, and the risk of selective non-reporting is low. The absence of a prespecified analysis plan precludes a definitive low-risk judgement.
Overall risk of bias	High	Driven by Domain 1. A predetermined randomisation schedule without described concealment constitutes a fundamental methodological limitation. Combined with the very small subgroup size and absence of a prespecified protocol (Domain 5), the overall risk of bias is high.

SAFE 2004

Citation: Finfer et al. (SAFE Study Investigators), NEJM, 2004; ICM, 2011

Outcome assessed: 28-day mortality (septic shock subgroup, via Xu et al. 2014)

Analysis framework: Effect of assignment to intervention

Domain	Judgement	Support for judgement
Domain 1: Bias arising from the randomisation process	Low	Central computer-based minimisation algorithm with stratification by site. Allocation concealment maintained through a centrally controlled web-based system. Study fluids (4% albumin and 0.9% saline) supplied in identical, blinded containers. Baseline

		characteristics well balanced in the overall trial and severe sepsis subgroup.
Domain 2: Bias due to deviations from intended interventions	Low	Double-blinded with masked study fluids. For all-cause mortality, deviations from intended interventions are unlikely to have introduced bias. Non-study fluid use was similar between arms (albumin 8.8%, saline 10.7%) and is not expected to differentially affect mortality.
Domain 3: Bias due to missing outcome data	Low	Follow-up for 28-day mortality was near-complete in the overall trial (99.25% albumin, 98.82% saline analysed). Losses were minimal and balanced. Completeness within the septic shock subgroup is assumed comparable.
Domain 4: Bias in measurement of the outcome	Low	All-cause mortality at 28 days is objective. The double-blind design provides additional protection, although blinding is not required for this endpoint.
Domain 5: Bias in selection of the reported result	Some concerns	The septic shock subgroup mortality data (70/209 vs 90/229) are not reported in any SAFE publication. They were obtained by Xu et al. through personal communication with Prof. Finfer and published as a secondary source. While the main trial had a published protocol and was registered, the specific septic shock subgroup analysis was not prespecified in the SAFE protocol, and the event counts cannot be independently verified from published sources.
Overall risk of bias	Some concerns	The SAFE trial is methodologically rigorous with low risk across Domains 1 through 4. The overall judgement is driven solely by Domain 5: the septic shock subgroup data derive from an unpublished, non-prespecified subgroup analysis communicated personally to a third-party research group.

EARSS 2011

Citation: Charpentier et al. (EARSS Study Group), ICM conference abstract, 2011

Outcome assessed: 28-day mortality

Analysis framework: Effect of assignment to intervention

Domain	Judgement	Support for judgement

Domain 1: Bias arising from the randomisation process	Some concerns	The abstract confirms that the trial was randomised and multicentre but provides no detail on sequence generation or allocation concealment. The large sample size (n = 798) and multicentre design are reassuring that a formal randomisation procedure was employed, but the mechanism cannot be verified from the abstract alone.
Domain 2: Bias due to deviations from intended interventions	Some concerns	The trial was open-label. While all-cause mortality is objective, the abstract provides no information on control-arm albumin contamination, protocol adherence, or co-interventions. The absence of verifiable detail on deviations from intended interventions in an open-label setting warrants some concerns, even for an objective endpoint.
Domain 3: Bias due to missing outcome data	Some concerns	The abstract reports results for the full randomised cohort (399 vs 399). However, no information on losses to follow-up, withdrawals, or the analysis population (ITT vs per-protocol) is available. Completeness of 28-day follow-up is plausible in an ICU population but cannot be confirmed from the abstract.
Domain 4: Bias in measurement of the outcome	Low	All-cause mortality at 28 days is an objective endpoint ascertained identically in both arms regardless of blinding status.
Domain 5: Bias in selection of the reported result	High	No trial registration, published protocol, or statistical analysis plan is available. The possibility that only selected results were presented cannot be excluded. The abstract was presented in 2011 and has never been published as a full manuscript, raising additional concern about selective non-publication of a large completed trial.
Overall risk of bias	High	Driven by Domain 5. The abstract-only status is the major driver of this high-risk classification due to incomplete reporting; this trial is included in the prespecified sensitivity analysis excluding studies at high risk of bias. Domains 1 through 3 carry some concerns on their own merits due to unverifiable methodological detail, while Domain 4 is low risk for the mortality outcome.

ALBIOS 2014

Citation: Caironi et al. (ALBIOS Study Investigators), NEJM, 2014

Outcome assessed: 90-day mortality (septic shock subgroup: cardiovascular SOFA 3 or 4)

Analysis framework: Effect of assignment to intervention

Domain	Judgement	Support for judgement
Domain 1: Bias arising from the randomisation process	Low	Computer-generated randomisation sequence with central allocation via a blinded, interactive web-based system, stratified by centre. Baseline characteristics well balanced. Prospectively registered (ClinicalTrials.gov NCT00707122) with a published protocol.
Domain 2: Bias due to deviations from intended interventions	Some concerns	Open-label design. For the objective mortality endpoint, open-label is not expected to directly bias outcome ascertainment. However, control-arm albumin contamination was substantial: 336 of 907 patients (37.1%) in the crystalloid-alone group received albumin at least once during the study period. While this biases the estimate toward the null rather than inflating the treatment effect, it represents a meaningful deviation from intended interventions.
Domain 3: Bias due to missing outcome data	Low	90-day follow-up available for >99% of patients in both groups. Intention-to-treat analysis. Losses minimal and balanced.
Domain 4: Bias in measurement of the outcome	Low	All-cause mortality at 90 days is objective. Open-label design does not affect ascertainment.
Domain 5: Bias in selection of the reported result	Some concerns	The trial was prospectively registered with a published protocol, and 90-day mortality was a prespecified secondary outcome. However, the result meta-analysed here is the septic shock subgroup (cardiovascular SOFA 3 or 4), which was a post hoc analysis rather than the parent trial's prespecified primary comparison. This introduces some concerns about selective reporting of a favourable subgroup result.
Overall risk of bias	Some concerns	Driven by Domains 2 and 5. Control-arm albumin contamination (37.1%) biases toward the null, and the meta-analysed result is a post hoc subgroup

		rather than the prespecified primary comparison. The trial is otherwise methodologically rigorous.

