Cureus
Part of SPRINGER NATURE

Open Access Review Article

Digital Twins and Offline Reinforcement Learning for Mechanical Circulatory Support Decisions: What the Current Evidence Can and Cannot Support

Review began 09/11/2026
Review ended 09/22/2026
Published 09/30/2026

© **Copyright** 2026
Li et al. This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

DOI: 10.7759/cureus.117234

Oscar Mingtian Li [1], Yijiang Li [1], Summer Xia Tian Li [1]

1. Zweistein Innovation Lab, Active Care and Education in Sports and Outreach (ACESO) Foundation, Dublin, USA

Corresponding author: Oscar Mingtian Li, oscarmtli113@gmail.com

Abstract

Mechanical circulatory support (MCS) is not a single intervention but a sequence of decisions: whether to initiate support, which device to use, how to titrate it while other therapies change, when to test for recovery, and whether to wean, escalate, bridge, or withdraw. Each decision has its own eligible population, time origin, action set, and cost of error. Interest is growing in patient-specific digital twins and offline reinforcement learning to support these decisions, and the first device-specific probabilistic twins and safety-regularized weaning policies have now been published. This narrative review examines what that literature currently establishes. Cardiac digital twins have produced their most convincing results in electrophysiology and resynchronization, where the modeled substrate is anatomically stable, and the intervention is discrete; circulatory support is harder because it involves a continuously titrated device acting on a rapidly changing patient under unrecorded co-interventions. The published MCS work is small, largely single-device, restricted to patients already receiving support, and evaluated against hemodynamic surrogate rewards inside simulators estimated from the same trajectories used for training. Three problems recur and appear to be the limiting factors rather than model capacity: predictive accuracy for an endpoint does not establish that an action is beneficial; a simulator cannot independently certify a policy optimized against it; and information available to a retrospective analyst is not information that was available at the decision time. The defensible near-term claim is improved state estimation and forecasting with calibrated uncertainty and explicit abstention. The initiation question requires target-trial emulation rather than a post-implant cohort, and public critical-care datasets can validate chronology and transport but cannot validate device titration or weaning policy.

Categories: Cardiology, Healthcare Technology, Cardiac/Thoracic/Vascular Surgery
Keywords: cardiogenic shock, clinical decision support, digital twin, extracorporeal membrane oxygenation, heart failure, mechanical circulatory support, myocardial recovery, offline reinforcement learning, off-policy evaluation, ventricular assist device

Introduction And Background

Mechanical circulatory support (MCS) is usually described through its endpoints: survival, myocardial recovery, transplantation, explantation, or successful weaning. The decisions that produce those endpoints are trajectories rather than single choices. A team recognizes a syndrome and grades its severity, decides whether support is indicated at all, selects and initiates a device, adjusts the level of support while inotropes, vasopressors, ventilation, volume, and anticoagulation all change around it, tests whether the native heart is recovering, and then weans, escalates, bridges to a durable device or transplant, or moves to comfort-focused care. Contemporary guidelines and scientific statements treat these as related but distinct decisions, with different eligibility, different evidence, and different consequences of error [1-7].

This structure matters for any computational system intended to help. A model that accurately predicts a favorable terminal outcome can still be clinically unhelpful or harmful if it uses information that was not available when the decision had to be made, if it responds after the window in which action would have mattered, or if it recommends an action that was rarely or never taken in comparable patients and therefore has no observational support. Discrimination for an endpoint and usefulness at a decision node are different properties, and one does not imply the other.

Two computational approaches have attracted particular attention. The first is the patient-specific digital twin: a model that maintains a latent physiological state, updates it as new observations arrive, forecasts how that state will evolve under candidate actions, and reports uncertainty [8,9]. The second is offline reinforcement learning, which attempts to learn a treatment policy from historical trajectories without experimenting on patients [10-12]. The two have recently converged on circulatory support: probabilistic twins of pump-level physiology have been proposed for smart weaning, and safety-regularized offline policies have been trained and evaluated against those twins [13-15].

That convergence is a reasonable place to ask what has actually been established. The purpose of this review is not to catalog architectures. It is to separate what the current literature supports from what it is often read as supporting, and to state which evidence would be needed to move each claim up one level. We take the position that the binding constraints on this field at present are not model capacity or data volume but three specific evidentiary problems: the gap between prediction and action benefit, the circularity of certifying a policy against a simulator built from the same data, and the difference between information available retrospectively and information available at the decision time.

Review

Review methodology and scope

This is a narrative review, not a systematic review, and it is reported with reference to the Scale for the Assessment of Narrative Review Articles (SANRA) criteria for narrative review quality rather than the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [16]. No protocol was registered, and no quantitative synthesis was performed because the included literature spans clinical trials, prediction-model studies, machine-learning methods papers, and reporting guidance, which share no common effect measure.

The reference base was assembled in two stages. An initial corpus was compiled by the authors during earlier work on this topic. It was then supplemented by targeted searching of PubMed-indexed literature and by general web search, both performed on September 8 and 9, 2026, with citation chaining from the retrieved articles. Searches were run by a single reviewer without blinding. Because a substantial part of the digital twin and offline reinforcement learning literature appears as conference proceedings, workshop papers, and preprints that are not indexed in PubMed, general web search rather than PubMed was the route by which that part of the corpus was identified.

Searches were executed as free-text Boolean queries run iteratively rather than as a single locked strategy. The queries were combinations of three term blocks. Block A, support modality: "mechanical circulatory support" OR "ventricular assist device" OR "extracorporeal membrane oxygenation" OR "Impella" OR "micro-axial flow pump" OR "cardiogenic shock." Block B, computational approach: "digital twin" OR "reinforcement learning" OR "offline reinforcement learning" OR "off-policy evaluation" OR "machine learning" OR "prediction model." Block C, decision node: "weaning" OR "explantation" OR "myocardial recovery" OR "ramp test" OR "titration" OR "candidacy" OR "initiation." The principal queries were (A) AND (B), (A) AND (C), and (A) AND (B) AND (C). Trials, reporting guidance, and public critical-care datasets named within the retrieved articles were then retrieved individually by name.

