

Digital twin applications in adult critical care: A scoping review of current development and implementation trends

DIGITAL HEALTH
Volume 12: 1–20
© The Author(s) 2026
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20552076261438961
journals.sagepub.com/home/dhj


Yeonwoo Kim¹ , Jiin Kim¹ , Yeonju Kim¹ , and Mona Choi^{2,3,4} 

Abstract

Objective: Digital twins (DTs) show promise in critical care by enabling personalised treatment and optimising clinical decision-making. Despite the complexity and data-intensive nature of critical care, the implementation of DTs in this setting remains under-investigated. This scoping review aimed to summarise DT research in critical care and identify current evidence gaps. **Methods:** Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) guidelines, seven electronic databases were searched. Studies reporting the development or evaluation of DT models in adult critical care were included. Data were extracted on study characteristics and DT development features, including modelling approaches, levels of data integration, and key findings. **Results:** Twenty-three studies were included, with most originating from North America and Europe. Retrospective designs using hospital datasets derived from intensive care unit and emergency department settings were common. Data integration predominantly corresponded to the digital model level of the DT maturity, whereas fully automated DT implementations were rare. Regarding modelling approaches, mathematical models were most frequently developed, followed by machine learning-based predictive models. DT application primarily focused on predictive modelling and virtual patient simulations to enhance personalised treatment, support clinical decision-making, and optimise organisational resource allocation. **Conclusion:** DT technologies in critical care remain in the exploratory and early stages of development and implementation. Further research incorporating higher levels of data integration, real-time deployment, and longitudinal external validation is warranted, alongside broader consensus on ethical governance and data privacy.

Keywords

critical care, clinical decision support, digital twin, digital health, intensive care units, scoping review

Received: 24 November 2025; Revised: 28 February 2026; Accepted: 19 March 2026

Introduction

Digital twins (DTs) are dynamic virtual representations of physical systems updated using real-time data.¹ DTs originated from simulation models developed by the National Aeronautics and Space Administration in the 1970s to support the prevention and resolution of Apollo mission accidents^{2,3} and have since evolved across academic domains.^{4,5} Since the

¹College of Nursing and Brain Korea 21 FOUR Project, Yonsei University, Seoul, Republic of Korea

²College of Nursing and Mo-Im Kim Nursing Research Institute, Yonsei University, Seoul, Republic of Korea

³JBI Yonsei Evidence Based Nursing Centre of Korea, University of Adelaide, Adelaide, South Australia, Australia

⁴Yonsei Institute for Digital Health, Yonsei University, Seoul, Republic of Korea

Corresponding author:

Mona Choi, Yonsei University College of Nursing, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 03722, Republic of Korea.

Email: monachoi@yuhs.ac



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

2000s, DTs have gained considerable attention in the context of the Fourth Industrial Revolution, particularly for managing product lifecycles and creating virtual counterparts for physical systems for real-time synchronisation, simulation, and predictive modelling.⁶ Information and communication technology advances, along with the increasing adoption of precision medicine, have expanded DT applications in the healthcare industry, enabling the exploration of their potential across diverse clinical contexts.⁶

In clinical settings, DT technology has been implemented at multiple levels, ranging from organisational systems to patient-specific organs.⁷ Additionally, at the organisational-level, AI-supported DT systems have been deployed in the radiology department to address the challenges posed by clinical complexity, ageing infrastructure, workflow delays, and growing patient demand.⁸ At the patient level, DT interventions for type 2 diabetes create patient-specific virtual replicas by integrating continuous glucose monitoring, nutrition, activity, and sleep data, enabling the prediction of postprandial glucose responses and tailored recommendations, leading to improvements in haemoglobin A1c levels, medication use, and overall metabolic outcomes.⁹ These examples demonstrate the broad utility of DTs across hospital management and personalised medicine.¹⁰

Although DT technology has shown promise in various healthcare settings, its application in critical care remains in its infancy, with limited empirical research and clinical implementation owing to significant barriers in data integration, regulatory governance, and scalability.⁷ While critical care is often synonymous with the intensive care unit (ICU), this review explicitly encompasses high-acuity environments such as emergency departments (EDs) and step-down units, acknowledging that the continuum of critical illness frequently begins with emergency stabilisation.¹¹ This broader scope is essential to capture DT applications across the full trajectory of acute and critical illness.¹² Critical care environments are characterised by continuous data streams, rapid patient status fluctuations, and an ongoing influx of multivariate clinical information.¹³ Notably, clinicians must interpret these signals, make high-stakes decisions in real time, and deliver life-sustaining therapies to critically ill patients, thereby increasing the need for advanced informatics solutions.^{13,14} Further, information overload and patient heterogeneity limit the effectiveness of standardised treatment protocols, which may not adequately capture individual physiological responses. These limitations highlight the need for precise medical approaches tailored to critical illness.^{15,16} Importantly, DTs can address this gap by providing a clearer physiological representation of patient states, enabling more consistent and effective care.¹⁷

DTs integrate diverse data sources, including bedside monitoring, imaging, pharmacokinetic/pharmacodynamic models, and electronic health records (EHRs), to construct and continuously refine patient-specific virtual models.^{18,19} These models can facilitate dynamic risk assessments, treatment simulations, and individualised care pathways. By synthesising multimodal patient data, DTs offer a framework for simulating disease progression, predicting treatment responses, and conducting *in silico* evaluations of therapeutic strategies without patient risk, thereby supporting workflow optimisation and decision support.^{18,20} Biomedical data exponential growth, along with advances in data science, artificial intelligence (AI), and cyberinfrastructure, has accelerated DT research and development in healthcare,²¹ with governmental and major funding bodies prioritising research in this domain. Recent initiatives, such as the U.S. National Science Foundation's 2024 Digital Twin Research Program, have further demonstrated the rapid field expansion.²²

Despite growing interest, inconsistent use of the term DT in the literature risks conflating it with simpler digital representations, such as digital models or shadows.²³ To address this conceptual ambiguity, DT-related technologies can be categorised operationally into three levels based on the direction and automation of data exchange between physical and digital entities: Digital Model (DM), Digital Shadow (DS), and Digital Twin (DT).²⁴ A DM is a static digital representation with no automated data exchange between physical and digital entities. A DS enables unidirectional data flow from the physical system to the digital representation, allowing real-time updates without feedback to the physical entity. In contrast, a DT is characterised by automated bidirectional data exchange, where the physical and digital entities dynamically interact, enabling real-time simulation and decision support.²⁴ Given that many studies described as DT may actually correspond to DM or DS, this classification was used to accurately assess the level of technical implementation in critical care research.

Recently, several reviews have examined DT use in healthcare, including a systematic review of precision health outcomes regarding DT technology,²⁵ a broad scoping review of DT applications across health domains,⁶ and a narrative review outlining the potential roles of DT in critical and acute care medicine.²⁶ Nonetheless, evidence specific to critical care remains limited, as the only review specific to this area remains conceptual insights. Preliminary studies suggest that DTs can support the early detection of patient deterioration, guide therapy optimisation, and enable the *in silico* simulation of clinical strategies in critical settings. Nevertheless, these findings remain fragmented across diverse modelling approaches and use cases.²⁶

To date, no systematic or scoping review has provided a comprehensive overview of the development and implementation of DT specifically in critical care settings. Given the unique complexity of these environments, a focused synthesis of DT-related research is warranted. Therefore, this scoping review aimed to systematically summarise published studies on DTs across all development and clinical application stages in critical care. This review comprehensively charts the landscape of

DT applications in adult critical care by examining study characteristics, aims, target populations, modelling approaches, and the level of data integration—classified as DM, DS, or DT. Through this analysis, we evaluated the extent of technological development and clinical implementation, provided an overview of the current state of research and identified existing gaps. By integrating evidence from proof-of-concept models and early clinical evaluations, this review provides insights to guide future research and advance the clinical translation of DT technologies in critical care.

Methods

Design

This scoping review was conducted in accordance with the Joanna Briggs Institute (JBI) methodological guidance for scoping reviews and the procedures described in Chapter 10 of the JBI Manual for Evidence Synthesis.^{27,28} Additionally, it was reported in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension of the Scoping Reviews) checklist.²⁹ The PRISMA-ScR checklist for this review is presented in Supplementary File 1 (Table S1). In accordance with the scoping review framework, a formal risk-of-bias assessment was not performed, as the objective was to summarise methodological characteristics rather than to evaluate causal effect estimates.

The population, concept, and context (PCC) framework for this scoping review is as follows: (1) population: adult critically ill patients, (2) concept: digital twin (DT) technologies, and (3) context: critical care settings. The scoping review protocol was registered in the Open Science Framework Registry (registration number: <https://doi.org/10.17605/OSF.IO/62HJD>).

