Journal Club: ICRAKI (Gaudry et al., NEJM 2026)

Strukturierte Bewertung · Andreas Fazekas · luftikus.org · persönliche Lehrunterlage, keine Stellungnahme des Arbeitgebers · Version 1.2 · 03.11.2026 (öffentliche Fassung) · Quellenstand: Volltext und Research Summary (NEJM, 14.10.2026); Supplement nicht eingesehen

Kernaussage

- Initiale intermittierende Hämodialyse (IHD) war gegenüber kontinuierlicher Nierenersatztherapie (CRRT) beim primären Endpunkt MAKE90 nicht nichtunterlegen. Die Überlegenheit der CRRT war ebenfalls nicht signifikant: RR 1,09 (95 %-KI 0,98–1,20).
- Die 90-Tage-Mortalität war unter IHD höher: HR 1,30 (1,04–1,63). Dieser Endpunkt ist sekundär, nicht multiplizitätsadjustiert und stammt aus einer vorzeitig abgebrochenen Studie. Die Richtung ist belastbarer als die Größe.
- Die Vertrauenswürdigkeit der Mortalitätsaussage ist nach Eigenbewertung niedrig (GRADE; moderat, wenn der Frühabbruch nur einmal abgewertet wird). Gründe: offenes Design, Ausschlüsse nach Randomisierung, Frühabbruch ohne Abbruchregeln, Fragility Index 3.
- Für ICUs mit kontinuierlicher HDF als Standard folgt: Die initiale kontinuierliche HDF bei beatmeten oder katecholaminpflichtigen Patient*innen wird gestützt. Ab Tag 4 bleibt der Wechsel auf IHD eine offene Frage.

1 Systeme zur strukturierten Bewertung eines RCT

Die Instrumente beantworten unterschiedliche Fragen. Eine Checkliste für die Berichtsqualität misst nicht die Validität. Ein Risk-of-Bias-Urteil ersetzt keine Aussage zur Anwendbarkeit.

Instrument	Frage	Inhalt und Einsatz im Journal Club	Ref.
CONSORT 2025	Wurde vollständig berichtet?	Checkliste und Flussdiagramm, neuer Abschnitt Open Science. Prüft Berichtsqualität, nicht Bias. Gut als Leseleitfaden.	[2]
CONSORT-Erweiterung Nichtunterlegenheit	Sind Margin, Analysepopulation und Hypothese korrekt berichtet?	Pflichtmodul bei NI- und Äquivalenzstudien: Begründung der Margin, ITT- und Per-Protocol-Analyse, KI gegen Margin.	[3]
JAMA Users' Guides / CASP RCT	Valide? Welche Ergebnisse? Lokal anwendbar?	Didaktisches Grundgerüst in drei bzw. vier Abschnitten. CASP ist bewusst ohne Punktescore. Geeignet als Struktur des Vortrags.	[4, 5]
Cochrane RoB 2	Wie hoch ist das Verzerrungsrisiko pro Endpunkt?	Fünf Domänen mit Signalfragen und Algorithmus: Randomisierung, Abweichungen von der Intervention, fehlende Daten, Messung, selektives Berichten. Urteil getrennt je Endpunkt. Referenzstandard.	[6]
JBI-Tool RCT (2023)	Wie hoch ist das Verzerrungsrisiko?	Alternative zu RoB 2, ebenfalls nach Bias-Domänen geordnet. Einfacher, für Anfänger*innen geeignet.	[7]
Designmodule	Was verzerrt speziell dieses Design?	Nichtunterlegenheit [8], Frühabbruch [9, 10], kombinierte Endpunkte [11], Fragility Index [12], Bayes-Reanalyse [13].	[8–13]
PRECIS-2	Wie pragmatisch ist die Studie?	Neun Domänen von explanatorisch bis pragmatisch, zum Beispiel Einschluss, Setting, Flexibilität der Intervention. Strukturiert die Übertragbarkeit auf die eigene ICU.	[14]

GRADE	Wie sicher ist die Evidenz insgesamt?	Abwertung bei Bias, Inkonsistenz, Indirektheit, Unpräzision und Publikationsbias. Ergebnis: hoch, moderat, niedrig oder sehr niedrig. Bewertet den Endpunkt über alle Studien, nicht die einzelne Studie.	[15]

Empfohlener Pfad für den Journal Club (im Template umgesetzt): PICO → Design-Steckbrief → CASP-Screening → RoB 2 für die zwei wichtigsten Endpunkte → passende Designmodule → Effektgrößen mit KI, NNT/NNH → Kontext der Vorevidenz → Anwendbarkeit (PRECIS-2, lokal) → GRADE → Fazit und Praxisentscheidung.