Selection was purposive rather than protocol-driven. Articles were included when they defined a decision node, reported a clinical trial or registry result bearing on one, introduced a modeling method used in this domain, or set out reporting or regulatory expectations. Non-English articles, conference abstracts without an accessible full record, and articles addressing respiratory support only were not sought. Screening was performed in a single pass by one reviewer: titles and abstracts of retrieved records were read, candidates were retrieved in full text or, for conference papers, as the full proceedings record, and inclusion was decided against those criteria. There was no second screener, no adjudication procedure, and no risk-of-bias instrument.

Because the search was iterative and purposive rather than executed once against a locked protocol, no screening log was kept, and the numbers of records identified and screened at each stage were not recorded prospectively. They are therefore not reported here, since any figure reconstructed after the fact would not describe what actually occurred. What can be stated exactly is the endpoint: 73 references are cited, of which three are the support-specific digital twin, simulator, or learned-policy studies whose design characteristics are set out in Table *1*. Two further computational studies on adjacent questions are discussed in the text but are not included in Table *1*: a registry prediction model of myocardial recovery on durable support, which estimates risk without simulating physiology or learning a policy, and an offline reinforcement learning study in venovenous extracorporeal support for respiratory failure, which concerns respiratory rather than circulatory support.

Study	Population and device	Sample size	Decision or quantity modeled	Outcome or reward	Validation approach as reported	Principal limitation for decision-support claims
Sun et al., 2024 [14]	Patients undergoing high-risk percutaneous coronary intervention supported with a percutaneous micro-axial flow pump	211 patients; 43,814 thirty-minute time series	Forward simulator of the aortic pressure response to pump support level; no policy is learned	Predicted mean aortic pressure with calibrated uncertainty	Each training batch split into context and target points; no patient-level held-out cohort and no external dataset reported	This is a forecasting model only. The authors state that further clinical metrics are required before generalizability can be assessed.
Sun et al., 2025 [13]	Cardiogenic shock patients supported with percutaneous micro-axial flow pumps	Not reported in the workshop paper	Probabilistic Transformer digital twin of circulatory dynamics, then used as the environment in which offline reinforcement learning methods are benchmarked.	Forecast error against baseline dynamics models; offline reinforcement learning benchmarked on clinically framed metrics	Twin compared against baseline models; candidate policies evaluated inside the twin.	Policy comparison occurs entirely within a learned simulator. Because this is a workshop paper, the depth of peer review is limited.
Tumay et al., 2026 [15]	Cardiogenic shock patients supported with percutaneous micro-axial flow pumps	379 patients, mean record 65.5 hours, 12 device-derived features	Weaning: choice of discrete pump support level	Composite of a hemodynamic reward, a penalty on abrupt action changes, and a weaning score conditioned on hemodynamic stability	Patients split 65:15:20 into training, validation, and testing sets for the dynamics model; policy value then estimated by model-based off-policy evaluation inside that learned twin, with additional synthetic trajectories rolled out from it.	Reward components are author-designed proxies that the authors state await clinician review. The policy is never valued outside a model fitted to its own training distribution.

TABLE 1: Design characteristics and principal limitations of the support-specific digital twin, simulator, and learned-policy studies in mechanical circulatory support.

Studies are those cited in this review that develop or evaluate a digital twin, simulator, or learned treatment policy using data from patients on mechanical circulatory support. A registry prediction model of myocardial recovery and an offline reinforcement learning study in venovenous extracorporeal support for respiratory failure are discussed in the text but are not included because the first is a prediction model and the second concerns respiratory rather than circulatory support. Entries record what each study reports; fields not reported are marked as such rather than inferred. Sources: [13-15].

Every reference in this article was individually checked against a retrieved bibliographic record for author, title, journal, year, volume, pages, and digital object identifier before inclusion. Five candidate references that could not be confirmed, or whose retrieved record contradicted the citation, were corrected or removed during this process. Readers should note the limitations that follow from a purposive, single-reviewer search: the corpus is not exhaustive, and the risk of selection favoring work consistent with the authors' argument cannot be excluded.

MCS is a sequence of distinct decisions

It is useful to name the decision nodes explicitly, because much of the confusion in this literature comes from treating them as one problem. Candidacy or initiation asks whether a patient who is not yet supported should receive a device and, if so, which one. Its eligible population includes every support-eligible patient, including those who were never implanted. On-support management asks how to titrate flow or speed, how to coordinate that with inotropes, vasoactive drugs, ventilation, volume, and anticoagulation, and when to escalate, combine devices, or recognize futility. Its population is patients already receiving support. Recovery assessment asks whether the native ventricle has improved enough to be tested, typically through structured ramp or low-support protocols [17,18]. Weaning or explantation asks whether to continue support, step down, remove the device, bridge to a durable device or transplant, or transition to palliation [19,20].

These nodes have genuinely different evidence bases. Structured Impella weaning protocols and echocardiographic ramp testing are described in single-center series and algorithmic proposals rather than

randomized trials [18,19,21]. Myocardial recovery on durable support has a longer literature, from the early Harefield combination-therapy work through RESTAGE-HF, with registry analyses showing that sustained recovery sufficient for explantation is uncommon and that the opportunity may be under-recognized [22-28]. Durable device outcomes themselves have been transformed by fully magnetically levitated pumps [29]. A model trained and validated at one node does not transfer to another simply because both involve the same device. Table *2* sets out, for each node, the population in which the question arises, the strongest evidence currently available, and what remains unestablished.