Search strategy (Information sources)

A comprehensive search strategy was employed, encompassing seven electronic databases (i.e., PubMed, IEEE Xplore, CINAHL, Cochrane Library, Embase, Scopus, and Web of Science). These databases were systematically searched in accordance with the PCC framework. Database-specific keywords and indexed terms were used without date restrictions, applying the “All fields” option where appropriate (Supplementary File 2, Table S2). Considering the heterogeneity in related terminology (e.g., virtual patient, physiologic simulator, digital avatar, digital model), the search was intentionally restricted to the explicit “digital twin*” to maintain conceptual specificity and ensure the retrieval of studies aligned with established DT criteria. The search strategy was reviewed and confirmed by an experienced health sciences librarian. The primary search was conducted on 20 January 2025 and subsequently updated on 17 July 2025.

An example of the full PubMed search string was as follows: “digital twin*” [All Fields] AND (“Critical care” [MeSH Terms] OR “Critical Illness” [MeSH Terms] OR “Intensive Care Units” [MeSH Terms] OR “Critical Care Nursing” [MeSH Terms] OR “Critical care” [All Fields] OR “critical illness” [All Fields] OR ICU [All Fields] OR “intensive care” [All Fields] OR “intensive care unit*” [All Fields])

Eligibility criteria

Using the PCC framework, the following inclusion criteria were established: (1) population—studies that utilised patient data collected prospectively or retrospectively in real-world clinical settings or that included virtual patients simulated or generated from clinical data; (2) concept—research applying DT technologies or describing the development of a DT model for healthcare, provided that the system created an explicit virtual representation of a critical care entity (e.g., patient, organ system, workflow, or facility) and enabled dynamic modelling of state transitions in response to simulated or real-world clinical interventions; and (3) context—studies conducted in critical care settings such as ICUs and high-acuity care settings (e.g. ED).

To capture emerging developments in this rapidly evolving field, both peer-reviewed journal articles and relevant grey literature (e.g., preprints and conference papers) were considered eligible, provided they presented original research with sufficient methodological detail to enable appraisal and data extraction.

The exclusion criteria were as follows: (1) population—studies that included infant or paediatric patients; (2) concept—research that solely describes the theoretical frameworks of DT technologies, or that are limited to static prediction models, risk scoring tools, or data visualisation dashboards without an explicit virtual representation and state-transition mechanism; (3) context—studies not conducted in critical care settings (e.g. general wards, outpatient clinics, rehabilitation facilities, long-term care), including those in nonhospital or community-based contexts; (4) language—research not written in English; and (5) study design—review articles, editorials, commentaries, letters, book chapters, dissertations, and other non-original research articles.

Study selection

All the studies identified in the comprehensive search were exported to EndNote 21 (Clarivate Analytics) for reference. After removing duplicates using EndNote, the remaining studies were screened using Google Sheets (Google LLC, Mountain View, CA, USA). The studies were independently screened by two researchers (Yeonw K and JK) according to the eligibility criteria. Any disagreements were resolved through discussion with a third researcher (Yeonj K or MC). Studies for which consensus was achieved were included in the final selection.

Data extraction and analysis

Two researchers (Yeonw K and JK) independently extracted data from the selected studies. The process was facilitated by Elicit, an AI-based literature review assistant,³⁰ previously applied in a scoping review.³¹ Elicit was used to identify the pre-defined data fields (e.g., study design, study objectives, datasets, participant counts, and outcomes). All AI-generated outputs were treated as provisional and were not directly incorporated into the review. Each extracted data element was manually cross-validated against the original full-text articles by two independent reviewers. Discrepancies were resolved through manual correction and discussion with a third researcher (Yeonj K). No AI-generated data were retained without independent human verification.

Data extraction forms were constructed for each included study and comprised three parts: (1) overall study characteristics, including year, continent, study design, study setting, data sources, and level of data integration of DTs; (2) study aims and clinical context, including dataset characteristics, critical care setting, clinical conditions, sample size, and intended end users; and (3) DT development and evaluation features, including modelling approaches, level of data integration of DTs, clinical input data, validation strategy, evaluation metrics, and key findings.

The extracted data were synthesised descriptively and organised into a tabular format (Tables 1–3) to present study characteristics, clinical context, and DT modelling and evaluation features. Modelling approaches were categorised into four groups, mathematical, machine learning (ML)-based, simulation-based, or rule-based, to identify methodological trends. Studies were classified as DM, DS, or DT according to their level of data integration. Clinical input data, validation strategies, evaluation metrics, and key findings were summarised narratively to capture variations across studies.

Results

Study selection

A systematic search conducted on 20 January 2025 and updated on 17 July 2025 identified 1,631 records. Following the removal of 163 duplicates, 1,468 titles and abstracts were screened, leading to exclusion of 1,344 records. Of the 116 full-text articles assessed for eligibility, 93 were excluded, resulting in 23 studies included in the final synthesis. Of these, 18 were classified as primary studies (16 peer-reviewed journal articles and two preprints), while five conference papers were presented separately in Supplementary File 3 (Table S3). The full selection process is illustrated in the PRISMA flow diagram (Figure 1).

Study characteristics

Table 1 summarises the overall characteristics of the included peer-reviewed articles and preprints. The publication years ranged from 2020 to 2025, with a significant majority of the studies published after 2024 (11/18, 61.1%).^{12,38–47} Geographically, half of the studies were conducted in North America, specifically the United States (9/18, 50%).^{12,18,32,33,35,40,42,44,45}

Regarding study design, most were retrospective (11/18, 61.1%),^{12,18,33,34,36,37,39,41,43,45,46} although two involved secondary analyses of prospectively collected cohort data.^{41,46} Half of the included studies were conducted in single-centre setting (9/18, 50%).^{18,32,33,35,39,41,45–47} “Not applicable” was assigned to studies based solely on simulation environments, rule-based modelling frameworks, or virtual platforms without real-world clinical data sources.^{40,42,44}

Most studies used hospital datasets (11/18, 61.1%), including single-centre hospital cohorts or multicentre clinical trial datasets,^{12,18,32,34,35,37,41,43,45–47} whereas publicly available datasets such as the Medical Information Mart for Intensive Care (MIMIC) were used in three studies^{33,36,39} (3/18, 16.7%). A small number of studies relied on expert rule-based^{42,44} or simulation-generated data,⁴⁰ and one randomised controlled trial did not explicitly report its data source.³⁸

In terms of the level of data integration of DTs, the majority were classified as DM (15/18, 83.3%),^{6,31–35,37,39–44,47} whereas one study was categorised as a DS³⁸ and two as full DT.^{40,47}

Table 1. Characteristics of the included studies (n=18).

Category	Number of studies (n, %)	Reference
Year		
2020	1 (5.6)	32
2021	2 (11.1)	33,34
2022	4 (22.2)	18,35–37
2024	8 (44.4)	12,38–44
2025	3 (16.7)	45–47
Continents		
Asia	3 (16.7)	37,39,47
Europe	5 (27.8)	36,38,41,43,46
North America	9 (50.0)	12,18,32,33,35,40,42,44,45
Oceania	1 (5.6)	34
Study designs		
Retrospective studies	11 (61.1)	12,18,33,34,36,37,39,41,43,45,46
Prospective studies	2 (11.1)	32,47
Simulation/Proof-of- concept	3 (16.7)	35,40,44
Usability test	1 (5.6)	42
Randomised controlled trial	1 (5.6)	38
Study settings		
Single-centre	9 (50.0)	18,32,33,35,39,41,45–47
Multicentre	6 (33.3)	12,34,36–38,43
Not applicable ^a	3 (16.7)	40,42,44
Data sources		
Hospital dataset	11 (61.1)	12,18,32,34,35,37,41,43,45–47
Publicly available dataset	3 (16.7)	33,36,39
Expert rule-based	2 (11.1)	42,44
Simulation-generated data	1 (5.6)	40
Not reported [*]	1 (5.6)	38
Level of data integration of DTs		
Digital Model	15 (83.3)	12,18,32–36,39,41–46,48
Digital Shadow	1 (5.6)	38
Digital Twin	2 (11.1)	40,47

^a“Not applicable” under Study Setting refers to studies conducted exclusively in simulation environments, rule-based modelling frameworks, or virtual platforms without the use of real-world clinical data.

^{*}“Not reported” under Data Sources indicates that the data sources used for model development were not explicitly described, typically because the studies evaluated pre-existing clinical systems.