ICARUS-ED 2025

Citation: Williams et al. (ICARUS-ED), Annals of Emergency Medicine, 2025

Outcome assessed: 30-day mortality (prespecified refractory hypotension subgroup)

Analysis framework: Effect of assignment to intervention

Domain	Judgement	Support for judgement
Domain 1: Bias arising from the randomisation process	Low	Computer-generated block randomisation with stratification by site, using a central web-based allocation system. Allocation concealment maintained through the electronic system. Baseline characteristics balanced. Prospectively registered (ACTRN12619001059190).
Domain 2: Bias due to deviations from intended interventions	Low	Open-label design with standard-care comparator. The outcome is all-cause mortality, which is objective. The ED-based intervention was short (4 hours, 400 mL of 20% albumin), limiting the window for differential co-interventions. Protocol adherence >95%.
Domain 3: Bias due to missing outcome data	Low	30-day mortality follow-up reported for the full randomised cohort. Conducted within a healthcare system with linked records; vital status ascertainment expected to be near-complete.
Domain 4: Bias in measurement of the outcome	Low	All-cause mortality at 30 days is objective and not susceptible to measurement bias from unblinded assignment.
Domain 5: Bias in selection of the reported result	Some concerns	The trial was prospectively registered and the refractory hypotension subgroup was prespecified. However, as acknowledged in the publication, mortality, invasive ventilation, renal replacement therapy, and ward emergency calls were not included as endpoints on the original registry protocol; these were added before study commencement. This post-registration addition of clinical endpoints, combined with the study being a feasibility trial not powered for mortality, introduces some concerns about

		whether results would have been reported had they been unfavourable.
Overall risk of bias	Some concerns	Driven by Domain 5. The post-registration addition of mortality endpoints and the feasibility-trial design (underpowered for clinical outcomes) introduce some concerns about selective reporting. The trial is methodologically sound across Domains 1 through 4.

Cusack et al. 2025

Citation: Cusack et al., Journal of Critical Care, 2025

Outcome assessed: ICU mortality (all analysed patients)

Analysis framework: Effect of assignment to intervention

Domain	Judgement	Support for judgement
Domain 1: Bias arising from the randomisation process	Low	Prospective single-centre RCT. Randomisation was performed using an online computer system (supplemental materials). Allocation concealment was maintained via opaque envelopes (supplemental materials). Prospectively registered (NCT05357339). Three patients excluded post-randomisation (1 consent withdrawal, 2 poor image quality). Baseline characteristics balanced (Table 1). Small sample size (n = 100) but adequate sequence generation and concealment.
Domain 2: Bias due to deviations from intended interventions	Low	Open-label design with blinded microcirculation image analysis. For the objective all-cause mortality endpoint, open-label assignment does not directly bias ascertainment. Resuscitation was carried out per local procedures; all other treatments were at treating physician discretion. No information on control-arm albumin contamination is provided. The short, acute resuscitation intervention limits the window for differential co-interventions.
Domain 3: Bias due to missing outcome data	Low	Of 103 randomised patients, 3 were excluded (1 consent withdrawal, 2 poor image quality), leaving 100 analysed (50 per group). ICU survival was reported for all 100 patients. The 3 exclusions are documented with reasons; losses are minimal and balanced.
Domain 4: Bias in measurement of the outcome	Low	All-cause mortality (ICU survival) is an objective endpoint. Open-label design does not introduce measurement bias for this outcome.
Domain 5: Bias in selection of the reported result	Some concerns	The trial was prospectively registered (NCT05357339). The primary outcome was microcirculation, not mortality. 28-day mortality was listed as a secondary endpoint in the methods section but event counts were not separately reported in the results; only ICU survival (survived to

		ICU discharge) was tabulated. The study was powered for microcirculation (n = 42 per group), not clinical outcomes. The absence of the declared 28-day mortality counts, combined with the non-primary status of mortality, introduces some concerns about selective reporting.
Overall risk of bias	Some concerns	Driven by Domain 5. Randomisation used an online computer system with allocation concealment via opaque envelopes (supplemental materials); Domain 1 is low risk. 28-day mortality event counts were not reported despite being a declared secondary endpoint. Domain 2 is low risk given the objective mortality outcome and short intervention window. Domains 3 and 4 are low risk. Funded by an unrestricted grant from Grifols (albumin manufacturer).