Decision node	Population in which the question arises	Strongest evidence currently available	Not established
Candidacy and initiation	All support-eligible patients, including those never implanted	Randomized trials of initiation strategies in infarct-related cardiogenic shock: IABP-SHOCK II was neutral, including at six years; ECLS-SHOCK and ECMO-CS were neutral; DanGer Shock reduced 180-day mortality, with a higher rate of serious adverse events; an individual patient data meta-analysis of temporary support [30-35]	Which individual patient benefits. No support-specific computational study addresses this node because every cohort in Table 1 is conditioned on the implantation decision.
On-support management: titration and co-intervention	Patients already receiving support	Probabilistic digital twins and offline reinforcement learning applied to micro-axial flow-pump support levels, with policies evaluated inside learned simulators [13-15]	Whether any titration policy improves patient outcomes. The rewards optimized are hemodynamic surrogates, not survival, recovery, or complication burden.
Recovery assessment	Patients already receiving support	Structured ramp and low-support protocols described in single-center series and algorithmic proposals rather than randomized trials [17,18,21]	Prospective or external validation of any recovery-prediction model, with calibration reported alongside discrimination.
Weaning, explantation, and bridging	Patients already receiving support	Weaning and explantation literature from the early Harefield combination-therapy work through RESTAGE-HF, with registry analyses showing sustained recovery to be uncommon and possibly under-recognized [19,20,22-28]; durable device outcomes transformed by fully magnetically levitated pumps [29]	Optimal timing for an individual patient. No learned weaning policy has been run prospectively, silently or otherwise.

TABLE 2: Decision nodes in mechanical circulatory support, the evidence available at each node, and what remains unestablished.

The four decision nodes are those defined in the text. Evidence cited in the third column is the strongest identified in this review and is not exhaustive. Sources: [13-15,17-35]. Digital twin, simulator, and learned-policy studies cited in the third column are described in full in Table 1.

ECLS, extracorporeal life support; ECMO, extracorporeal membrane oxygenation; IABP, intra-aortic balloon pump; DanGer Shock, Danish-German Cardiogenic Shock trial; RESTAGE-HF, Remission from Stage D Heart Failure

What the randomized evidence does and does not settle

Any claim that an algorithm can improve MCS decisions has to be positioned against what randomized trials have shown, which is more specific and more sobering than it is often rendered. In infarct-related cardiogenic shock, intra-aortic balloon counterpulsation did not reduce mortality in IABP-SHOCK II, and the neutral result persisted at six years [30,31]. Routine early venoarterial extracorporeal life support did not reduce 30-day mortality in ECLS-SHOCK, and immediate extracorporeal membrane oxygenation did not improve the composite outcome in ECMO-CS [32,33]. A microaxial flow pump did reduce 180-day all-cause mortality in DanGer Shock, at the cost of a significantly higher rate of serious adverse events, including bleeding, limb ischemia, and renal replacement therapy [34]. An individual patient data meta-analysis of randomized trials of temporary support in this population has since been reported [35].

Three points follow. First, these are average treatment effects in defined populations under defined strategies; none of them licenses a general statement that a device class is good or bad. Second, the DanGer Shock result in particular is a benefit-and-harm result rather than a clean win, which is precisely the situation in which per-patient decision support is attractive and also the situation in which a miscalibrated recommendation does the most damage. Third, and most important for this review, none of these trials

answers the question that a titration or weaning algorithm is trying to answer. They compare strategies at initiation; they do not establish what the optimal support level is for a given patient 18 hours into support, nor when that patient should be tested for recovery.

What a digital twin must specify to earn the name

The term digital twin is applied loosely enough that it is worth stating what the concept requires. A latent vector attached to a patient record is not a twin. A usable formulation specifies a latent physiological state, an observation model linking that state to what is measured, a transition model describing how the state evolves, an explicit action set that can alter transitions, a mechanism for updating patient-specific parameters as data arrive, and calibrated uncertainty in its forecasts [8,9].

For circulatory support, the candidate state variables are reasonably well agreed upon: contractility, preload, afterload, systemic and pulmonary vascular resistance, ventricular interdependence, congestion, end-organ perfusion, and some representation of a recovery regime. Observations include device telemetry, invasive pressures and flows, laboratory values, echocardiography, electrocardiography, and clinical documentation. The action set must include support-level changes and, critically, the co-interventions that are changed at the same time. The quantity of interest is the posterior distribution over physiological state conditional on the information available at a decision node, which is not the same object as a retrospective summary of the whole admission. That distinction sounds pedantic and is not: it is the difference between a model that could have been run at the bedside and one that could not.

Where cardiac digital twins have actually delivered

The strongest cardiac digital twin results to date come from electrophysiology and resynchronization. Personalized virtual-heart models built from imaging have been used to identify ablation targets in infarct-related ventricular tachycardia [36], and patient-specific models have been used to predict left ventricular reverse remodeling after cardiac resynchronization therapy by virtually pacing the model [37]. These are genuine successes, and it is worth being precise about why they were tractable. The modeled substrate is anatomically stable over the relevant horizon, the intervention is discrete and well specified, the model is personalized from a rich imaging study, and the prediction target is a remodeling outcome measured months later rather than a minute-to-minute control decision.

Circulatory support inverts most of these favorable conditions. The substrate changes hour by hour, the intervention is continuously titrated rather than discrete, personalization must come from streaming telemetry rather than a single imaging study, and the decisions of interest occur on a scale of hours. A recent scoping review of digital twin applications in adult critical care found the field still largely at the level of development and feasibility work rather than implementation [38]. Enthusiasm for the electrophysiology results should therefore be discounted when transferred to the circulatory support setting.

Sequence models and the limits of endpoint prediction

A large methodological literature exists for irregular longitudinal clinical data, and it provides the baselines against which any twin should be judged. Recurrent architectures handling missingness, attention-based interpretable models, transformer encoders for electronic health records, and multi-horizon forecasting architectures all address parts of the problem [39-42]. Dynamic survival models produce updated risk estimates as new observations arrive [43,44], and joint longitudinal and time-to-event models remain a strong, interpretable statistical reference point [45]. Classical state-space and hidden-state approaches are the simplest credible starting points and are too often skipped [46,47]. Self-exciting point processes and distributed-lag models offer principled ways to handle event clustering and delayed effects in irregularly observed data [48,49]. Within MCS specifically, machine learning has been applied to predict myocardial recovery resulting in pump explantation among 20,270 patients with durable continuous-flow left ventricular assist devices in the Interagency Registry for Mechanically Assisted Circulatory Support (INTERMACS), with an area under the receiver operating characteristic curve of 0.813 to 0.824 on an internal 70:30 split, exceeding established recovery risk scores at 0.744 to 0.796 [50]. That study is a prediction model rather than a digital twin, simulator, or learned policy: it estimates the probability of an outcome, does not model how physiology responds to an action, and reports discrimination without calibration or external validation.