Study aims and clinical context

Table 2 presents the study aims and the clinical context, including dataset, critical care setting, clinical conditions, sample sizes, and intended end user. ICUs were the most common environments examined settings, accounting for 83.3% (15/18) of the included studies.^{18,32–34,36–43,46,47} One study was conducted in an ED.¹² Two did not report the type of care unit. One of these involved hospitalised patients with trauma-related acute respiratory distress syndrome,³⁵ whereas the other implemented simulation-based scenarios representing critically ill patients and was therefore categorised as “Not reported”.⁴⁴

Regarding clinical conditions, eight studies (44.4%) focused on respiratory conditions, including patients requiring mechanical ventilation or high-flow nasal cannula therapy.^{34–38,41,43,46} Two studies each focused on sepsis^{32,39} and haemodynamic instability.^{12,47} One study examined ischaemic stroke progression³³ and another focused on patients with heart failure who underwent left ventricular assist device (LVAD) implantation.⁴⁵ The remaining studies (4/18, 22.2%) addressed mixed critical illness scenarios without focusing on specific patient conditions. Of these, two studies focused on simulating ICU environments for critical care delivery optimisation or nurse-robot collaboration,^{18,40} and two were designed for simulation-based education using rule-based models.^{42,44}

Sample sizes were extracted as reported by author and reflected the cohorts used for model development or clinical evaluation, depending on study design. Four studies included more than 1,000 samples, all of which were retrospective in design and utilised large-scale datasets, such as the publicly available MIMIC-IV and eICU Collaborative Research Database

Table 2. Summary of study aims and clinical context of included studies (n=18).

Author, country, & year	Aim of the study	Dataset	Critical care setting	Clinical conditions	Sample sizes	End user
Lal, United States, 2020 ³²	To develop a decision-support DT model for critically ill patients to predict response to specific treatments during the first 24 h of sepsis	EHRs (Mayo clinic)	Medical ICU	Sepsis	29 patients	Physicians
Allen, United States, 2021 ³³	To build a generative model for simulating clinical trajectories in ischaemic stroke patients	EHRs (MIMIC-IV)	ICUs	Ischaemic stroke	1,216 patients	Physicians
Zhou, New Zealand, 2021 ³⁴	To develop a generalised digital clone model to accurately predict lung mechanics in response to MV setting changes	Multicentre EHRs (Eight hospitals in German and Maastricht University Medical Centre)	ICUs	Patients with MV	32 patients	Physicians
Ang, Malaysia, 2022 ³⁷	To present a framework for generating virtual patients for testing model-based decision support, and direct use in MV treatment	Ventilator data (Two clinical cohorts)	ICUs	Respiratory failure patients receiving invasive MV	35 patients	Physicians
Chakshu, United Kingdom, 2022 ³⁶	To develop a DT system to prioritise ICU and ventilator support for pneumonia patients, using severity indices	EHRs (eICU-CRD v2.0, MIMIC-III)	ICUs	Severe pneumonia with MV	Approximately 60,000 admissions	Physicians
Cockrell ^a , United States, 2022 ³⁵	To demonstrate the generation of synthetic time-series data for pro- and anti-inflammatory cytokines using a mechanism-based CIDT	EHR (Walter Reed National Military Medical Centre)	Not reported	ARDS	199 trauma patients	ML/AI model developers
Zhong, United States, 2022 ¹⁸	To develop and test a hybrid simulation model as a step toward a full-fledged ICU DT for improving critical care delivery	EHRs (Mayo Clinic ICU DataMart)	Medical ICU	Critically ill ICU patients	2,809 patients	ICU operational decision-makers
Patel, United Kingdom, 2024 ³⁸	To investigate the clinical efficacy of a physiological model DT-based ventilatory DSS to personalise therapy in ARDS patients	Multicentre ICU cohort	Adult ICUs	ARDS with invasive MV	Randomised, 95 patients	Physicians

(continued)

Table 2. (continued)

Author, country, & year	Aim of the study	Dataset	Critical care setting	Clinical conditions	Sample sizes	End user
Danesh, South Korea, 2024 ³⁹	To develop a personalized, real-time sepsis prediction framework by integrating DT technology with a novel, layered ML model based on dynamic ensemble learning	EHRs (MIMIC-IV)	ICUs	Sepsis	13,948 patients	Physicians
Anyene, United States, 2024 ⁴⁰	To describe the integration of a DT and DES to demonstrate the benefits associated with this co-simulation on healthcare planning and management	Simulation-based data (Jump Simulation Center)	Tele-ICU	Critically ill ICU patients	Three virtual patient scenarios	Healthcare managers
Cannon, United States, 2024 ¹²	To create a mechanistic DT model that simulates individualised resuscitation strategies for haemorrhagic shock, linking preclinical and clinical data	PROMMTT clinical cohort and porcine experimental data	ER	Haemorrhagic shock	Sixteen pigs and 68 patients	Physicians
Rovati, United States, 2024 ⁴²	To assess the usability, workload, and acceptance of the DT application software for critical care education in a cohort of internal medicine residents	Expert rule-based	Medical ICU	Critically ill ICU patients	35 internal medicine residents	Internal medicine residents
Weaver, United Kingdom, 2024 ⁴¹	To clarify the mechanistic basis for the success or failure of non-invasive ventilation in AHRF by using the DT	EHRs (University Hospital of Modena)	Respiratory ICU	AHRF patients with non-invasive trial	25 patients	Physicians
Sun, Belgium, 2024 ⁴³	To estimate patient-specific inspiratory work using a validated DT lung mechanics model in assisted MV modes to optimise patient-ventilator interaction in the ICU	Ventilator data (Two hospitals)	ICUs	Patients with MV	22 patients	Physicians

(continued)

Table 2. (continued)

Author, country, & year	Aim of the study	Dataset	Critical care setting	Clinical conditions	Sample sizes	End user
Trevena, United States, 2024 ⁴⁴	To develop a scalable full-stack architecture for cross-platform patient simulation using graph-based models and present a proof-of-concept	Expert rule-based	NR	Critically ill ICU patients	One virtual patient (case study)	Physicians
Shamohammadi, Italy, 2025 ⁴⁶	To use DTs to investigate the physiological basis behind the varying effects of increased HFNC flow rates on different patients	EHRs (Fondazione IRCCS Ca'Granda Ospedale Maggiore Policlinico)	ICUs	Non-intubated AHRF patients	10 patients	Physicians
Nair, India, 2025 ⁴⁷	To develop an AI-enhanced closed-loop noradrenaline infusion system using patient clustering and DT cardiac models to personalise drug delivery in ICU	Vital signs and inotrope infusion data (Amrita Institute of Medical Sciences and Research Centre)	Cardiac, gastric surgery, neurosurgery and infection control ICUs	Critically ill patients requiring inotropic support	31 patients	Clinicians (physicians and nurses)
Umo ^a , United States, 2025 ⁴⁵	To validate PSCOPE's predictions of post-surgical physiology against clinical data in the context of HeartMate 3 LVAD implantation	Interagency Registry for Mechanically Assisted Circulatory Support database	Cardiothoracic/advanced heart-failure ICU	Heart failure patients who underwent LVAD implantation	Three patients	Physicians

Note. "Sample size" indicates the number of subjects included in model development and internal validation.

^aPreprint.

AHRF: acute hypoxaemic respiratory failure; AI: artificial intelligence; ARDS: acute respiratory distress syndrome; CIDT: Critical Illness Digital Twin; DES: discrete-event simulation; DSS: decision support system; DT: digital twin; EHR: electronic health record; eICU-CRD: eICU Collaborative Research Database; ER: emergency department; HFNC: high flow nasal cannula; ICU: intensive care unit; ITU: intensive therapy unit; LVAD: left ventricular assist device; MIMIC: Medical Information Mart for Intensive Care; ML: machine learning; MV: mechanical ventilation; PROMMTT: Prospective, Observational, Multicenter, Major Trauma Transfusion; PSCOPE: Physiology Simulation Coupled Experiment; RCT: randomised controlled trial; Tele-ICU: tele-intensive care unit.

or institutional electronic health record databases (e.g., Mayo Clinic).^{18,33,36,39} Most studies included fewer than 50 samples (11/18, 61.1%),^{32,34,37,40–47} with two prospective studies including 29 and 31 patients.^{32,47}

Regarding intended end users, most studies targeted physicians, either explicitly or as part of broader clinician groups (13/18, 72.2%).^{12,32–34,36–39,41,43–46} A smaller number of studies identified other professional users, including internal medicine residents in ICUs,⁴² and multidisciplinary ICU clinicians (physicians and nurses).⁴⁷ One study demonstrated active nursing involvement in DT use.³⁸ Beyond direct clinical users, several studies were designed for non-bedside stakeholders, such as ICU operational decision-makers, healthcare managers, and ML/AI model developers.^{18,35,40}

Modelling approaches and key findings

Table 3 synthesises the DT modelling approaches, levels of data integration, clinical input data, validation strategies, evaluation metrics, and key findings of the included studies (n = 18). The modelling approaches were heterogeneous, encompassing mathematical models, ML-based models, simulation-based frameworks, and rule-based systems. Mathematical models were the most common approach, representing physiological dynamics (8/18, 44.4%),^{12,34,37,38,41,43,45,46} followed by ML approaches (4/18, 22.2%) primarily used for predictive modelling and risk stratification.^{33,36,39,47}

Table 3. Synthesis of digital twin modelling, evaluation features, and key findings (n=18).