ARISS 2026

Citation: Sakr et al. (ARISS), JAMA Network Open, 2026

Outcome assessed: 90-day mortality (mITT population)

Analysis framework: Effect of assignment to intervention

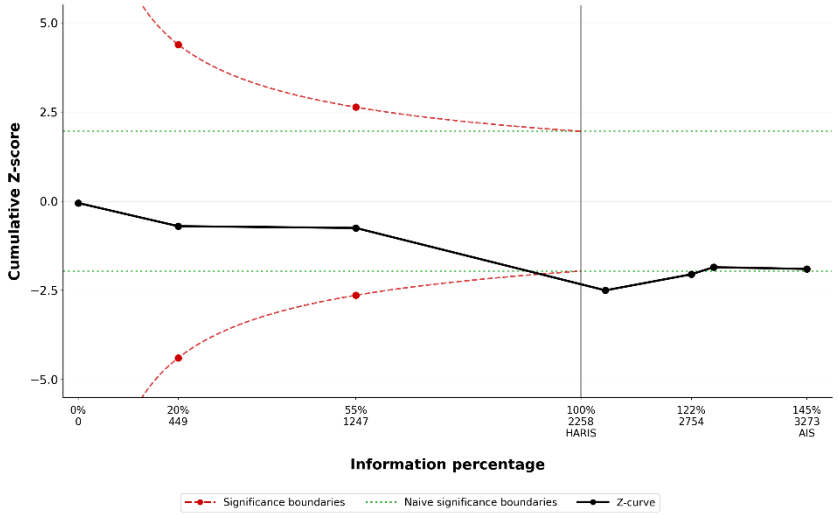
Domain	Judgement	Support for judgement
Domain 1: Bias arising from the randomisation process	**Low**	Computer-generated randomisation with central allocation, stratified by site. Allocation concealment maintained through a web-based system. Baseline characteristics well balanced. Prospectively registered (NCT03869385) with a published protocol.
Domain 2: Bias due to deviations from intended interventions	**Some concerns**	Open-label design. For the objective mortality endpoint, this does not directly bias ascertainment. However, albumin was administered to 106 of 218 control-group patients (48.6%) during their ICU stay at clinician discretion. This degree of contamination is substantial and severely dilutes the between-arm treatment contrast. While the direction of bias is toward the null, a 48.6% contamination rate represents a major deviation from intended interventions.
Domain 3: Bias due to missing outcome data	**Low**	The modified intention-to-treat population (210 vs 209) excludes patients who withdrew consent or were found ineligible post-randomisation. 90-day mortality reported for the full mITT population with minimal loss.
Domain 4: Bias in measurement of the outcome	**Low**	All-cause mortality at 90 days is objective. Open-label design does not introduce measurement bias.
Domain 5: Bias in selection of the reported result	**Some concerns**	The trial was prospectively registered with a published protocol and the primary outcome (90-day mortality) was prespecified. However, the trial was terminated prematurely due to low enrolment (440 of a planned 1662 patients) without a formal statistical stopping rule or data safety monitoring board recommendation. Premature termination increases the risk that the observed result is a

		chance finding and raises concerns about the completeness and reliability of the reported results.
Overall risk of bias	Some concerns	Driven by Domains 2 and 5. The 48.6% control-arm albumin contamination severely dilutes the treatment contrast (biasing toward the null), and premature termination without a formal stopping rule raises concerns about the reliability of the result. The trial is otherwise well designed with low risk across Domains 1, 3, and 4.

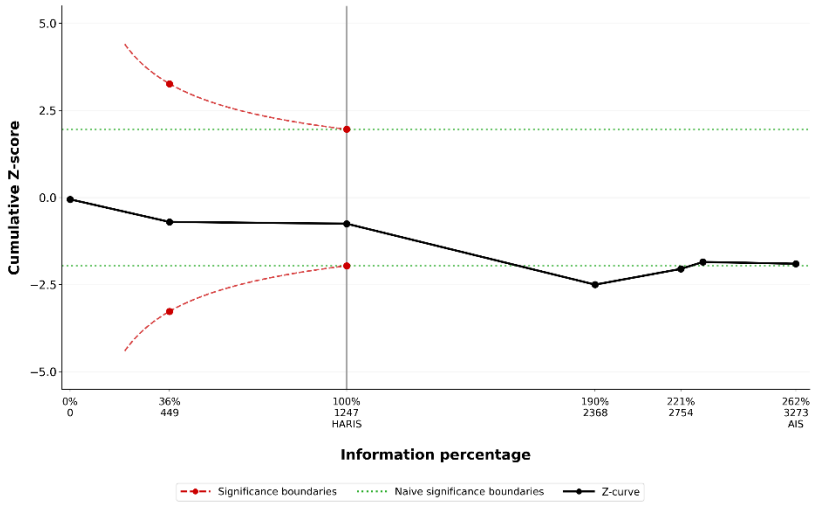
- S2: Trial Sequential Analysis — Sensitivity Analyses

Trial sequential analysis (TSA) was performed at two prespecified sensitivity effect sizes (RRR 10% and RRR 20%) to complement the primary TSA conducted at RRR 15% (reported in the main manuscript, eFigure S1a). All analyses used a two-sided α of 0.05, power of 80%, and an empirical control-arm event rate. The accrued information size (AIS) across all included trials was 3,273 participants.