The recurring evaluation failure in this area is the use of area under the receiver operating characteristic curve as the primary or sole metric. For a probabilistic forecasting system, this is inadequate: it is insensitive to calibration, it says nothing about interval coverage, it is computed over a whole admission rather than at a decision node, and it can be high while the model is systematically late. Proper scoring rules, which are minimized only by the true predictive distribution, are the appropriate primary criterion, supplemented by calibration curves, interval coverage, and multi-step forecast error at clinically meaningful horizons [51]. A twin earns an engineering claim only when it beats strong mechanistic, state-space, and sequence baselines on those measures, on held-out patients, and preferably on external data.

Offline reinforcement learning for device weaning, and what it currently

shows

Reinforcement learning has been applied to sequential treatment problems in critical care for several years, most visibly in sepsis, and the methodological cautions were articulated early [10-12]. Recent work has extended the approach to venovenous extracorporeal membrane oxygenation for respiratory failure, jointly optimizing ventilator and extracorporeal settings in 184 adults at a single tertiary center [52]. That study concerns respiratory rather than circulatory support and is cited here as the nearest methodological analogue, not as evidence about circulatory support decisions. The MCS-specific line of work is newer and small: a data-driven simulator for MCS built with domain-adversarial neural processes [14], a probabilistic digital twin proposed for data-driven smart weaning [13], and a guardian-regularized safe offline reinforcement learning method for weaning that constrains the learned policy toward the region supported by observed data [15].

The design characteristics of this support-specific literature, including the cohorts, rewards, and validation approaches on which it rests, are set out in Table *1*. This work is careful and is honestly reported by its authors, and the criticisms that follow are about what the evidence can bear, not about the quality of the modeling. Four constraints apply. The cohorts are small and drawn from patients already receiving support from a single device family, so nothing in them speaks to who should have been implanted. The reward functions are hemodynamic surrogates rather than patient outcomes, so a reported improvement in reward is an improvement in a proxy whose relationship to survival, recovery, or complication burden is assumed rather than demonstrated. The policies are evaluated in simulators estimated from the same trajectories that trained them, which is discussed below. And there has been no prospective deployment, silent or otherwise, and therefore no evidence about clinician interaction, alert burden, or behavior change.

Why a simulator cannot certify the policy trained against it

The most consequential methodological issue in this literature is circularity. If a transition model is estimated from a set of observed trajectories, and a policy is then optimized to perform well under that transition model, evaluating the policy in the same model measures the consistency of the pipeline with itself. It does not provide independent evidence of clinical benefit. Wherever the learned policy departs from historical practice, which is exactly where its value would lie, the simulator is extrapolating into state-action regions it never observed, and its error there is unconstrained.

It is therefore material whether the cited support-specific studies evaluated their policies against data independent of the data used to build their simulators, and on the published record, none of them did in the sense that matters. In the micro-axial flow pump work, patients were split 65:15:20 into training, validation, and testing sets, so the transition model was fitted on one set of patients and its forecasting accuracy assessed on held-out patients, which is a genuine held-out evaluation of the twin [15]. Policy value, however, was then estimated by model-based off-policy evaluation inside that same learned twin, supplemented by trajectories rolled out from it, rather than against independent real trajectories or an independent value estimator [13,15]. The earlier simulator paper reports no patient-level held-out cohort and no external dataset, splitting each training batch into context and target points instead [14]. The respiratory analogue shows the same pattern: in the venovenous extracorporeal work, fitted Q evaluation was used, and the learned policy was compared with clinician action distributions on the same single-center cohort from which it was learned [52]. Held-out patients therefore support the claim that these dynamics models forecast well; they do not support the claim that the learned policies would act well, because no policy in this literature has yet been valued outside a model estimated from its own training distribution.

Model-based offline policy optimization methods address this explicitly, typically by penalizing the objective in proportion to model uncertainty so that the policy is discouraged from exploiting regions where the dynamics model is unreliable [53]. Off-policy evaluation estimators, including doubly robust constructions that combine importance weighting with a learned value model, give partially independent value estimates but degrade as the learned policy diverges from the behavior policy, and their variance in long-horizon clinical problems is often severe enough to make the estimate uninformative [54]. The practical implications are that behavior-policy overlap must be reported rather than assumed, that value estimates should be accompanied by the effective sample size on which they rest, and that a policy that recommends an action never observed in similar patients should abstain rather than recommend.

Information availability, leakage, and the problem of three clocks

A separate failure mode is temporal rather than causal. Retrospective clinical records collapse several distinct times into one. The time an event physiologically occurred, the time its result first became usable by a clinician or system, and the time it was entered into the record are different, and in intensive care they can differ by hours. A laboratory value timestamped at the moment the specimen was drawn is not knowable at that moment; a note documenting a deterioration is written after the deterioration. Any model trained on a naive extraction that uses recorded time as though it were available time has been given information from the future, and its apparent performance is inflated by an amount that is generally not estimable after the fact.

This is not a hypothetical concern in a domain where the whole proposition is timeliness. A weaning

algorithm that appears to anticipate decompensation may simply be reading documentation created because the decompensation had already been recognized. Guarding against it requires that each event carry event time, available time, and recorded time separately; that features at a decision node be constructed only from what was available at that node; and that validation split patients, sites, and calendar periods rather than randomly splitting observations. Negative-control outcomes and exposures, and deliberate future-information oracles that quantify how much performance a leak would buy, are practical diagnostics [55].

The initiation question requires a target trial, not a post-implant cohort

The most common category error in this field is using a cohort of supported patients to make claims about whether support should be given. A post-implant cohort is conditioned on the implantation decision. It contains no patients who were eligible and not implanted, and therefore contains no information about the counterfactual. No amount of modeling recovers a comparison group that was never sampled.