Author, year	Modelling approach	Level of data integration of DTs	Clinical input data	Validation strategy	Evaluation metrics	Key findings
Lal, 2020 ³²	Bayesian networks (DAG), ABMS, DES	DM	Demographics, V/S, GCS, lab data, SOFA score, source of sepsis, type of interventions	Clinical validation (comparison with real data)	Kappa coefficient: 0.41 – 0.65	Feasibility of developing and prospectively testing a causal AI model for sepsis treatment response prediction was confirmed, showing fair primary and good secondary response agreement.
Allen, 2021 ³³	β -VAE	DM	Lab data, demographic, medical history	Internal validation (80% training, 10% CV, 10% testing)	AUROC_adversary = 0.51	The β -VAE model accurately predicted disease progression, producing simulated trajectories statistically indistinguishable from the real data
Zhou, 2021 ³⁴	Hysteresis loop model	DM	Pressure-Volume loop data (one breath at baseline PEEP)	Clinical validation (comparison with real data)	R ² values (0.83–0.90)	Virtual patient models accurately predicted lung responses to PEEP changes in both volume and pressure-control ventilation for key lung mechanical parameters
Ang, 2022 ³⁷	Single compartment lung model.	DM	Demographics, diagnosis, ventilation mode, days of recording	Internal validation (patient-level)	Prediction errors of 3.26–6.80%, R ² values (0.74–0.92)	Virtual patient models simulated lung mechanical responses with low prediction errors and high correlation between clinical data and predicted responses

(continued)

Table 3. (continued)

Author, year	Modelling approach	Level of data integration of DTs	Clinical input data	Validation strategy	Evaluation metrics	Key findings
Chakshu, 2022 ³⁶	Deep learning using MLP and RNN (LSTM)	DM	Demographics, V/S, lab data, blood gases, clinical status, GCS, MV settings	Internal validation (80% train/20% test; 10-fold cross-validation)	AUROC (0.83-0.89 for mortality and requiring MV)	The DT-based three-tier system accurately predicted the probability of death and the need for MV, enabling effective patient prioritisation for ICU admission and ventilator support and outperforming conventional scoring systems
Cockrell ^a , 2022 ³⁵	Mechanistic agent-based simulation	DM	V/S, labs and cytokine time-series	Same-dataset calibration validation	Visually reported	Synthetic trajectories match clinical ranges and provide richer dynamics; ARDS and non-ARDS parameter spaces overlap but have distinct regions
Zhong, ^{1,8} 2022	Hybrid Simulation, ABS and DES	DM	Patient data, clinician factors and task data, system and resource data	System level validation (comparison with real data)	95% CI: census (12.7-13.2), deaths (0.22-0.24), discharges (4.07-4.27)	The DT model simulated real-life ICU events and provided insights into resource allocation and patient safety; interventions such as reducing task durations and adopting a new discharge policy improved performance metrics
Patel, 2024 ³⁸	Mathematical physiological modelling	DS	NR	Prospective RCT-based clinical validation	Driving pressure (p = 0.72); significant secondary effects (p = 0.037; p = 0.0068); ~60% advice adherence	Applying the DSS advice did not change average driving pressure but improved the oxygenation index and ventilatory ratio in support and control modes, respectively, leading to more frequent adjustments of pressure and respiratory rate

(continued)

Table 3. (continued)

Author, year	Modelling approach	Level of data integration of DTs	Clinical input data	Validation strategy	Evaluation metrics	Key findings
Danesh, 2024 ³⁹	Dynamic Ensemble machine learning	DM	V/S, glucose, and SPO2	Internal validation (80% training, 20% testing, 5 hold-out validation)	Accuracy 0.944, precision 0.945, recall 0.942, F1-score 0.943, AUC 0.978	Within a DT dynamic ensemble learning framework for sepsis prediction, the KNOP model outperformed traditional ensembles by providing accurate and medically interpretable predictions
Anyene, 2024 ⁴⁰	DT + DES	DT	Virtual data from Jump Simulation Center – facilities, human & robot agents, motion & task parameters	Simulation-based validation (Functional validation via sensitivity analysis)	Interpretation of quantitative evaluation results	Virtual reality-based DT and DES simulated ICU patient-care scenarios to quantify staff and robot workload and availability, showing that team composition (nurse, robot, HCT) had significant nonlinear effects on productivity
Cannon, 2024 ⁴²	Three-compartment ODE	DM	Porcine pre-clinical data and human clinical data; V/S, lab data, injury severity scores, intervention scores, clinical outcomes	Internal validation (sixteen pigs: 12 fitting, 4 verification, 68 patients: 35 fitting, 17 verifications, 16 hold-out validation)	Time-of-death prediction error (0.67± 0.82 min), statistically significant differences in AUC Damage (p=0.009)	A DT model accurately predicted physiological responses and outcomes in haemorrhagic shock, demonstrating that early plasma plus red blood cell resuscitation improved survival and that the timing of early infusions critically influenced outcomes
Rovati, 2024 ⁴²	Expert rule-based, physiology-driven	DM	70 total expert rules from literature and current clinical practice guidelines	NR (pilot usability test)	SUS: 70 (IQR 62.5–82.5), NASA-TLX 29.2 (IQR 22.5–34.2)	The DT application demonstrated good usability and low-to-moderate workload scores, while residents expressed interest in using it during ICU rotations and identified areas for improvement

(continued)

Table 3. (continued)

Author, year	Modelling approach	Level of data integration of DTs	Clinical input data	Validation strategy	Evaluation metrics	Key findings
Weaver, 2024 ⁴¹	Mechanistic computational models	DM	Respiratory data, ventilation/oxygen settings	Clinical validation (comparison with real data)	MAPE: 0.44~10.02%	DTs served as a methodological framework and a simulation tool based on the physiological data of patients, enabling simulations to elucidate the mechanisms of NIV success and failure in AHRF. The model can noninvasively estimate patient-specific spontaneous breathing effort, showing a strong correlation with Eadi in NAVA and a better association with tidal volume in PSV.
Sun, 2024 ⁴³	Bouc-Wen DT with b-spline effort estimator	DM	Airway pressure, tidal volume, airway flow, PEEP, Eadi	Internal validation (patient level)	Median R ² : 0.55 (QR: 0.40–0.70)	The case study demonstrated a scalable, reliable, and fault-tolerant architecture for a graph-based patient simulation, validating its viability by modelling patient trajectories. DT accurately predicted patient responses to HFNC flow adjustments: responders demonstrated improved lung compliance and reduced lung injury risk, while non-responders exhibited deteriorating lung mechanics and increased injury risk.
Trevena, 2024 ⁴⁴	Expert rules, Graph-based model using DAG	DM	Expert rules	Scenario-based simulation validation	Qualitative assessment of rule-based decision outputs	
Shamoh ammadi, 2025 ⁴⁶	Multi-compartmental mechanistic cardiopulmonary simulator	DM	Demographics, fraction of inspired oxygen, and HFNC flow rate	Clinical validation (comparison with real data)	MAPE: 1.42%–7.76%	

(continued)

Table 3. (continued)

Author, year	Modelling approach	Level of data integration of DTs	Clinical input data	Validation strategy	Evaluation metrics	Key findings
Nair, 2025 ⁴⁷	LSTM-based time-series models + AI decision model	DT	Time series infusion rates, blood pressure values	Internal validation (5-fold)	R^2 score: 0.96–0.97; MAPE: 5.2%–5.3%	The AI-enhanced closed-loop infusion system with cluster-specific decision models and a cardiac DT outperformed conventional models, providing safer and more personalised noradrenaline control with reduced clinician workload
Umo ^a , 2025 ⁴⁵	Hybrid experimental-computational modelling	DM	HR, SVR, PVR, preoperative haemodynamics, baseline speed, in-vitro flow/pressure signals	Case-based clinical validation	Validation residuals: 0-44.7%	PSCOPE predictions accurately indicated the absence of right ventricular failure and captured key haemodynamic trends, although they were less accurate in predicting pulsatile parameters like pulse pressures.

^aPreprint.