eFigure S1a: TSA sensitivity analysis at RRR 15%

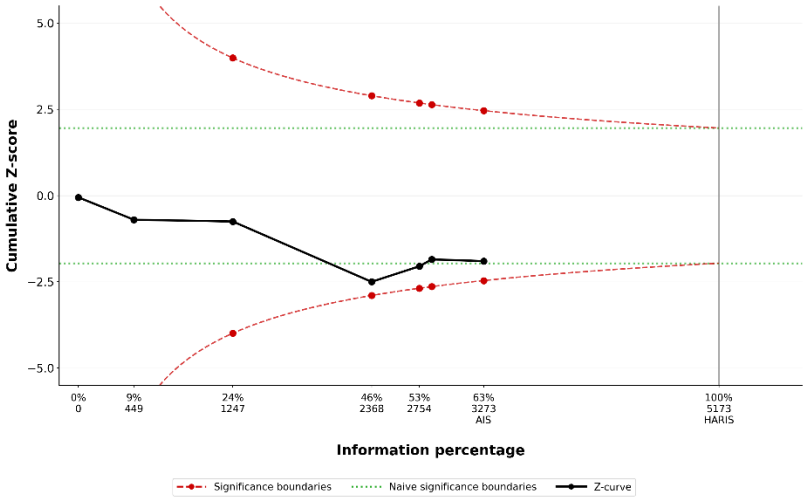


eFigure S1b: TSA sensitivity analysis at RRR 20%



At an assumed relative risk reduction of 20% (eFigure S1b), the heterogeneity-adjusted required information size (HARIS) was 1,247 participants. The AIS of 3,273 participants exceeded HARIS by 162%. The cumulative Z-curve crossed both the naïve and the TSA-adjusted significance boundaries in favour of albumin. These findings confirm that if the true treatment effect is a 20% relative risk reduction or larger, the existing evidence is more than sufficient to detect it, and the observed benefit is robust to the inflated type I error risk inherent in sequential testing.

eFigure S1c: TSA sensitivity analysis at RRR 10%



At an assumed relative risk reduction of 10% (eFigure S1c), the HARIS was 5,173 participants. The AIS of 3,273 participants reached only 63% of HARIS. The cumulative Z-curve crossed the naïve significance boundary but did not cross the wider TSA-adjusted significance boundary. This indicates that the accrued evidence is insufficient to confirm or reliably exclude a treatment effect as small as a 10% relative risk reduction; further trials enrolling approximately 1,900 additional participants would be required to reach the information threshold at this effect size.

- S3: Secondary outcome details

Secondary outcomes: organ-support-free days

The prespecified secondary outcome of organ-support-free days was reported inconsistently across studies and could not be meta-analysed. No included trial reported a harmonisable composite organ-support-free-day endpoint to day 28. Instead, studies reported a range of non-equivalent proxy outcomes: duration of mechanical ventilation, duration of renal replacement therapy, time to vasopressor or inotrope discontinuation, binary organ-support requirements at prespecified timepoints, and ICU-based ventilator- or vasopressor-free days defined to ICU discharge or death rather than to a fixed day-28 horizon.

Vasopressor-related outcomes

ARISS was the only trial to report vasopressor-free days explicitly: median 2 days (IQR 0–7) in the albumin group versus 2 (IQR 0–6) in controls, not statistically significant. In the ALBIOS overall severe sepsis population—not reported separately for the septic shock subgroup—the time to suspension of vasopressor or inotropic agents was significantly shorter in the albumin group (median 3 days, IQR 1–6) than in the crystalloid group (median 4 days, IQR 2–7; p=0.007), and the average cardiovascular SOFA subscore was lower throughout the study period (p=0.03). Supplementary data from ALBIOS add early semi-quantitative detail consistent with this signal: at 6 hours, achievement of MAP ≥65 mmHg was more frequent with albumin (86.0% vs 82.5%), and at day 1 the mean norepinephrine dose was lower despite similar proportions receiving norepinephrine (0.29±0.35 vs 0.33±0.36 μg/kg/min; 56.2% vs 59.1%). In ICARUS-ED, significantly fewer albumin-treated patients in the overall trial cohort required vasopressor therapy at any time during admission (25.5% vs 32.9%; difference −7.4%, 95% CI −14.5 to −0.0); within the refractory hypotension subgroup, the direction of effect was consistent (26.9% vs 32.8%) though not separately tested. In the same subgroup, systolic blood pressure at 6 hours was higher with albumin (median 108 vs 102 mmHg), although fluid and vasopressor requirements were not significantly different. In the SAFE severe sepsis subgroup, the proportion of patients receiving vasopressors was similar between groups during the first three days (albumin: day 1, 382/597; day 2, 372/558; day 3, 261/466 vs saline: day 1, 402/613; day 2, 395/565; day 3, 275/458).[14] The EARSS abstract reported that the number of days without catecholamines was significantly higher in the albumin group, without providing extractable numeric data. In Cusack et al., vasopressor days were numerically longer in the albumin group (mean 7.87, range 1–39, vs 6.02, range 1–29; p=0.22), opposite in direction to the ALBIOS and EARSS signals, although the trial was small (n=100) and underpowered for this endpoint.

Respiratory outcomes

ARISS reported ventilator-free days until ICU discharge or death: median 4 (IQR 1–7) in the albumin group versus 3 (IQR 0–7) in controls, not statistically significant. Duration of mechanical ventilation was similar between groups in ALBIOS (median 6 days in both arms; p=0.50) and the SAFE severe sepsis subgroup (mean 6.0±7.2 vs 5.4±6.2 days; mean difference −0.56, 95% CI −1.31 to 0.20; p=0.15). In the ICARUS-ED refractory hypotension subgroup, invasive ventilation within 72 hours was required in 19 of 197 albumin-treated patients (9.7%) versus 12 of 189 controls (6.4%). In ARISS, early 6-hour procedure rates were similar between groups, with invasive mechanical ventilation in 64.8% versus 65.1% (albumin

vs control patients). In Cusack et al., mechanical ventilation days were similar (mean 9.9, range 0–48, vs 10.9, range 0–81; p=0.67); all patients were intubated at enrolment by design.