The appropriate framework is target-trial emulation: specify the eligibility criteria, the time zero, the treatment strategies being compared, the assignment procedure, the outcome, and the follow-up, and then emulate that protocol in observational data, with time-varying confounding handled by methods designed for it, such as marginal structural models [56,57]. Doing this for MCS initiation requires a cohort assembled at the point of eligibility rather than at the point of implantation, and it requires that the co-interventions and clinical reasoning that drove the implantation decision be recorded well enough to be adjusted for. Psychosocial evaluation is a formal part of candidacy assessment for durable support and is associated with outcomes, yet is rarely available in structured form in the datasets used for modeling, which is a concrete example of a determinant of the treatment decision that is systematically missing from the data used to study it [58,59].

What public critical-care datasets can and cannot validate

Public intensive-care datasets have been indispensable to methodological progress. MIMIC-IV, the eICU Collaborative Research Database, and HiRID provide large, well-documented, differently structured cohorts that support development, multicenter held-out evaluation, and higher-frequency physiological work, respectively [60-63]. Emerging common data standards and new cardiovascular longitudinal cohorts make cross-source work more tractable [64,65].

Their limits for this specific problem should be stated plainly. They do not contain synchronized device telemetry at the resolution at which support is actually titrated; they do not record the alternative actions that were considered and rejected; they capture co-interventions incompletely, and they contain no expert-defined set of acceptable actions against which a policy could be judged. They can therefore support claims about state estimation, event recognition, chronology handling, and transport across sites. They cannot validate device-specific initiation, titration, or weaning policy, and a paper that trains a weaning policy on them is making a claim its data cannot carry. Governed multicenter telemetry with synchronized co-interventions and adjudicated outcomes is the missing resource, and its absence, rather than any algorithmic limitation, is what currently caps the field.

Reporting standards, evaluation design, and the regulatory frame

The reporting infrastructure for this work already exists and is underused. TRIPOD+AI covers prediction models developed with regression or machine learning [66], PROBAST+AI provides the corresponding risk-of-bias and applicability assessment [67], DECIDE-AI covers early-stage live clinical evaluation of decision support systems [68], and SPIRIT-AI and CONSORT-AI cover protocols and reports of trials of AI interventions [69,70]. An American Heart Association scientific statement sets out expectations for artificial intelligence in cardiovascular care [71], and general guidance on benchmarking medical AI agents has recently been published [72]. For any system intended to inform device management, the United States Food and Drug Administration's guidance on assessing the credibility of computational modeling and simulation in device submissions defines the question a regulator will ask: what is the specific context of use, and what evidence establishes credibility for that context [73].

Read together, these documents describe an evidence ladder that the current MCS literature has climbed only the first few rungs of. State estimation and forecasting can be established retrospectively with proper scoring rules and external validation. Policy value cannot, and requires overlap-aware off-policy evaluation with honest uncertainty, followed by silent prospective running in which the system's recommendations are recorded but not shown. Only then do human-factors and workflow studies become meaningful, and only after those does a randomized or stepped-wedge evaluation against patient outcomes become appropriate. Skipping rungs is the failure mode to watch for, and the clearest warning sign in a manuscript is a policy-value claim supported only by simulator performance.

Separating what circulatory support evidence shows from what methodology recommends

Because this review draws on two distinct bodies of work, it is worth marking which conclusions rest on

which. Four of its claims are supported directly by MCS evidence. First, the randomized trials establish average effects of initiation strategies in defined populations and establish nothing about titration or weaning, which follows from the trial protocols themselves [30-35]. Second, the support-specific computational literature is confined to cohorts of patients already receiving support, which follows from the cohorts as reported (Table *1*) and is what makes initiation claims unavailable to it. Third, the reward functions used in that literature are hemodynamic surrogates rather than patient outcomes, which is stated by the authors of those studies [13,15]. Fourth, no system in this literature has been run prospectively, silently or otherwise, which follows from the absence of any such report.

The remaining positions taken here are methodological recommendations imported from the offline reinforcement learning, causal inference, and prediction-model literatures, and they are not findings about MCS. That a policy should abstain outside the support of the behavior policy, that off-policy value estimates should be reported with the effective sample size on which they rest, that proper scoring rules should displace area under the receiver operating characteristic curve as the primary forecasting criterion, that event time and available time must be recorded separately, and that initiation questions require target-trial emulation are general results applied here by analogy [51,53-57]. They are, in the authors' view, the right requirements to impose on this field, but they have not themselves been tested in MCS. A reader is entitled to accept the first set of claims while disputing the second.

Limitations

Several limitations should temper the conclusions drawn here. The limitations of the search and selection process are set out in the methodology section above and are not repeated here. The MCS-specific digital twin and offline reinforcement learning literature is small and rapidly moving, and part of it exists as conference proceedings and workshop papers rather than journal articles, which limits both the depth of peer review behind it and the stability of the record. The review is restricted to circulatory support and does not systematically address respiratory extracorporeal support, which is cited only where it is the nearest available analogue [52]. Finally, the argument advanced here is methodological; it identifies what current evidence supports but does not itself provide new empirical evidence, and the evaluation program it implies has not been executed.

Conclusions

In its published record, the support-specific digital twin and offline reinforcement learning literature for MCS establishes improved patient-specific state estimation and short-horizon forecasting with calibrated uncertainty, but nothing yet about whether a learned policy benefits patients. The requirements for a stronger claim, namely, outcome-based rewards, overlap-aware off-policy evaluation, separate recording of event and available time, target-trial emulation for initiation, and silent prospective running before any interventional study, come from the wider methodological literature and remain untested in this setting. Until they are met, investigators should state which decision node a model addresses and what was knowable at that node, and readers, editors, and reviewers should not treat an improvement in simulated reward as a clinical result.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Oscar Mingtian Li, Yijiang Li

Acquisition, analysis, or interpretation of data: Oscar Mingtian Li, Summer Xia Tian Li

Drafting of the manuscript: Oscar Mingtian Li

Critical review of the manuscript for important intellectual content: Yijiang Li, Summer Xia Tian Li

Supervision: Yijiang Li

Disclosures

Conflicts of interest: In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** Yijiang Li declare(s) stock/stock options from Medigenius Investment Inc. Yijiang Li holds an equity interest (stock/stock options) in Medigenius Investment Inc., which is the assignee of pending patent applications relating to blood pump and ultrasonic anticoagulation technologies. No payments or royalties have been received. **Intellectual property info:** Oscar Mingtian Li and Yijiang Li are named inventors on pending patent applications relating to blood pump

and ultrasonic anticoagulation technologies, which are assigned to Medigenius Investment Inc. No patents have issued and no royalties have been received. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Acknowledgements

The generative artificial intelligence tool Claude (Anthropic PBC, San Francisco, CA) was used during the preparation of this manuscript for literature searching, verification of bibliographic records against source databases, and language editing and structural revision of author-written text. All scientific content, argument and interpretation are the authors' own. Every reference was independently checked against a retrieved bibliographic record by the authors, and no text was accepted without author review. The authors take full responsibility for the content of this article. No figures were generated by artificial intelligence; this article contains no figures.