ABMS: agent-based modelling and simulation; ABS: agent-based simulation; AHRF: acute hypoxaemic respiratory failure; AI: artificial intelligence; AUROC: area under the receiver operating characteristic curve; β -VAE: beta variational autoencoder; Ci: confidence interval; CV: cross validation; DAG: directed acyclic graph; DES: discrete-event simulation; DM: digital model; DS: digital shadow; DSS: Decision Support System; DT: digital twin; EAdi: electrical activity of the diaphragm; GCS: Glasgow Coma Scale; HCT: healthcare technician/assistant; HFNC: high flow nasal cannula; HR: heart rate; ICU: intensive care unit; IQR: interquartile range; LSTM: long short-term memory; MAPE: mean absolute percentage error; MLP: multi-layer perceptron; MW: mechanical ventilation; NASA-TLX: NASA Task Load Index; NAVA: neurally adjusted ventilatory assist; NIV: non-invasive ventilation; NR: not reported; ODE: Ordinary Differential Equations; PEEP: positive end-expiratory pressure; PSCOPE: Physiology Simulation Coupled Experiment; PSV: pressure support ventilation; PVR: pulmonary vascular resistance; RCT: randomised controlled trial; RNN: recurrent neural network; SOFA: Sequential Organ Failure Assessment; SpO₂: peripheral oxygen saturation; SUS: System Usability Scale; SVR: systemic vascular resistance; V/S: vital signs.

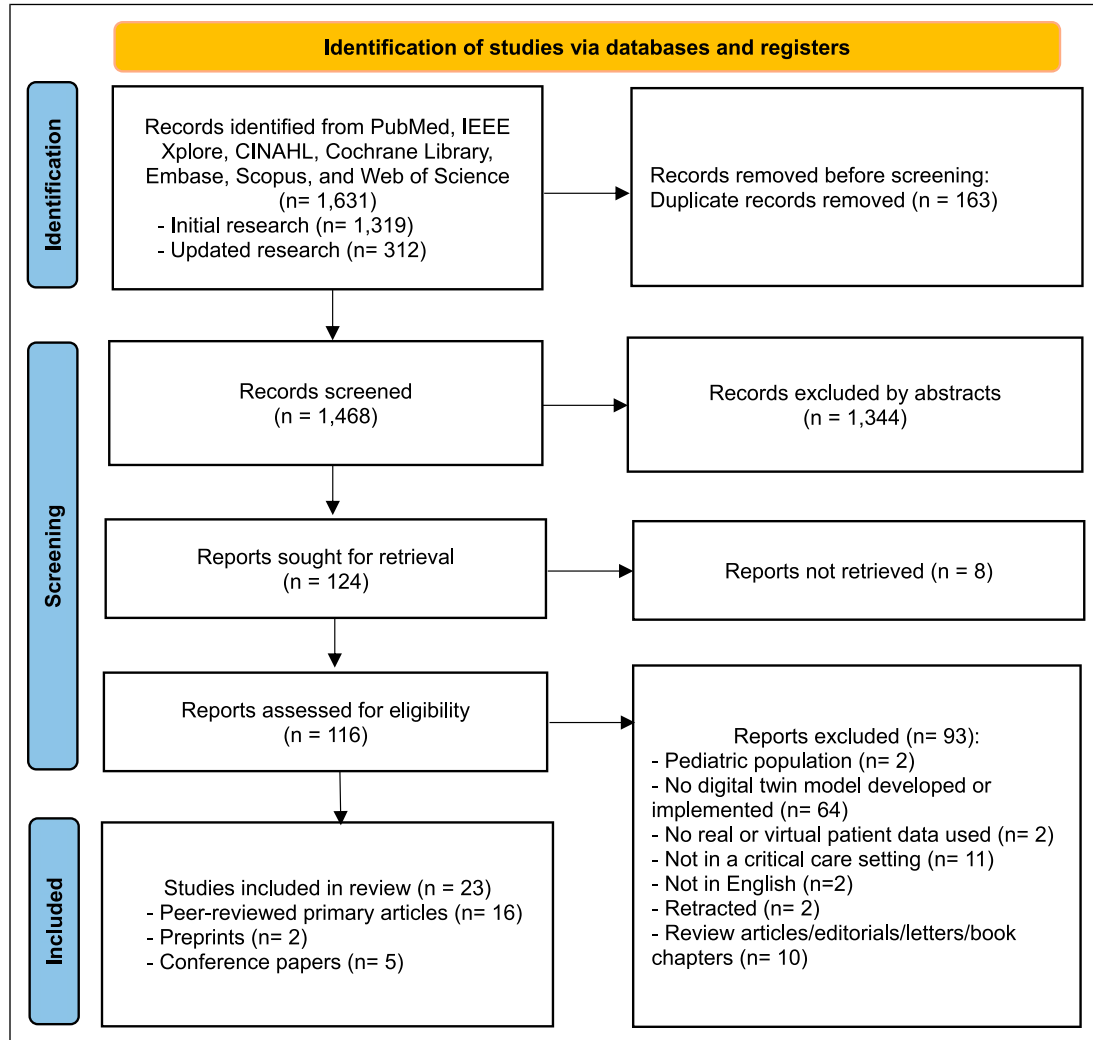


Figure 1. Preferred reporting for systematic reviews and meta-analysis (PRISMA) flow diagram.

Simulation-based approaches mostly represented care delivery processes (3/18, 16.7%),^{18,35,40} and expert rule-based approaches were typically applied in educational or scenario-based simulations (2/18, 11.1%).^{42,44} One study employed mixed methods: a hybrid causal AI model that integrates expert rule-based Bayesian networks with two simulation-based models.³²

Most studies utilised EHR-derived clinical inputs, such as demographic information, laboratory results, and vital signs. A few articles incorporated clinical severity scores.^{32,36} Research targeting patients requiring mechanical ventilation or high-flow nasal cannula therapy used respiratory device parameters, such as ventilator and flow-rate data, as model inputs.^{9,34,36,37,41,43,46} Two studies modelled the entire ICU environment using hospitalisation records, clinical task types and durations, and information on capacity and resource utilisation.^{18,40}

No studies reported external validation using independent datasets. A small number conducted internal validation procedures such as splitting the datasets into training and validation subsets.^{33,36,39,47} Under these internal validation settings, classification models reported high Area Under the Receiver Operating Characteristic Curve (AUROC) values, and while regression-based models demonstrated low prediction errors, including mean absolute percentage error (MAPE) values approximating 5%.

However, several studies conducted clinical or system-level validation, comparing model outputs with real patient data to assess clinical plausibility and performance. In these cases, validation was typically involved by comparing simulated outputs against observed clinical parameters. To evaluate model performance, various quantitative metrics were reported depending on the modelling objective. Models predicting continuous clinical or operational variables presented prediction accuracy using measures such as MAPE,^{41,46} coefficient of determination (R^2),^{34,43,49} or confidence intervals to demonstrate

no statistically significant differences between simulated and observed values.¹⁸ Agreement-based metrics such as kappa coefficients were also used in one study.³²

Despite these validation efforts and comparisons with real-world data, most systems did not demonstrate automated real-time data exchange or bidirectional integration with clinical infrastructures and were therefore classified as DM.

Among the studies that achieved DS or DT levels of integration, implementation depth varied substantially. Patel et al. (2024)³⁸ was classified as a DS despite conducting a prospective RCT-based clinical validation, as the decision support system operated in a unidirectional manner within a human-in-the-loop framework. In contrast, studies classified as DT, including those by Anyene et al. (2024)⁴⁰ and Nair et al. (2025),⁴⁷ demonstrated bidirectional or closed-loop architectures. However, these studies were conducted within simulation environments rather than fully integrated into routine real-time clinical workflows.

Discussion

Principal findings

This scoping review synthesised the current landscape of DT applications in adult critical care by examining study characteristics, clinical context, and DT modelling and evaluation features. Across the 18 included studies, the findings suggest that DT research in critical care remains at an early developmental stage, often characterised by small sample sizes, single-centre designs, and an absence of external validation.

Because DTs represent a relatively novel concept, no studies reporting the development and evaluation of functional DT models in critical care were identified before 2020. Since then, an upward trend has been observed, with more than 60% of studies published after 2024, as reported in this review. This surge likely reflects a growing interdisciplinary interest in digital health innovation and biomedical simulation research. This momentum has been further accelerated by the coronavirus disease 2019 (COVID-19) pandemic, during which DTs were applied to simulate disease spread, optimise healthcare operations, evaluate treatment effects, and support drug development.⁵⁰ Although most studies originated from high-income regions, particularly the United States and Europe, the continued expansion of DT research across diverse healthcare contexts will be important to ensure broader applicability and equitable adoption.