2 ICRAKI im Überblick

P	Erwachsene auf der ICU mit AKI und Indikation zur Nierenersatztherapie (RRT), invasiv beatmet oder katecholaminpflichtig. Ausgeschlossen waren unter anderem Hirnschädigung mit erhöhtem Hirndruck und dialysierbare Toxine. Mittlerer SOFA-Score 12, septischer Schock 42,5–46,7 %.
I / C	Initiale IHD (rein diffusiv) gegenüber CRRT (CVVH, CVVHD oder CVVHDF nach ärztlicher Wahl). Die Zuteilung galt mindestens bis Tag 3, danach war ein Wechsel erlaubt.
O	Primär MAKE90: Tod, RRT-Abhängigkeit oder Kreatinin mehr als 25 % über Ausgangswert. Sekundär unter anderem Mortalität an Tag 28, 60 und 90.
Design	Offen, randomisiert, Nichtunterlegenheit mit Margin 7,5 Prozentpunkten. 42 französische ICUs, davon 38 rekrutierend. Öffentlich finanziert. Geplant waren n = 1.000, randomisiert wurden 554, ausgewertet 535 (mITT).
Abbruch	Das DSMB empfahl den Abbruch nach drei Zwischenanalysen der 28-Tage-Mortalität. Geplant waren zwei, ohne vordefinierte Entscheidungsregeln.

3 Bewertung entlang des Pfads

3.1 Verzerrungsrisiko nach RoB 2 (Eigenbewertung)

Domäne	Mortalität d90	MAKE90	Begründung
D1 Randomisierung	**gering**	**gering**	Zentral, webbasiert, stratifiziert, Blockgröße verdeckt. Baseline-Imbalanzen zugunsten der CRRT-Gruppe: CKD 21,9 vs. 15,4 %, Notfall-OP 18,1 vs. 9,5 %, Katecholamine bei Aufnahme 85,4 vs. 77,8 %. Am ehesten Zufall. Die post hoc adjustierte Analyse soll konsistent sein (Tab. S7, nicht eingesehen).
D2 Abweichungen von der Intervention	**einige Bedenken**	**einige Bedenken**	Offenes Design. 19 Randomisierte wurden ausgeschlossen (12 IHD, 7 CRRT): ohne erhaltene RRT, Ablehnung der Datennutzung durch Angehörige oder rechtliche Gründe. Im Artikel heißt diese Population ITT; tatsächlich ist es eine mITT. Wechsel der Methode: 21,9 % (IHD) vs. 42,5 % (CRRT), davon 10 bzw. 9 protokollwidrig vor Tag 4. Der Behandlungswechsel verdünnt den Effekt Richtung Null.
D3 Fehlende Daten	**gering**	**einige Bedenken**	Vitalstatus zu 99,3 % vollständig. Kreatinin fehlte bei 20 von 213 Überlebenden ohne RRT und wurde multipel imputiert.

D4 Messung des Endpunkts	gering	einige Bedenken	Tod ist objektiv. Über RRT-Abhängigkeit entschieden unverblindete Behandler*innen; Kreatinin ist bei kritisch Kranken artefaktanfällig.
D5 Selektives Berichten	gering	gering	Protokoll und Analyseplan wurden vorab publiziert, die Studie ist registriert (NCT06032884); die Ergebnisse werden gemäß Analyseplan berichtet. Die zusätzliche Zwischenanalyse (drei statt zwei, ohne Abbruchregel) ist eine Abweichung in der Durchführung, kein selektives Berichten; sie wird unter Frühabbruch (3.2) bewertet.
Gesamt	einige Bedenken	einige Bedenken	

3.2 Designmodule

Nichtunterlegenheit. Die Margin von 7,5 Prozentpunkten entspricht etwa 10 % relativ [1]. MAKE90 wird von der Mortalität dominiert. Eine Margin dieser Größe akzeptiert damit implizit bis zu 7,5 Prozentpunkte zusätzliche Todesfälle. Das ist ein Diskussionspunkt, ebenso die Mortalitätsmargin von HR 1,27. Behandlungswechsel gleichen die Gruppen einander an; die ITT-Analyse ist bei NI-Studien deshalb nicht konservativ, sondern begünstigt Nichtunterlegenheit [3, 8]. Dass die Nichtunterlegenheit trotzdem verfehlt wurde, stärkt die Richtung des Befunds.