References

1. Tedford RJ, Leacche M, Lorts A, Drakos SG, Pagani FD, Cowger J: Durable mechanical circulatory support: JACC scientific statement. J Am Coll Cardiol. 2023, 82:1464-81. 10.1016/j.jacc.2023.07.019
2. Geller BJ, Sinha SS, Kapur NK, et al.: Escalating and de-escalating temporary mechanical circulatory support in cardiogenic shock: a scientific statement from the American Heart Association. Circulation. 2022, 146:e50-68. 10.1161/CIR.0000000000001076
3. Bernhardt AM, Copeland H, Deswal A, Gluck J, Givertz MM: The International Society for Heart and Lung Transplantation/Heart Failure Society of America guideline on acute mechanical circulatory support. J Heart Lung Transplant. 2023, 42:e1-64. 10.1016/j.healun.2022.10.028
4. Heidenreich PA, Bozkurt B, Aguilar D, et al.: 2022 AHA/ACC/HFSA guideline for the management of heart failure: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. Circulation. 2022, 145:e895-1032. 10.1161/CIR.0000000000001063
5. McDonagh TA, Metra M, Adamo M, et al.: 2021 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure. Eur Heart J. 2021, 42:3599-726. 10.1093/eurheartj/ehab368
6. Potapov EV, Antonides C, Crespo-Leiro MG, et al.: 2019 EACTS expert consensus on long-term mechanical circulatory support. Eur J Cardiothorac Surg. 2019, 56:230-70. 10.1093/ejcts/ezz098
7. Naidu SS, Baran DA, Jentzer JC, et al.: SCAI SHOCK stage classification expert consensus update: a review and incorporation of validation studies: This statement was endorsed by the American College of Cardiology (ACC), American College of Emergency Physicians (ACEP), American Heart Association (AHA), European Society of Cardiology (ESC) Association for Acute Cardiovascular Care (ACVC), International Society for Heart and Lung Transplantation (ISHLT), Society of Critical Care Medicine (SCCM), and Society of Thoracic Surgeons (STS) in December 2021. J Am Coll Cardiol. 2022, 79:933-46. 10.1016/j.jacc.2022.01.018
8. Corral-Acero J, Margara F, Marciniak M, et al.: The "digital twin" to enable the vision of precision cardiology . Eur Heart J. 2020, 41:4556-64. 10.1093/eurheartj/ehaa159
9. Trayanova NA, Prakosa A: Up digital and personal: how heart digital twins can transform heart patient care . Heart Rhythm. 2024, 21:89-99. 10.1016/j.hrthm.2023.10.019
10. Komorowski M, Celi LA, Badawi O, Gordon AC, Faisal AA: The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care. Nat Med. 2018, 24:1716-20. 10.1038/s41591-018-0213-5
11. Gottesman O, Johansson F, Komorowski M, Faisal A, Sontag D, Doshi-Velez F, Celi LA: Guidelines for reinforcement learning in healthcare. Nat Med. 2019, 25:16-8. 10.1038/s41591-018-0310-5
12. Liu S, See KC, Ngiam KY, Celi LA, Sun X, Feng M: Reinforcement learning for clinical decision support in critical care: comprehensive review. J Med Internet Res. 2020, 22:e18477. 10.2196/18477
13. Sun S, Tumay A, Fereidooni S, Dumas A, Jortberg E, Yu R: Probabilistic digital twin for data-driven smart weaning of mechanical circulatory devices. NeurIPS 2025 Workshop on Learning from Time-Series for Health. 2025,
14. Sun S, Chen W, Zhou Z, Fereidooni S, Jortberg E, Yu R: Data-driven simulator for mechanical circulatory support with domain adversarial neural process. Proc Mach Learn Res. 2024, 242:1513-25.
15. Tumay A, Sun S, Fereidooni S, Dumas A, Jortberg E, Yu R: Guardian-regularized safe offline reinforcement learning for smart weaning of mechanical circulatory devices. Proc Mach Learn Res. 2026, 297:1269-96.
16. Baethge C, Goldbeck-Wood S, Mertens S: SANRA-a scale for the quality assessment of narrative review articles. Res Integr Peer Rev. 2019, 4:5. 10.1186/s41073-019-0064-8
17. Uriel N, Morrison KA, Garan AR, et al.: Development of a novel echocardiography ramp test for speed optimization and diagnosis of device thrombosis in continuous-flow left ventricular assist devices: the Columbia ramp study. J Am Coll Cardiol. 2012, 60:1764-75. 10.1016/j.jacc.2012.07.052
18. Tschöpe C, Spillmann F, Potapov E, et al.: The "TIDE"-algorithm for the weaning of patients with cardiogenic shock and temporarily mechanical left ventricular support with Impella devices. A cardiovascular physiology-based approach. Front Cardiovasc Med. 2021, 8:563484. 10.3389/fcvm.2021.563484
19. Slack RJ, McGain F, Cox N, et al.: Structured weaning from the Impella left ventricular micro-axial pump in acute myocardial infarction with cardiogenic shock and protected percutaneous coronary intervention: experience from a non-cardiac surgical centre. Heart Lung Circ. 2024, 33:460-9. 10.1016/j.hlc.2023.12.007
20. Monteagudo Vela M, Rial Bastón V, Panoulas V, Riesgo Gil F, Simon A: A detailed explantation assessment protocol for patients with left ventricular assist devices with myocardial recovery. Interact Cardiovasc Thorac Surg. 2021, 32:298-305. 10.1093/icvts/ivaa259
21. Matassini MV, Marini M, Angelozzi A, et al.: Clinical outcomes and predictors of success with Impella weaning in cardiogenic shock: a single-center experience. Front Cardiovasc Med. 2023, 10:1171956. 10.3389/fcvm.2023.1171956
22. Birks EJ, Tansley PD, Hardy J, et al.: Left ventricular assist device and drug therapy for the reversal of heart