Renal support outcomes

No trial reported renal-replacement-therapy (RRT)-free days as a harmonised endpoint. RRT use was similar between groups in ALBIOS (24.6% vs 21.4%; p=0.11) and the SAFE severe sepsis subgroup (18.7% vs 18.2%; OR 1.04, 95% CI 0.78–1.38; p=0.98).[14] In the ICARUS-ED refractory hypotension subgroup, RRT within 72 hours was required in 6 of 197 albumin-treated patients (3.1%) versus 1 of 189 controls (0.5%), though absolute numbers were small. In Cusack et al., acute kidney injury was equally prevalent (50% in both groups at admission), but RRT use was not separately reported. Rackow et al. did not report organ-support outcomes. In ARISS, early 6-hour renal replacement therapy was 25.6% versus 21.1% (Albumin vs control patients).

Overall trend

These data do not provide evidence of a clear effect of albumin on protocol-defined organ-support-free outcomes. Interpretation is limited by indirect endpoint definitions, non-uniform reporting windows, incomplete subgroup reporting for septic shock, and the fact that several available measures were surrogate rather than protocol-equivalent secondary outcomes. Among the available proxies, the most consistent signal was for earlier vasopressor liberation or reduced vasopressor requirement with albumin-containing resuscitation.

- S4: Exploratory data

Exploratory haemodynamic and resuscitation outcomes showed a more coherent directional signal than the secondary organ-support outcomes, although reporting was heterogeneous and often incomplete. These findings should therefore be interpreted cautiously.

Fluid balance and fluid administration

In ALBIOS, the median cumulative net fluid balance to day 7 was significantly lower in the albumin group than in the crystalloid group (347 mL, IQR −3266 to 4042, vs 1220 mL, IQR −2767 to 5034; p=0.004), despite similar total administered volumes; these data are for the overall severe sepsis population and were not reported for the septic shock subgroup separately. In ICARUS-ED, total fluid administered by 72 hours (including albumin volume) was significantly lower in the albumin arm (median 4275 vs 5000 mL; difference −700 mL, 95% CI −1284 to −116); within the refractory hypotension subgroup, the difference was directionally consistent but smaller (4345 vs 4900 mL). In ARISS, cumulative fluid balance at 24 hours was numerically lower in the albumin group (median 1365 vs 1666 mL). The SAFE severe sepsis subgroup received 43% more study fluid by volume in the saline arm, consistent with the expected oncotic advantage of albumin. Comparable cumulative fluid-balance data were not available for Rackow et al. or EARSS. In Cusack et al., 24-hour fluid balance was similar between groups (mean 2033, range −853 to 8863 mL, vs 2626, range −1462 to 11,294 mL; p=0.40), as was the volume of fluid bolus administered during the acute resuscitation phase (405 vs 481 mL; p=0.15).

Vasopressor exposure

No trial reported norepinephrine-equivalent doses specifically for the septic shock population. The most informative measure was in ALBIOS, where time to vasopressor or inotropic cessation was significantly shorter in the albumin group (median 3 vs 4 days; p=0.007), albeit in the overall severe sepsis cohort and not reported separately for the shock subgroup. In ARISS, early haemodynamics were broadly similar between groups, with mean arterial pressure 75±12 versus 73±11 mmHg and median norepinephrine dose 0.28 versus 0.32 µg/kg/min at 6 hours post-randomisation. In ICARUS-ED, albumin was associated with lower inpatient vasopressor use (25.5% vs 32.9%) and lower vasopressor use at 24 hours (24.7% vs 32.5%), with a consistent direction in the refractory hypotension subgroup. The EARSS abstract reported significantly more catecholamine-free days with albumin but provided no extractable effect estimate. In Cusack et al., baseline noradrenaline dose was numerically lower in the albumin group (0.111 vs 0.168 µg/kg/min; p=0.08), and serial haemodynamic data (MAP, heart rate, lactate, $ScvO_2$) showed no between-group differences at any timepoint from baseline through 24 hours; vasopressor days were numerically longer with albumin (7.87 vs 6.02; p=0.22), although this was underpowered and non-significant.

Serum albumin separation

All four trials reporting serial serum albumin demonstrated significant between-arm separation after randomisation: from day 1 to day 7 in the SAFE severe sepsis subgroup, day 1 to day 28 in ALBIOS, day 0 to day 4 in ICARUS-ED, and from day 1 through ICU stay in ARISS. Cusack et al. reported baseline serum albumin (mean 26.7 vs 27.7 g/L; p=0.42) but did not report post-intervention albumin concentrations. ARISS now adds useful quantitative detail: median serum albumin was similar at randomisation (22.0 vs

22.2 g/L) but diverged to 27.0 versus 21.0 g/L on day 1, 28.0 versus 19.5 g/L on days 2–7, and 29.0 versus 17.9 g/L on days 8–28. In ARISS, however, more than half of albumin-group patients did not achieve the target serum albumin concentration of 30 g/L during their ICU stay, despite a dosing protocol designed to do so; this may reflect the high severity of illness and the dilutional effect of large-volume resuscitation in septic shock.

Control-arm albumin contamination

Control-arm albumin exposure was substantial in the open-label trials. In ALBIOS, 336 of 907 patients (37.1%) in the crystalloid group received albumin at least once during the study period, accounting for 958 of 9697 patient-days (9.9%) of protocol violations. In ARISS, 106 of 218 controls (48.6%) received clinician-directed albumin during their ICU stay, as permitted by current guidelines for refractory shock. Such contamination would be expected to bias treatment effect estimates toward the null and may have attenuated between-group separation within individual trials, for mortality, secondary and exploratory outcomes.