In this expanding field of research, a substantial proportion of the included studies developed DT frameworks that utilised patient physiological data to simulate lung mechanics, predict treatment outcomes, and explore patient-specific strategies. Notably, mechanical ventilation is particularly well-suited for quantitative physiological modelling owing to continuous data collection and real-time monitoring. Recently, efforts have been made to integrate deep learning technologies to enhance decision support and implement automated decision-making in complex ventilator management.³⁶ This respiratory focus contrasts with previous DT reviews, where most studies addressed specific organs, such as the heart, bones, and joints, or broader biological systems, such as the endocrine and immune systems.⁵¹ While research on non-respiratory conditions, such as sepsis, stroke, and cardiovascular dysfunction, remains limited, future DT studies should increasingly encompass these groups to better reflect critical care diversity. The acquisition of high-quality, real-time clinical data is therefore important, as data sources and technological tools shape the methodological approaches used in DT development.

Physicians were the primary end users, predominantly in a clinical decision support context. Although nursing staff were involved in the implementation and training process in one prospective study, the system itself remained largely physician-oriented.³⁸ Another study examining infusion rate optimisation demonstrated the potential relevance of DT tools for nursing-related medication management tasks, suggesting that certain applications may extend beyond physician-centred use.⁴⁷ At the organisational-level, DT applications demonstrated potential for workflow optimisation, staffing allocation, and resource management, as illustrated by simulation-based studies such as nurse–robot collaboration.⁴⁰ To be effectively integrated into practice, DT interfaces and workflows must be tailored to the distinct cognitive demands and professional accountabilities of each user group. Bedside physicians require real-time synchronisation and actionable physiological insights for high-stakes decisions, whereas organisational decision-makers prioritise aggregated data simulations for resource management. For trainees, DTs function as high-fidelity, risk-free environments that enable the exploration of ‘what-if’ scenarios, bridging theoretical knowledge and practical clinical competence.⁶ Given the early stage of DT development in healthcare, further research is needed to clarify how role-specific workflow integration and accountability structures should be operationalised across multidisciplinary ICU teams, spanning three broad functional domains: predictive diagnostics, patient-specific treatment simulation, and organisational-level optimisation.

Several DTs focused on predictive tasks, such as forecasting physiological deterioration or estimating response to therapy using retrospective data. Others implemented patient-specific simulation frameworks to model ventilation strategies, haemodynamic management, or glycaemic control, primarily within in-silico or controlled prospective settings. A smaller subset of studies extended beyond bedside decision support toward hospital-level optimisation, including workflow

modelling and resource allocation simulations. This functional categorisation indicates that, although most current implementations remain model-centric and limited in real-world validation, DT research in critical care spans both individual patient management and broader organisational decision-making contexts.

In this scoping review, DT studies were categorised into mathematical (mechanistic), data-driven ML, simulation-based, and expert rule-based models. Mechanistic models offer physiological interpretability and precise representation of patient-specific states but are limited in capturing highly dynamic, non-linear systems such as acute haemorrhage.¹² In contrast, ML approaches accommodate complex pattern recognition and adaptive modelling, yet often lack transparency and depend on large-scale data. Although explainability techniques (e.g., SHAP) have been applied in one study,³⁹ limited interpretability remains a barrier to clinical trust in high-stakes environments. For advanced DT applications, real-time adaptive control via online sequential learning with multi-objective model selection has been shown to reduce control error in streaming environments,⁵² aligning with the core DT requirement of continuous synchronisation.

The persistent challenge of “data sparsity” in healthcare must be addressed. Similar to strategies in the engineering domain—where scarce fault data are supplemented by finite element simulations⁵³—mechanistic simulation models can generate high-fidelity synthetic patient trajectories, providing robust in-silico datasets to mitigate data sparsity and high-dimensional complexity in ML model training.³⁵ Together, these trade-offs indicate that integrating mechanistic insight with data-driven adaptability may represent a more viable pathway for advancing DT implementation.

Despite these technical advances, the translation of DT models into clinically reliable systems remains constrained by fundamental study design limitations. Most DT studies in critical care have been conducted in single-centre settings with small sample sizes, reflecting the time and resource-intensive nature of developing, debugging, and validating these models.^{32,34,41,43,47} Although several studies have used large-scale public datasets or institutional cohorts to enhance model robustness,^{18,33,36,39} the limited variable diversity of these datasets hinders their ability to capture the complexity of real-world critical care. The occurrence of ‘not reported’ or ‘not applicable’ items regarding data provenance and study settings raises concerns about reproducibility and external validity, as unclear description limits the ability to replicate findings and assess their generalisability. This highlights the need for greater methodological transparency and consistency in reporting practices in DT research.

Across the included studies, validation was predominantly limited to internal approaches such as split-sample or cross-validation, while no study reported independent external validation using separate datasets. In several cases, simulation outputs were compared with retrospective clinical data to assess internal consistency and calibration; however, such strategies do not establish real-world reliability. Notably, model performance declined when applied to external hospitals,⁵⁴ underscoring the need for multi-institutional collaboration, and prospective validation under routine clinical conditions to enhance generalisability.

To characterise implementation maturity studies were positioned along a continuum of data integration and clinical embedding. At the lowest level, several studies functioned as DM, evaluated solely in simulation environments without direct patient data integration. Others incorporated retrospective EHR data, representing partial clinical linkage. One emerging preprint described a hardware-integrated simulation platform linking device-level data from an LVAD to a model; although classified as DM in this review, its architecture borders on a DS.⁴⁵ Only one study progressed to prospective human-in-the-loop evaluation, testing a physiological model-based decision support system in a multicentre RCT context.³⁸ Closed-loop bidirectional DT mechanisms were largely confined to simulated settings, and no study reported fully automated, real-time DT deployment integrated into routine critical care practice.

Emerging work presented in conference papers^{48,49,55–57} (see Supplementary File 3, Table S3) suggests ongoing movement toward greater integration and autonomy. Recent simulation-based studies have implemented stochastic modelling of mechanical ventilation protocols within closed-loop in-silico environment,^{48,49} while a prospective study reported clinical validation of a stochastic glycaemic control protocol in ICU patients, demonstrating safe and effective glucose management under controlled clinical conditions.⁵⁵ Although these findings indicate methodological progression, most developments remain confined to simulation-based or controlled prospective evaluation rather than routine automated deployment in clinical practice.

Predictive DT models generally reported high performance; however, heterogeneous reporting and inconsistent calibration metrics limited cross-study comparison. Many relied on surrogate or simulation-based outcomes rather than prospective clinical endpoints, indicating that predictive accuracy alone does not establish clinical readiness. Notably, one prospective study demonstrated both physiological improvements and clinician adherence of approximately 60%,³⁸ highlighting that usability and workflow integration are as critical as predictive performance for real-world implementation.

From an implementation perspective, most DT systems operated in batch or retrospective modes rather than true real-time processing, with limited reporting of run-time performance, computational scalability, or interoperability with existing ICU monitors and EHR systems. Only a small number provided any indication of processing time,^{47,58} which is worth noting given the time-critical nature of ICU decision-making. Human-in-the-loop configurations were more common than fully

automated closed-loop systems, reflecting technical, infrastructural, and governance-related constraints that limit routine clinical integration.

Despite current limitations, DT technology holds translational potential to enhance safer and more precise clinical care.^{6,7} In critical care, life-support management and weaning follow established protocols.⁵⁹ Integrating patient-specific, real-time physiological and device-derived data into simulation-based decision support could augment these protocols during weaning. Progression from controlled validation to embedded workflow integration and prospective real-world evaluation will be critical for routine DT deployment.

Challenges and future directions

These findings indicate that integrating DTs into ICU settings remains a challenging endeavour, requiring concerted efforts to address persistent challenges in data quality, interoperability, and ethical governance. First, the reliance on small sample sizes and single-centre datasets, which constrain model performance and generalisability, was a commonly identified limitation across the included studies. Second, the lack of standardised data structures and interoperability across health systems remains a fundamental barrier to real-time data integration and scalable DT development. Third, ethical concerns surrounding patient privacy, informed consent, and data ownership necessitate robust governance frameworks and adherence to regulatory standards.

Despite these challenges, ICU-specific DTs may offer a structured pathway toward clinical integration. A staged roadmap may support ICU-specific DT integration. For example, development may progress from conceptual modelling of high-frequency physiological signals and pharmacodynamic responses (Stage 1), to virtual patient simulation and counterfactual analysis (Stage 2), bedside shadow-mode prediction (Stage 3), clinician-in-the-loop decision support (Stage 4), and ultimately tightly bounded automation under predefined safety constraints (Stage 5).

Across these stages, ICU-specific challenges such as handling non-stationary high-frequency data, ensuring safe clinician oversight of automation, and maintaining real-time reliability must be addressed.