failure. N Engl J Med. 2006, 355:1873-84. 10.1056/NEJMoa053063
23. Birks EJ, George RS, Hedger M, et al.: Reversal of severe heart failure with a continuous-flow left ventricular assist device and pharmacological therapy: a prospective study. Circulation. 2011, 123:381-90. 10.1161/CIRCULATIONAHA.109.933960
24. Birks EJ, Drakos SG, Patel SR, et al.: Prospective multicenter study of myocardial recovery using left ventricular assist devices (RESTAGE-HF [Remission from Stage D Heart Failure]): medium-term and primary end point results. Circulation. 2020, 142:2016-28. 10.1161/CIRCULATIONAHA.120.046415
25. Wever-Pinzon O, Drakos SG, McKellar SH, et al.: Cardiac recovery during long-term left ventricular assist device support. J Am Coll Cardiol. 2016, 68:1540-53. 10.1016/j.jacc.2016.07.743
26. Topkara VK, Garan AR, Fine B, et al.: Myocardial recovery in patients receiving contemporary left ventricular assist devices: results from the Interagency Registry for Mechanically Assisted Circulatory Support (INTERMACS). Circ Heart Fail. 2016, 9:10.1161/CIRCHEARTFAILURE.116.003157
27. Phan K, Huo YR, Zhao DF, Yan TD, Tchantchaleishvili V: Ventricular recovery and pump explantation in patients supported by left ventricular assist devices: a systematic review. ASAIO J. 2016, 62:219-31. 10.1097/MAT.0000000000000328
28. Itagaki S, Moss N, Toyoda N, et al.: Incidence, outcomes, and opportunity for left ventricular assist device weaning for myocardial recovery. JACC Heart Fail. 2024, 12:893-901. 10.1016/j.jchf.2023.12.006
29. Mehra MR, Uriel N, Naka Y, et al.: A fully magnetically levitated left ventricular assist device - final report . N Engl J Med. 2019, 380:1618-27. 10.1056/NEJMoa1900486
30. Thiele H, Zeymer U, Neumann FJ, et al.: Intraaortic balloon support for myocardial infarction with cardiogenic shock. N Engl J Med. 2012, 367:1287-96. 10.1056/NEJMoa1208410
31. Thiele H, Zeymer U, Thelemann N, et al.: Intraaortic balloon pump in cardiogenic shock complicating acute myocardial infarction: long-term 6-year outcome of the randomized IABP-SHOCK II trial. Circulation. 2019, 139:395-403. 10.1161/CIRCULATIONAHA.118.038201
32. Thiele H, Zeymer U, Akin I, et al.: Extracorporeal life support in infarct-related cardiogenic shock . N Engl J Med. 2023, 389:1286-97. 10.1056/NEJMoa2307227
33. Ostadal P, Rokyta R, Karasek J, et al.: Extracorporeal membrane oxygenation in the therapy of cardiogenic shock: results of the ECMO-CS randomized clinical trial. Circulation. 2023, 147:454-64. 10.1161/CIRCULATIONAHA.122.062949
34. Møller JE, Engstrøm T, Jensen LO, et al.: Microaxial flow pump or standard care in infarct-related cardiogenic shock. N Engl J Med. 2024, 390:1382-93. 10.1056/NEJMoa2312572
35. Thiele H, Moller JE, Henriques JPS, et al.: Temporary mechanical circulatory support in infarct-related cardiogenic shock: an individual patient data meta-analysis of randomised trials with 6-month follow-up. Lancet. 2024, 404:1019-28. 10.1016/S0140-6736(24)01448-X
36. Prakosa A, Arevalo HJ, Deng D, et al.: Personalized virtual-heart technology for guiding the ablation of infarct-related ventricular tachycardia. Nat Biomed Eng. 2018, 2:732-40. 10.1038/s41551-018-0282-2
37. Koopsen T, Gerrits W, van Osta N, et al.: Virtual pacing of a patient's digital twin to predict left ventricular reverse remodelling after cardiac resynchronization therapy. Europace. 2023, 26:10.1093/europace/euae009
38. Kim Y, Kim J, Kim Y, Choi M: Digital twin applications in adult critical care: a scoping review of current development and implementation trends. Digit Health. 2026, 12:20552076261438961. 10.1177/20552076261438961
39. Che Z, Purushotham S, Cho K, Sontag D, Liu Y: Recurrent neural networks for multivariate time series with missing values. Sci Rep. 2018, 8:6085. 10.1038/s41598-018-24271-9
40. Choi E, Bahadori MT, Sun J, Kulas J, Schuetz A, Stewart W: RETAIN: an interpretable predictive model for healthcare using reverse time attention mechanism. Adv Neural Inf Process Syst. 2016, 29:
41. Li Y, Rao S, Solares JR, et al.: BEHRT: transformer for electronic health records . Sci Rep. 2020, 10:7155. 10.1038/s41598-020-62922-y
42. Lim B, Arik SO, Loeff N, Pfister T: Temporal fusion transformers for interpretable multi-horizon forecasting . Int J Forecast. 2021, 37:1748-64. 10.1016/j.ijforecast.2021.03.012
43. Lee C, Zame WR, Yoon J, van der Schaar M: DeepHit: a deep learning approach to survival analysis with competing risks. Proc AAAI Conf Artif Intell. 2018, 32:2314-21. 10.1609/aaai.v32i1.11842
44. Lee C, Yoon J, Schaar MV: Dynamic-DeepHit: a deep learning approach for dynamic survival analysis with competing risks based on longitudinal data. IEEE Trans Biomed Eng. 2020, 67:122-33. 10.1109/TBME.2019.2909027
45. Papageorgiou G, Mauff K, Tomer A, Rizopoulos D: An overview of joint modeling of time-to-event and longitudinal outcomes. Annu Rev Stat Appl. 2019, 6:223-40. 10.1146/annurev-statistics-030718-105048
46. Kalman RE: A new approach to linear filtering and prediction problems . J Basic Eng. 1960, 82:35-45. 10.1115/1.3662552
47. Rabiner LR: A tutorial on hidden Markov models and selected applications in speech recognition . Proc IEEE. 1989, 77:257-86. 10.1109/5.18626
48. Hawkes AG: Spectra of some self-exciting and mutually exciting point processes . Biometrika. 1971, 58:83-90. 10.1093/biomet/58.1.83
49. Gasparrini A, Armstrong B, Kenward MG: Distributed lag non-linear models. Stat Med. 2010, 29:2224-34. 10.1002/sim.3940
50. Topkara VK, Elias P, Jain R, Sayer G, Burkhoff D, Uriel N: Machine learning-based prediction of myocardial recovery in patients with left ventricular assist device support. Circ Heart Fail. 2022, 15:e008711. 10.1161/CIRCHEARTFAILURE.121.008711
51. Gneiting T, Raftery AE: Strictly proper scoring rules, prediction, and estimation . J Am Stat Assoc. 2007, 102:359-78. 10.1198/016214506000001437
52. Song J, Dave SB, Yang Y, et al.: rECMOmender: reinforcement learning for decision support in venovenous extracorporeal membrane oxygenation management. Crit Care Explor. 2026, 8:e1369. 10.1097/CCE.0000000000001369
53. Yu T, Thomas G, Yu L, et al.: MOPO: model-based offline policy optimization . Adv Neural Inf Process Syst. 2020, 33:14129-42.