This clinician-directed use is the most relevant contamination metric for interpretation of treatment separation; the protocol-deviation table separately identified study-treatment administration in only 2 control patients (0.9%), which should be interpreted as accidental protocol deviation rather than the full contamination rate. The double-blind design of the SAFE trial minimised contamination by design. Contamination rates were not reported for EARSS, or Cusack et al., and minimal in ICARUS-ED.

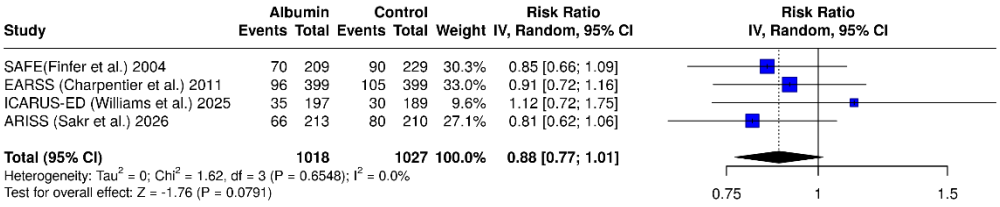
Safety outcomes

Where reported, safety outcomes did not suggest excess harm with albumin-containing resuscitation. In ALBIOS, the incidence of acute kidney injury by RIFLE criteria was similar between groups (21.9% vs 22.7%; p=0.71), as was RRT use (24.6% vs 21.4%; p=0.11). In the SAFE severe sepsis subgroup, RRT use was nearly identical (18.7% vs 18.2%; OR 1.04, 95% CI 0.78–1.38; p=0.98). In ARISS, the overall frequency and severity of adverse events were comparable between groups (267 events in 54.5% of albumin-group patients vs 202 events in 48.1% of controls), with no signal of albumin-related harm. The ARISS supplement further reported severe adverse events in 17.1% versus 18.8% and serious adverse events in 13.5% versus 15.6% of albumin and control patients, respectively; sepsis-related events were broadly similar, with renal events numerically lower in the albumin group. The EARSS abstract reported similar renal and pulmonary tolerance between groups without providing numeric data. In Cusack et al., there were no significant differences between groups in ICU length of stay (mean 20.2, range 2–190, vs 20.9, range 2–122 days; p=0.89), hospital length of stay (48 vs 47.2 days; p=0.93), mechanical ventilation days (9.9 vs 10.9; p=0.67), vasopressor days (7.87 vs 6.02; p=0.22), or ICU survival (35/50 vs 38/50; p=0.49); the trial did not report pulmonary oedema or adverse event frequency. Pulmonary oedema was listed as a monitored possible adverse effect in ARISS but was infrequently and inconsistently reported across the remaining trials, precluding meaningful comparative synthesis; no consistent trial-level pulmonary-harm signal emerged across the included studies. Interpretation of subjective secondary and exploratory outcomes is additionally complicated by the substantial control-arm contamination described above.

- S5: 28-day mortality: detailed results

Four trials (n = 2045; 1018 albumin, 1027 crystalloid) contributed data to the prespecified 28-day mortality sensitivity analysis. The pooled risk ratio was 0.88 (95% CI 0.77–1.01; p = 0.08; I^2 = 0%), suggestive of a mortality benefit consistent in direction and magnitude with the primary analysis.

eFigure S5a: Forest plot — albumin vs crystalloid, 28-day mortality



Footnote: Mortality at 28 days or nearest equivalent. 28-day mortality (SAFE, EARSS, ARISS); 30-day mortality (ICARUS-ED). ALBIOS was excluded from this analysis because the 28-day mortality data were not reported separately for the septic shock subgroup. Of note, Cusack et al. listed 28-day mortality as a secondary endpoint, but event counts were not separately reported. Thus, Cusack et al. and Rackow et al. (study-period mortality) did not have a calendar-fixed endpoint and were excluded from this pooling. ARISS contributes 213 albumin / 210 control to the 28-day analysis, whereas the 90-day primary analysis uses the modified intention-to-treat population (210 / 209); the difference reflects the analysis populations reported by Sakr et al. at each timepoint.

Bayesian analysis: 28-day mortality (refer to Supplement 2 for further details)

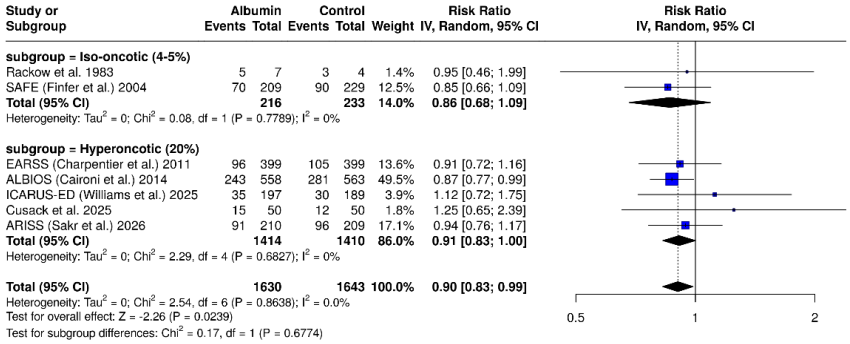
Bayesian analysis of 28-day mortality under the primary weakly informative prior yielded a posterior median RR of 0.89 (95% CrI 0.774--1.024), with P(RR<1) = 94.9% and P(RR<0.95) = 82.2%. These probabilities were consistent across alternative priors: sceptical (P[RR<1] = 94.4%; P[RR<0.95] = 80.4%), optimistic (96.0%; 83.4%), and pessimistic (92.2%; 74.2%). Leave-one-out analyses demonstrated stability: the posterior remained most sensitive to exclusion of ARISS (P[RR<1] dropping to 84.3%) and least affected by exclusion of ICARUS-ED (P[RR<1] = 97.7%). The fixed-effect model yielded stronger posterior probabilities (P[RR<0.95] = 84.6%; P[RR<1] = 95.9%), consistent with the pattern observed in the primary analysis. Complete 28-day Bayesian results are also reported in Supplement 2.