Frühabbruch. Abgebrochen wurde wegen Schadens bei einem sekundären Endpunkt, ohne formale Abbruchgrenzen. Abgebrochene Studien überschätzen Effekte systematisch; das Verhältnis der relativen Risiken gegenüber nicht abgebrochenen Studien betrug insgesamt 0,71 (0,65–0,77), die größten Überschätzungen traten bei weniger als 500 Ereignissen auf [10]. ICRAKI hat 311 Todesfälle. Diese Daten stammen aus Studien mit Abbruch wegen Nutzens, die Übertragung auf einen Abbruch wegen Schadens ist eine Analogie. Die Autor*innen benennen die mögliche Überschätzung selbst [1].

Kombinierter Endpunkt. MAKE90 wird von Todesfällen getragen. Unter Überlebenden war die Rate der Nierenkomponenten gleich: 39/96 (40,6 %) vs. 50/128 (39,1 %; eigene Berechnung aus Tab. 2). Weil Tod ein konkurrierendes Ereignis ist, lässt sich zu Nierenfolgen keine Aussage machen [1, 11].

Plausibilität. Die Mortalitätskurven trennen sich früh (Tag 28: 53,6 vs. 44,8 %). Post hoc traten unter IHD häufiger Katecholamin-Eskalationen auf. Das ist hypothesengenerierend und passt zu hämodynamischem Stress [1]. Hypotonie war in beiden Gruppen gleich häufig (68,8 vs. 66,5 %). Die Durchführung der IHD bleibt im Haupttext unklar: Dosis unter dem Ziel, keine Angaben auf Sitzungsebene zu Dialysattemperatur, Natrium oder Sitzungsdauer.

3.3 Effektgrößen (Rohdaten, eigene Berechnung)

Endpunkt	IHD	CRRT	Absolute Differenz (95 %-KI)	Relativ (95 %-KI)	NNH
MAKE90 (primär)	203/260 (78,1 %)	197/275 (71,6 %)	+6,28 PP (−1,22 bis 13,79), adjustiert [1]	RR 1,09 (0,98–1,20)	–
Tod bis Tag 90	164/260 (63,1 %)	147/275 (53,5 %)	+9,6 PP (1,3 bis 17,8), roh, Newcombe	HR 1,30 (1,04–1,63) [1]	10 (6–79)

Rechenweg: absolute Risikodifferenz (ARD) = 164/260 − 147/275 = 0,631 − 0,535 = 0,096. NNH = 1/ARD = 10,4. Die KI-Grenzen der NNH sind die Kehrwerte der KI-Grenzen der ARD: 1/0,178 = 5,6 und 1/0,013 = 79. Die Werte sind unadjustiert und durch den Frühabbruch vermutlich überschätzt. Fragility Index der Mortalität: 3. So viele zusätzliche Todesfälle im CRRT-Arm heben p unter 0,05 auf (Fisher-Test; roh p = 0,028) [12]. Approximative Bayes-Interpretation (Normalapproximation auf log-HR, SE 0,115): Mit flachem Prior beträgt P(HR > 1) 0,99 und P(HR > 1,1) 0,93. Mit skeptischem Prior (95 % der Masse zwischen HR 0,80 und 1,25) ergibt sich eine Posterior-HR von 1,14 (0,97–1,33) mit P(HR > 1) 0,95. Das ist eine Eigenrechnung zur Illustration, keine formale Reanalyse [13].

3.4 Kontext: Vorevidenz

Studie	Ergebnis

Mehta 2001, RCT, n = 166 [17]	Höhere Mortalität unter CRRT. Starke Baseline-Imbalanz zugunsten IHD; adjustiert OR 1,3 (0,6–2,7), nicht signifikant.
Vinsonneau 2006, RCT, n = 360 [16]	60-Tage-Überleben 32 vs. 33 %, kein Unterschied. IHD wurde nach strengen Verträglichkeitsprotokollen durchgeführt.
Gaudry 2022, Sekundäranalyse AKIKI/IDEAL-ICU, n = 543 [18]	CRRT war mit höherer Mortalität assoziiert (gewichtete HR 1,26; 1,01–1,60). Das widerspricht ICRAKI; dieselbe Arbeitsgruppe hat ICRAKI als Konsequenz geplant.
Wald 2023, Sekundäranalyse STARRT-AKI, n = 2.196 [19]	CRRT war mit weniger Tod oder RRT-Abhängigkeit assoziiert (OR 0,81; 0,66–0,99), getragen von der RRT-Abhängigkeit.
Leitlinien	KDIGO 2012, 5.6.2: CRRT bei hämodynamischer Instabilität (2B) [20]. S3-Leitlinie 2025/2026, 3.3: kontinuierliche Verfahren bei Instabilität (B) [21].