Limitations and strengths

This study had some limitations. First, the search was restricted to literature published in English, potentially excluding relevant works published in other languages. Second, only critically ill adult populations were included, limiting generalisability to neonatal, paediatric, or non-critical care contexts. Third, although the search strategy focused on the term “digital twin”, studies describing DT-like systems under alternative terminology may not have been fully captured. Fourth, although preprints and conference proceedings were included to capture emerging evidence, the latter were summarised in a supplementary table rather than subjected to full analytical synthesis, given their preliminary nature and absence of full peer review. Lastly, as a scoping review, this study did not formally appraise methodological quality or risk of bias and therefore does not provide comparative judgements regarding study robustness.

Nonetheless, this study had several strengths. To our knowledge, this represents the first scoping review to comprehensively summarise DT applications in critical care, encompassing both patient-centred applications for critically ill populations and system-level implementations for care delivery optimisation. By synthesising diverse methodological approaches and clinical implementation strategies, this review provides a comprehensive overview of the current state of DT research in critical care and identifies key areas that require further empirical investigation.

Conclusions

This scoping review represents the first scoping review to systematically summarise DT applications in adult critical care. Across the 18 studies, DT research remained at an early developmental stage, with most relying on retrospective or single-centre data and lacking external, real-world validation. Nevertheless, this review identified potential for DTs across three key areas: enhancing clinical prediction and personalised treatment for patients, supporting decision-making and education for clinicians, and optimising operational efficiency for healthcare organisations. Although current applications are primarily targeted towards physicians, emerging work suggests opportunities for broader implementation across multidisciplinary healthcare teams. As the field matures, multicentre collaborations, expanded clinical use cases, strengthened data interoperability, and robust ethical governance will be necessary to facilitate the progression of DT systems from simulation-based models toward clinically evaluated tools.

Acknowledgments

The authors acknowledge the research assistants YJ Choi and MS Lee, who helped organise and visualise the selected literature.

ORCID iDs

Yeonwoo Kim  <https://orcid.org/0009-0007-1904-2250>

Jiin Kim  <https://orcid.org/0009-0005-7050-8955>

Yeonju Kim  <https://orcid.org/0009-0001-6797-6716>

Mona Choi  <https://orcid.org/0000-0003-4694-0359>

Ethical considerations

This article did not require ethical board approval because it did not contain human or animal trials.

Author contributions

All authors contributed to the study's conceptualization and design. Yeonw K and Yeonj K performed the literature search, and Yeonw K and JK completed study selection and data extraction. Yeonw K, JK, Yeonj K conducted the formal analysis. All authors contributed to the interpretation of findings. Yeonw K and JK drafted the initial manuscript. Yeonj K and MC provided critical comments and substantive editorial feedback to the draft manuscript. All authors reviewed and approved the final manuscript. MC provided supervision and funding acquisition.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (Ministry of Science and ICT) (No. RS-2022-NR069414) and the Brain Korea 21 FOUR Project funded by the National Research Foundation (NRF) of Korea, Yonsei University College of Nursing.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and publication of this article.

Data Availability Statement

The dataset used and analysed for the present study is available upon reasonable request.

Guarantor

MC.

Supplemental material

Supplemental material for this article is available online.

References

1. Jones D, Snider C, Nassehi A, et al. Characterising the digital twin: a systematic literature review. *CIRP J Manuf Sci Technol* 2020; 29: 36–52. <https://doi.org/10.1016/j.cirpj.2020.02.002>
2. Piasecik B, Vickers J, Lowry D, et al. TA 12: Materials, structures, mechanical systems, and manufacturing roadmap. NASA, 2012.
3. Glaessgen E and Stargel D. The digital twin paradigm for future NASA and US air force vehicles. In: *53rd AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics and materials conference 20th AIAA/ASME/AHS adaptive structures conference 14th AIAA*, 2012, pp. 1818.
4. Grieves MW. Product lifecycle management: the new paradigm for enterprises. *Int J Prod Dev* 2005; 2: 71–84. <https://doi.org/10.1504/IJPD.2005.006669>
5. Grieves M. Digital Twin: Manufacturing Excellence through Virtual Factory Replication. *White Pap*, 2014. <https://doi.org/10.5281/zenodo.1493930>
6. Katsoulakis E, Wang Q, Wu H, et al. Digital twins for health: a scoping review. *NPJ Digit Med* 2024; 7: 77. <https://doi.org/10.1038/s41746-024-01073-0>
7. Ringeval M, Etindele Sosso FA, Cousineau M, et al. Advancing health care with digital twins: meta-review of applications and implementation challenges. *J Med Internet Res* 2025; 27: e69544. <https://doi.org/10.2196/69544>

8. Khan S, Arslan T and Ratnarajah T. Digital twin perspective of fourth industrial and healthcare revolution. *IEEE Access* 2022; 10: 25732–25754. <https://doi.org/10.1109/ACCESS.2022.3156062>
9. Shamanna P, Erukulapati RS, Shukla A, et al. One-year outcomes of a digital twin intervention for type 2 diabetes: a retrospective real-world study. *Sci Rep* 2024; 14: 25478. <https://doi.org/10.1038/s41598-024-76584-7>
10. Singh M, Srivastava R, Fuenmayor E, et al. Applications of digital twin across industries: a review. *Appl Sci* 2022; 12: 5727. <https://doi.org/10.3390/app12115727>
11. Vincent JL. The continuum of critical care. *Critical Care* 2019; 23: 122. <https://doi.org/10.1186/s13054-019-2393-x>
12. Cannon JW, Gruen DS, Zamora R, et al. Digital twin mathematical models suggest individualized hemorrhagic shock resuscitation strategies. *Commun Med (Lond)* 2024; 4: 113. <https://doi.org/10.1038/s43856-024-00535-6>
13. Gopalan PD and Pershad S. Decision-making in ICU – a systematic review of factors considered important by ICU clinician decision makers with regard to ICU triage decisions. *J Crit Care* 2019; 50: 99–110. <https://doi.org/10.1016/j.jcrc.2018.11.027>
14. James FR, Power N and Laha S. Decision-making in intensive care medicine - a review. *J Intensive Care Soc* 2018; 19: 247–258. <https://doi.org/10.1177/1751143717746566>
15. Vincent JL. The coming era of precision medicine for intensive care. *Critical Care* 2017; 21: 314. <https://doi.org/10.1186/s13054-017-1910-z>
16. Maslove DM, Tang B, Shankar-Hari M, et al. Redefining critical illness. *Nat Med* 2022; 28: 1141–1148. <https://doi.org/10.1038/s41591-022-01843-x>
17. Chase JG, Zhou C, Knopp JL, et al. Digital twins in critical care: what, when, how, where, why? *IFAC Pap OnLine* 2021; 54: 310–315. <https://doi.org/10.1016/j.ifacol.2021.10.274>
18. Zhong X, Babaie Sarijaloo F, Prakash A, et al. A multidisciplinary approach to the development of digital twin models of critical care delivery in intensive care units. *Int J Prod Res* 2022; 60: 4197–4213. <https://doi.org/10.1080/00207543.2021.2022235>
19. Bahrami F, Rossi RM, De Nys K, et al. An individualized digital twin of a patient for transdermal fentanyl therapy for chronic pain management. *Drug Deliv Transl Res* 2023; 13: 2272–2285. <https://doi.org/10.1007/s13346-023-01305-y>
20. Appuhamilage GDKK, Hussain M, Zaman M, et al. DT4CCU—A Digital Twins framework for Critical Care Unit. *Research Square [Preprint]*, 2024. <https://doi.org/10.21203/rs.3.rs-5010353/v1>
21. Tortora M, Pacchiano F, Ferraciolli SF, et al. Medical digital twin: a review on technical principles and clinical applications. *J Clin Med* 2025; 14: 324. <https://doi.org/10.3390/jcm14020324>
22. NSF. Foundations for Digital Twins as Catalyzers of Biomedical Technological Innovation (FDT-BioTech). *National Science Foundation [Internet]*, 2024. NSF; Alexandria (VA). <https://www.nsf.gov/funding/opportunities/fdt-biotech-foundations-digital-twins-catalyzers-biomedical>
23. Fuller A, Fan Z, Day C, et al. Digital twin: enabling technologies, challenges and open research. *IEEE Access* 2020; 8: 108952–108971. <https://doi.org/10.1109/ACCESS.2020.2998358>
24. Kritzinger W, Karner M, Traar G, et al. Digital twin in manufacturing: a categorical literature review and classification. *IFAC Pap OnLine* 2018; 51: 1016–1022. <https://doi.org/10.1016/j.ifacol.2018.08.474>
25. Shen M-d, Chen S-b and Ding X-d. The effectiveness of digital twins in promoting precision health across the entire population: a systematic review. *NPJ Digit Med* 2024; 7: 145. <https://doi.org/10.1038/s41746-024-01146-0>
26. Halpern GA, Nemet M, Gowda DM, et al. Advances and utility of digital twins in critical care and acute care medicine: a narrative review. *J Yeungnam Med Sci* 2024; 42: 49. <https://doi.org/10.12701/jyms.2024.01053>
27. Peters MDJ, Marnie C, Tricco AC, et al. Updated methodological guidance for the conduct of scoping reviews. *JBI Evid Synth* 2020; 18: 2119–2126. <https://doi.org/10.11124/JBIES-20-00167>
28. Peters MDJ, Godfrey C, McInerney P, et al. Scoping reviews (2020). In: Aromataris E, Lockwood C, Porritt K, et al. (eds). *JBI Manual for Evidence Synthesis*. 2024, JBI, <https://doi.org/10.46658/JBIMES-24-09>.
29. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med* 2018; 169: 467–473. <https://doi.org/10.7326/M18-0850>
30. Whitfield S and Hofmann MA. Elicit: AI literature review research assistant. *Public Services Quarterly* 2023; 19: 201–207. <https://doi.org/10.1080/15228959.2023.2224125>
31. Kim Y, Kim Y, Kim J, et al. Nursing surveillance for clinical deterioration among intensive care unit patients: a scoping review. *Intensive Crit Care Nurs* 2026; 92: 104218. <https://doi.org/10.1016/j.iccn.2025.104218>
32. Lal A, Li G, Cubro E, et al. Development and verification of a digital twin patient model to predict specific treatment response during the first 24 hours of sepsis. *Crit Care Explor* 2020; 2: e0249. <https://doi.org/10.1097/ccc.0000000000000249>
33. Allen A, Siefkas A, Pellegrini E, et al. A digital twins machine learning model for forecasting disease progression in stroke patients. *Applied Sciences (Switzerland)* 2021; 11: 5576. <https://doi.org/10.3390/app11125576>
34. Zhou C, Chase JG, Knopp J, et al. Virtual patients for mechanical ventilation in the intensive care unit. *Comput Methods Programs Biomed* 2021; 199: 105912. <https://doi.org/10.1016/j.cmpb.2020.105912>