- S6: Pre-specified subgroup analyses

Subgroup analysis: albumin concentration (iso-oncotic vs hyperoncotic)

In trials using iso-oncotic albumin (4–5%)—Rackow et al. and SAFE—the pooled risk ratio for mortality was 0.86 (95% CI 0.68–1.09), with no observed heterogeneity (I^2=0%). In trials using hyperoncotic albumin (20%)—EARSS, ALBIOS, ICARUS-ED, Cusack et al., and ARISS—the pooled risk ratio was 0.91 (95% CI 0.83–1.00), again with no observed heterogeneity (I^2=0%). The formal test for subgroup differences was not significant (χ^2=0.17, df=1, p=0.67), indicating no clear evidence of effect modification by albumin concentration. See eFigure S6a.

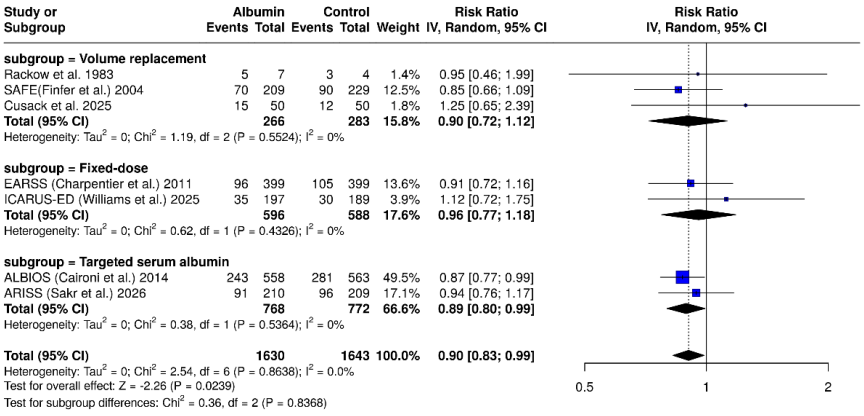
eFigure S6a: Forest plot — effect of different albumin concentrations for primary outcome



Subgroup analysis: dosing strategy

In trials using albumin as volume replacement—Rackow et al., SAFE, and Cusack et al.—the pooled risk ratio for mortality was 0.90 (95% CI 0.72–1.12), with no observed heterogeneity (I^2=0%). In trials using a fixed-dose strategy—EARSS and ICARUS-ED—the pooled risk ratio was 0.96 (95% CI 0.77–1.18), again with no observed heterogeneity (I^2=0%). In trials using a targeted serum albumin strategy—ALBIOS and ARISS—the pooled risk ratio was 0.89 (95% CI 0.80–0.99), with no observed heterogeneity (I^2=0%). The formal test for subgroup differences was not significant (χ^2=0.36, df=2, p=0.84), indicating no clear evidence of effect modification by dosing strategy. See eFigure S6b.

eFigure S6b: Forest plot — effect of different dosing strategies for primary outcome



Subgroup analysis: baseline hypoalbuminaemia

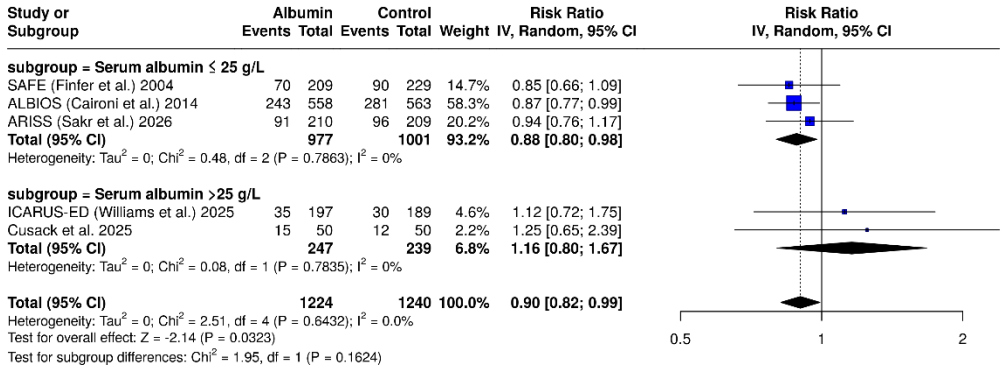
Baseline albumin subgrouping was performed at the study level and should not be interpreted as an individual-patient hypoalbuminaemia analysis. When septic shock-specific baseline albumin values were unavailable, the closest reported baseline population was used, as detailed in eTable S6.X. When baseline albumin was reported separately by treatment arm, classification was based on values across both randomised groups; when only a single overall baseline value was reported, that value was used. Accordingly, this subgroup analysis evaluates whether trial-level baseline albumin category modified the treatment effect, but cannot determine whether individual patients with hypoalbuminaemia derive greater benefit.