Die Vorevidenz ist inkonsistent. ICRAKI ist der erste große RCT mit indikationsbasiertem Start und widerspricht der älteren Annahme der Gleichwertigkeit.

3.5 Anwendbarkeit (PRECIS-2-orientiert)

- **Population:** schwer krank (SOFA 12, 84 % beatmet, Mortalität über 50 %). Bei stabileren Patient*innen ist der Befund indirekt, obwohl die Subgruppen keine Heterogenität zeigen.
- **Setting:** französische ICUs mit hoher IHD-Expertise. Das stärkt den Befund, weil eine schlechte IHD-Praxis als Erklärung unwahrscheinlich ist.
- **Intervention:** Die CRRT-Modalität war frei wählbar. Antikoagulation und Dosis der CRRT sind im Haupttext nicht berichtet. Für Citrat-CVVHDF ist die Übertragung plausibel, aber nicht gesondert geprüft.
- **Nicht beantwortet:** Methode ab Tag 4, SLED/PIRRT, Patient*innen mit Hirndruck oder Intoxikation.

3.6 GRADE für die Mortalität (Eigenbewertung, Einzelstudie)

Ausgangsniveau hoch (RCT). −1 für Risk of Bias: offen, mITT, Frühabbruch ohne Regeln. −1 für Unpräzision: Frühabbruch, KI von 1,04 bis 1,63, Fragility Index 3. Inkonsistenz zu den Vorstudien nur bei Beobachtungsdaten und Altstudien, daher keine weitere Abwertung. **Ergebnis: niedrig.** Wer den Frühabbruch nur einmal (unter Unpräzision) zählt, kommt auf moderat; diese Grenzentscheidung eignet sich für die Diskussion.

4 Implikationen und Diskussionsfragen

Praxis: keine Änderung. Die initiale kontinuierliche HDF bei instabilen, beatmeten Patient*innen entspricht der Studienlage und den Leitlinien. Wird ab Tag 4 bei hämodynamischer Stabilität und Mobilisierungsziel auf IHD gewechselt, ist das weiterhin eine Einzelfallentscheidung und sollte dokumentiert werden.

1. Ist eine NI-Margin von 7,5 Prozentpunkten bei einem mortalitätsdominierten Endpunkt ethisch und klinisch vertretbar?
2. Wie viel des Mortalitätseffekts ist Frühabbruch-Artefakt? Welche Effektgröße erwarten wir realistisch? Vergleich mit dem skeptischen Prior.
3. Warum widerspricht ICRAKI der eigenen Sekundäranalyse derselben Arbeitsgruppe von 2022 (Confounding by Indication)?
4. Welche IHD-Verträglichkeitsmaßnahmen (Vinsonneau 2006) wären bei uns Standard, und fehlten sie in ICRAKI?
5. Was messen wir lokal? Zum Beispiel Katecholamin-Eskalation beim Start und Filterlaufzeit.

Hinweis: Diese Bewertung und die abgeleiteten Zahlen (NNH, Fragility Index, Bayes-Approximation) ersetzen keine individuelle klinische Beurteilung. RoB-2- und GRADE-Urteile sind Eigenbewertungen auf Basis des Volltexts ohne Supplement.

Daneben aufgefallen

- Geplant waren zwei Zwischenanalysen (nach 300 und 600 Patient*innen), durchgeführt wurden drei. Ein gutes Beispiel für eine Abweichung zwischen Protokoll und Durchführung, die nach CONSORT berichtet werden muss.
- Rund 45 % der Studienpopulation hatten einen septischen Schock. Mit welcher Antikoagulation die CRRT lief, berichtet der Haupttext nicht; für Citrat gelten im Schock die Einschränkungen der jeweiligen Fachinformation.
- Das Supplement (Tab. S3/S4: Dosis, Modalität, Antikoagulation; S7: adjustierte Analyse) sollte vor dem Journal Club eingesehen werden.