35. Cockrell C, Schobel-McHugh S, Lisboa F, et al. Generating synthetic data with a mechanism-based critical illness digital twin: demonstration for post traumatic acute respiratory distress syndrome. *bioRxiv [Preprint]* 2023; <https://doi.org/10.1101/2022.11.22.517524>
36. Chakshu NK and Nithiarasu P. An AI based digital-twin for prioritising pneumonia patient treatment. *Proc Inst Mech Eng H* 2022; 236: 1662–1674. <https://doi.org/10.1177/09544119221123431>
37. Ang CYS, Lee JWW, Chiew YS, et al. Virtual patient framework for the testing of mechanical ventilation airway pressure and flow settings protocol. *Comput Methods Programs Biomed* 2022; 226: 107146. <https://doi.org/10.1016/j.cmpb.2022.107146>
38. Patel BV, Mumby S, Johnson N, et al. A randomized control trial evaluating the advice of a physiological-model/digital twin-based decision support system on mechanical ventilation in patients with acute respiratory distress syndrome. *Front Med (Lausanne)* 2024; 11: 1473629. <https://doi.org/10.3389/fmed.2024.1473629>
39. Danesh A, Juraev F, El-Sappagh S, et al. Integrating digital twin technology with dynamic ensemble learning for sepsis prediction in intensive care units. *J Intell Inf Syst* 2024; 30: 25–59. <https://doi.org/10.13088/jiis.2024.30.2.025>
40. Anyene G, Schultz C, Nepomuceno A, et al. Digital-twin co-simulation framework to support informed decision in healthcare planning and management. *Simulation* 2024; 101: 361–375. <https://doi.org/10.1177/00375497241283047>
41. Weaver L, Shamohammadi H, Saffaran S, et al. Digital twins of acute hypoxemic respiratory failure patients suggest a mechanistic basis for success and failure of noninvasive ventilation. *Crit Care Med* 2024; 52: e473–e484. <https://doi.org/10.1097/ccm.0000000000006337>
42. Rovati L, Gary PJ, Cubro E, et al. Development and usability testing of a patient digital twin for critical care education: a mixed methods study. *Front Med (Lausanne)* 2023; 10: 1336897. <https://doi.org/10.3389/fmed.2023.1336897>
43. Sun Q, Chase JG, Zhou C, et al. Estimating patient spontaneous breathing effort in mechanical ventilation using a b-splines function approach. *IFAC Journal of Systems and Control* 2024; 28: 100259. <https://doi.org/10.1016/j.ifacsc.2024.100259>
44. Trevena W, Zhong X, Lal A, et al. Model-driven engineering for digital twins: a graph model-based patient simulation application. *Front Physiol* 2024; 15: 1424931. <https://doi.org/10.3389/fphys.2024.1424931>
45. Abraham EU, Brett AW, Kilic A, et al. Clinical validation of the PSCOPE hybrid model prediction of left ventricular assist device implantation hemodynamics: three patient-specific cases. *medRxiv [Preprint]* 2025; <https://doi.org/10.1101/2025.03.10.25323688>
46. Shamohammadi H, Saffaran S, Tonelli R, et al. Digital twins suggest a mechanistic basis for differing responses to increased flow rates during high-flow nasal cannula therapy. *Intensive Care Med Exp* 2025; 13: 66. <https://doi.org/10.1186/s40635-025-00773-5>
47. Nair VS, Niranga GH, Aryalakshmi C, et al. Optimizing inotropic infusion with cluster specific AI decision models and digital twins. *IEEE Access* 2025. <https://doi.org/10.1109/ACCESS.2025.3581969>
48. Ang CYS, Chiew YS, Wang X, et al. Virtual clinical trials for mechanically ventilated respiratory failure patients under volume-controlled ventilation – the challenges of volume-control Protocols. *IFAC Pap OnLine* 2024; 58: 94–99. <https://doi.org/10.1016/j.ifacol.2024.11.018>
49. Ang CYS, Chiew YS, Wang X, et al. Comparative virtual trials: pressure-controlled versus volume-controlled stochastic integrated model-based mechanical ventilation protocols. *IFAC Pap OnLine* 2024; 58: 100–105. <https://doi.org/10.1016/j.ifacol.2024.11.019>
50. Łukaniszyn M, Majka Ł, Grochowicz B, et al. Digital twins generated by artificial intelligence in personalized healthcare. *Appl Sci* 2024; 14: 9404. <https://doi.org/10.3390/app14209404>
51. Drummond D and Gonsard A. Definitions and characteristics of patient digital twins being developed for clinical use: scoping review. *J Med Internet Res* 2024; 26: e58504. <https://doi.org/10.2196/58504>
52. Jin X, He T and Lin Y. Multi-objective model selection algorithm for online sequential ultimate learning machine. *EURASIP J Wirel Commun Netw* 2019; 2019: 156. <https://doi.org/10.1186/s13638-019-1477-2>
53. Liu X, Huang H and Xiang J. A personalized diagnosis method to detect faults in gears using numerical simulation and extreme learning machine. *Knowl Based Syst* 2020; 195: 105653. <https://doi.org/10.1016/j.knosys.2020.105653>
54. Rockenschaub P, Hilbert A, Kossen T, et al. The impact of multi-institution datasets on the generalizability of machine learning prediction models in the ICU. *Crit Care Med* 2024; 52: 1710–1721. <https://doi.org/10.1097/CCM.0000000000006359>
55. Uyttendaele V, Knopp JL, Desai T, et al. Clinical trial validation of the STAR-3D glycemetic control framework. *IFAC Pap OnLine* 2023; 56: 4758–4763. <https://doi.org/10.1016/j.ifacol.2023.10.1239>
56. Caljé-van der Klei T, Sun Q, Zhou C, et al. Functional residual capacity predictions through three personalized basis functions in a virtual patient model for PCV. *IFAC Pap OnLine* 2024; 58: 526–531. <https://doi.org/10.1016/j.ifacol.2024.11.092>
57. Hennigs C, Sauer J, Bigalke A, et al. Towards a digital twin based monitoring tool for ventilated patients. *IFAC Pap OnLine* 2024; 58: 43–48. <https://doi.org/10.1016/j.ifacol.2024.11.009>
58. Alhammad N, Alajlani M, Abd-alrazaq A, et al. Patients’ perspectives on the data confidentiality, privacy, and security of mHealth Apps: systematic review. *J Med Internet Res* 2024; 26: e50715. <https://doi.org/10.2196/50715>
59. Akella P, Voigt LP and Chawla S. To wean or not to wean: a practical patient focused guide to ventilator weaning. *J Intensive Care Med* 2022; 37: 1417–1425. <https://doi.org/10.1177/08850666221095436>