eTable S6.1. Source of baseline albumin values used for trial-level subgroup classification

Trial	Population source for baseline albumin	Arm source	Value used	Trial-level category
SAFE	Severe sepsis cohort; not separately reported for septic shock subgroup	By treatment arm	25.0 and 25.2 g/L	≤25 g/L
ALBIOS	Overall severe sepsis trial population; not separately reported for septic shock subgroup	By treatment arm	24.1 and 24.2 g/L	≤25 g/L
ARISS	Enrolled septic shock trial population	By treatment arm	22.0 and 22.2 g/L	≤25 g/L
ICARUS-ED	Overall trial population; not separately reported for refractory-hypotension subgroup	Overall value	31 g/L	>25 g/L
Cusack et al.	Enrolled septic shock population / analysed cohort	By treatment arm	26.7 and 27.7 g/L	>25 g/L
Rackow et al.	Not available / not used	Not applicable	Not available	Not included
EARSS	Not available / not used	Not applicable	Not available	Not included

Using a study-level threshold of baseline serum albumin ≤25 versus >25 g/L, ARISS, ALBIOS, and SAFE were classified in the ≤25 g/L category, whereas Cusack et al. and ICARUS-ED were classified in the >25 g/L category. In the ≤25 g/L subgroup, albumin was associated with lower mortality (RR 0.88, 95% CI 0.80 to 0.98; I^2 = 0%; 3 studies), whereas no clear benefit was observed in the >25 g/L subgroup (RR 1.16, 95% CI 0.80 to 1.67; I^2 = 0%; 2 studies) (eFigure S6c). However, the test for subgroup differences was not statistically significant (χ^2 = 1.95, df = 1, p = 0.16). The point estimates were therefore more favourable in trials with lower baseline albumin levels, but this study-level analysis does not provide clear evidence of effect modification and cannot identify individual-level treatment responsiveness.

eFigure S6c: Forest plot — effect of different baseline serum albumin levels for primary outcome



ICEMAN assessment

The Instrument to Assess the Credibility of Effect Modification Analyses (ICEMAN) was applied to all three subgroup comparisons. No subgroup comparison was judged as providing credible evidence of effect modification. Although trial-level albumin category was prespecified as the primary effect modifier and baseline hypoalbuminaemia also had biological plausibility, interaction tests were non-significant, subgroup-specific estimates were imprecise, and the number of contributing studies within each subgroup was small. In addition, the baseline hypoalbuminaemia analysis relied on study-level classification and, for several trials, on broader mixed-severity cohorts rather than septic shock-specific baseline albumin data. Overall, the ICEMAN assessment supports low credibility of effect modification by concentration, dosing strategy, or baseline hypoalbuminaemia within the current dataset.

- S7: Sensitivity analyses

Fixed-effect model

As expected, given the low heterogeneity in the random-effects model, the analysis using the fixed-effects model with inverse-variance weights yielded the same results (RR 0.90, 95% CI 0.83–0.99; p = 0.02). The near-identical pooled estimates under both models confirm that the primary finding is not sensitive to the choice of meta-analytic model.

High-risk-of-bias exclusion

The exclusion of trials deemed high-risk per RoB 2.0 assessment (Rackow et al. and EARSS) did not impact the pooled estimate. The random-effects model with inverse-variance weighting yielded an RR of 0.90 (95% CI, 0.82–0.99; p = 0.03). This confirms that the primary finding is not driven by the two methodologically weakest trials in the evidence base.

Bayesian leave-one-out sensitivity analyses

Leave-one-out Bayesian sensitivity analyses confirmed the robustness of the primary posterior estimates for longest follow-up mortality. Sequential exclusion of each trial yielded posterior median RRs ranging from 0.909 (excluding ICARUS-ED) to 0.953 (excluding ALBIOS), with P(RR<1) remaining above 77% in all refits. The posterior was most attenuated by removal of ALBIOS (P[RR<1] = 77.0%; P[RR<0.95] = 47.9%), the largest single trial contributing the most events, and was most robust to exclusion of ICARUS-ED (P[RR<1] = 96.1%; P[RR<0.95] = 80.3%) and Cusack et al. (P[RR<1] = 96.8%; P[RR<0.95] = 79.4%). No single trial exclusion shifted the posterior median RR above 1.0, indicating that the overall finding is not driven by any individual study. Detailed leave-one-out results for both endpoints are also presented in Supplement 2.

- ## S8: GRADE assessment details

High baseline certainty (randomised clinical trials). The evidence was not downgraded for risk of bias. Although two trials were judged at high risk of bias (Rackow et al. and EARSS), sensitivity analysis excluding these trials did not change the pooled estimate (RR 0.90, 95% CI 0.82–0.99). The evidence was not downgraded for inconsistency ($I^2 = 0\%$; $Q = 1.56$, $p = 0.91$). The evidence was downgraded one level for indirectness because a substantial proportion of the pooled data came from post hoc subgroup analyses or subgroup extractions rather than trials designed with septic shock as the primary population and mortality as the primary endpoint. The evidence was not downgraded for imprecision: the confidence interval excluded the null and the optimal information size was met at the prespecified 15% RRR threshold. Publication bias not downgraded. Funnel plot assessment was qualitative because fewer than 10 studies were available; visual inspection was largely equivocal and did not show clear or convincing asymmetry, and trial registry searches did not identify unreported completed trials.