Referenzen

1. Gaudry S, Boubaya M, Louis G, et al. Intermittent or continuous renal-replacement therapy in acute kidney injury. N Engl J Med. 2026;395(17):1677-90. doi:10.1056/NEJMoa2608397
2. Hopewell S, Chan AW, Collins GS, et al. CONSORT 2025 statement: updated guideline for reporting randomised trials. BMJ. 2025;389:e081123. doi:10.1136/bmj-2024-081123
3. Piaggio G, Elbourne DR, Pocock SJ, et al. Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. JAMA. 2012;308(24):2594-604. doi:10.1001/jama.2012.87802
4. Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. A. Are the results of the study valid? JAMA. 1993;270(21):2598-601. doi:10.1001/jama.270.21.2598
5. Critical Appraisal Skills Programme. CASP Checklist: Randomised Controlled Trials. 2024. casp-uk.net (PDF)
6. 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 · Werkzeug: riskofbias.info
7. Barker TH, Stone JC, Sears K, et al. The revised JBI critical appraisal tool for the assessment of risk of bias for randomized controlled trials. JBI Evid Synth. 2023;21(3):494-506. doi:10.11124/JBIES-22-00430
8. Mauri L, D'Agostino RB. Challenges in the design and interpretation of noninferiority trials. N Engl J Med. 2017;377(14):1357-67. doi:10.1056/NEJMra1510063
9. Montori VM, Devereaux PJ, Adhikari NK, et al. Randomized trials stopped early for benefit: a systematic review. JAMA. 2005;294(17):2203-9. doi:10.1001/jama.294.17.2203
10. Bassler D, Briel M, Montori VM, et al. Stopping randomized trials early for benefit and estimation of treatment effects: systematic review and meta-regression analysis. JAMA. 2010;303(12):1180-7. doi:10.1001/jama.2010.310
11. Ferreira-González I, Busse JW, Heels-Ansdell D, et al. Problems with use of composite end points in cardiovascular trials: systematic review of randomised controlled trials. BMJ. 2007;334(7597):786. doi:10.1136/bmj.39136.682083.AE
12. Walsh M, Srinathan SK, McAuley DF, et al. The statistical significance of randomized controlled trial results is frequently fragile: a case for a Fragility Index. J Clin Epidemiol. 2014;67(6):622-8. doi:10.1016/j.jclinepi.2013.10.019
13. Goligher EC, Tomlinson G, Hajage D, et al. Extracorporeal membrane oxygenation for severe acute respiratory distress syndrome and posterior probability of mortality benefit in a post hoc Bayesian analysis of a randomized clinical trial. JAMA. 2018;320(21):2251-9. doi:10.1001/jama.2018.14276
14. Loudon K, Treweek S, Sullivan F, et al. The PRECIS-2 tool: designing trials that are fit for purpose. BMJ. 2015;350:h2147. doi:10.1136/bmj.h2147
15. 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
16. Vinsonneau C, Camus C, Combes A, et al. Continuous venovenous haemodiafiltration versus intermittent haemodialysis for acute renal failure in patients with multiple-organ dysfunction syndrome: a multicentre randomised trial. Lancet. 2006;368(9533):379-85. doi:10.1016/S0140-6736(06)69111-3
17. Mehta RL, McDonald B, Gabbai FB, et al. A randomized clinical trial of continuous versus intermittent dialysis for acute renal failure. Kidney Int. 2001;60(3):1154-63. doi:10.1046/j.1523-1755.2001.0600031154.x
18. Gaudry S, Grolleau F, Barbar S, et al. Continuous renal replacement therapy versus intermittent hemodialysis as first modality for renal replacement therapy in severe acute kidney injury: a secondary analysis of AKIKI and IDEAL-ICU

studies. Crit Care. 2022;26(1):93. doi:10.1186/s13054-022-03955-9

19. Wald R, Gaudry S, da Costa BR, et al. Initiation of continuous renal replacement therapy versus intermittent hemodialysis in critically ill patients with severe acute kidney injury: a secondary analysis of STARRT-AKI trial. Intensive Care Med. 2023;49(11):1305-16. doi:10.1007/s00134-023-07211-8
20. Kidney Disease: Improving Global Outcomes (KDIGO) Acute Kidney Injury Work Group. KDIGO clinical practice guideline for acute kidney injury. Kidney Int Suppl. 2012;2:1-138. doi:10.1038/kisup.2012.1
21. Meersch-Dini M, et al. Multidisciplinary guidelines on renal replacement therapy in intensive care medicine. Crit Care. 2026;30:46. doi:10.1186/s13054-025-05817-6

Verifikation: [1] aus dem Volltext; [2–4, 6–19] über PubMed (03.11.2026); [5] über die Website von CASP; [20, 21] im Volltext verifiziert (Recherche 02.11.2026).