54. Jiang N, Li L: Doubly robust off-policy value evaluation for reinforcement learning . Proc Mach Learn Res. 2016, 48:652-61.
55. Lipsitch M, Tchetgen Tchetgen E, Cohen T: Negative controls: a tool for detecting confounding and bias in observational studies. Epidemiology. 2010, 21:383-8. 10.1097/EDE.0b013e3181d61eeb
56. Hernán MA, Robins JM: Using big data to emulate a target trial when a randomized trial is not available . Am J Epidemiol. 2016, 183:758-64. 10.1093/aje/kwv254
57. Robins JM, Hernán MA, Brumback B: Marginal structural models and causal inference in epidemiology . Epidemiology. 2000, 11:550-60. 10.1097/00001648-200009000-00011
58. Dew MA, DiMartini AF, Dobbels F, et al.: The 2018 ISHLT/APM/AST/ICCAC/STSW recommendations for the psychosocial evaluation of adult cardiothoracic transplant candidates and candidates for long-term mechanical circulatory support. Psychosomatics. 2018, 59:415-40. 10.1016/j.psym.2018.04.003
59. Choi JJ, Bhasin S, Levstik J, Walsh P, Oldham MA, Lee HB: Association between psychosocial factors and left ventricular assist device implant outcomes: a systematic review. Gen Hosp Psychiatry. 2024, 90:132-40. 10.1016/j.genhosppsych.2024.08.007
60. Johnson AE, Bulgarelli L, Shen L, et al.: MIMIC-IV, a freely accessible electronic health record dataset . Sci Data. 2023, 10:1. 10.1038/s41597-022-01899-x
61. Pollard TJ, Johnson AE, Raffa JD, Celi LA, Mark RG, Badawi O: The eICU Collaborative Research Database, a freely available multi-center database for critical care research. Sci Data. 2018, 5:180178. 10.1038/sdata.2018.178
62. Faltys M, Zimmermann M, Lyu X, Huser M, Hyland S, Ratsch G, Merz T: HiRID, A High Time-Resolution ICU Dataset (version 1.1.1). PhysioNet, 2021. 10.13026/nkwc-js72
63. Hyland SL, Faltys M, Hüser M, et al.: Early prediction of circulatory failure in the intensive care unit using machine learning. Nat Med. 2020, 26:364-73. 10.1038/s41591-020-0789-4
64. McDermott MB, Steinberg E, Fries JA, et al.: MEDS - an emerging data standard and ecosystem for health AI research. NEJM AI. 2026, 3:10.1056/AIra2501253
65. Zha L, Fu C, Sha X, Yin P, Li Y: CardioEHR: a longitudinal electronic health record dataset of cardiovascular patients from central China. Sci Data. 2026, 13:10.1038/s41597-026-06855-7
66. Collins GS, Moons KG, Dhiman P, et al.: TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ. 2024, 385:e078378. 10.1136/bmj-2023-078378
67. Moons KG, Damen JA, Kaul T, et al.: PROBAST+AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. BMJ. 2025, 388:e082505. 10.1136/bmj-2024-082505
68. Vasey B, Nagendran M, Campbell B, et al.: Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nat Med. 2022, 28:924-33. 10.1038/s41591-022-01772-9
69. Cruz Rivera S, Liu X, Chan AW, Denniston AK, Calvert MJ: Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. Nat Med. 2020, 26:1351-63. 10.1038/s41591-020-1037-7
70. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK: Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nat Med. 2020, 26:1364-74. 10.1038/s41591-020-1034-x
71. Armoundas AA, Narayan SM, Arnett DK, et al.: Use of artificial intelligence in improving outcomes in heart disease: a scientific statement from the American Heart Association. Circulation. 2024, 149:e1028-50. 10.1161/CIR.0000000000001201
72. Ruhrberg Estévez S, Ferber D, van der Schaar M, Kather JN: How to benchmark medical AI agents . PLoS Med. 2026, 23:e1005170. 10.1371/journal.pmed.1005170
73. Assessing the credibility of computational modeling and simulation in medical device submissions: guidance for industry and Food and Drug Administration staff. (2023). Accessed: September 9, 2026: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/assessing-credibility-computational-modeling....