Razavi Int J Med. 2023 November; 11(4): e1289 doi: 10.30483/RIJM.2023.254477.1289

Published online 2023 November **Review Article**

Artificial Intelligence Solutions for Risk Prediction of Healthcare Associated Infections during and after COVID-19 Pandemic: A Systematic Literature Review

Azadeh Saki[1], and Zahra Ebnehoseini[2,*]

[1] Department of Biostatistics, School of Health, Mashhad University of Medical Sciences, Mashhad, Iran
[2] Health Information Technology unit of Economic Health Department, Mashhad University of Medical Sciences, Mashhad, Iran

*** *Corresponding author:*** Zahra Ebnehoseini, Health Information Technology unit of Economic Health Department, Mashhad University of Medical Sciences, Mashhad, Iran. Email: Ebnehoseiniz1@mums.ac.ir

Received 2023 Aguste 26**; Accepted** 2023 October 28.

Abstract

Background: Healthcare-associated infections (HAIs) pose a significant challenge to patient safety and healthcare systems worldwide. These infections, acquired during medical care, can lead to prolonged long hospital stays, increased morbidity and mortality, and substantial healthcare costs. Identifying and managing risk factors associated with HAIs is crucial for effective prevention and control strategies.
Objectives: This study aims to systematically review the application of artificial intelligence (AI) techniques in Healthcare Associated Infections (HAIs).
Methods: A systematic review was performed that follows the Preferred Reporting Items for Systematic Reviews (PRISMA) guidelines. PubMed was used to search for HAI publications with an emphasis on AI that were published during and post-COVID-19 pandemic. The terms "artificial intelligence" and "HAIs" were used to search for the publications.
Results: A total of 29 articles were included in the systematic review. The most commonly studied healthcare-associated infections (HAIs) were ventilator-associated pneumonia (VAP) and hospital-acquired pneumonia (HAP). However, other HAIs such as hospital-acquired bloodstream infections (BSI), urinary tract infections (UTIs), surgical site infections (SSIs), Klebsiella pneumonia bloodstream infections (Kp-BSI), incubator infections, skin infections, central nervous system infections, meningitis, central line-associated bloodstream infections (CLABSIs), and tracheobronchitis were also examined, although to a lesser extent.
Conclusion: By providing a comprehensive overview of the current landscape of AI solutions in HAI research, this review seeks to facilitate knowledge exchange, promote further research collaborations, and ultimately contribute to the development of effective strategies for preventing and managing HAIs.

Keywords: Algorithms, Artificial intelligence, Electronic health records, Infection, Inpatient, Machine learning

1. Background

Hospital-acquired infections (HAIs) pose a major challenge to patient well-being and healthcare systems worldwide (1). These infections, acquired during medical care, can result in lengthy hospital stays, increased morbidity and death, and substantial healthcare costs. Identification and management of risk factors associated with healthcare-associated infections are critical for effective prevention and control strategies (2, 3).

When multiple infections caused by the same pathogen occur in close temporal proximity to each other in hospital, it raises concerns about potential transmission or outbreaks. Hospital epidemiologists in typically take a traditional approach to investigating such situations. This included defining cases, developing a list of potentially affected people, examining medical history for frequent exposures, conducting ecological surveys, and reviewing healthcare practices (4). In recent years, artificial intelligence has become a robust tool in healthcare and revolutionized various aspects of medical practice (5). Machine learning and deep learning are AI techniques that have been promising in predicting and analyzing complex medical outcomes (6). In the context of healthcare-associated infections, AI solutions are able to improve our comprehension of risk factors (7), diagnosis, and death prediction (8), understanding of disease outbreaks (9), staff training, hand hygiene, and environmental cleaning (10).

In this paper we present a thorough analysis of AI solutions for identifying and analyzing risk factors, forecast and mortality predictions related to HAIs. By summarizing and analyzing the existing literature, we aim to assess the current state of AI applications in this field, highlight key findings, and identify potential challenges and future directions. Overall, this review attempts to illuminate the capability of AI in improving risk assessment, prognosis, and mortality prediction for HAIs. Through a critical analysis of existing literature, we hope to identify knowledge gaps, highlight areas for further investigation, and ultimately contribute to the advancement of AI-driven approaches to HAI prevention and control.

2. Objectives

This study aims to systematically review the application of artificial intelligence (AI) techniques in Healthcare Associated Infections (HAIs).

3. Methods

The current systematic review was preformed

according guidelines. The PubMed and Medline, databases were searched for peer-reviewed articles in English using pre specified search terms. The search terms included "cross infection" OR "hospital infection" OR "nosocomial infection" OR "healthcare-associated infection" in combination with terms related to predictive models, artificial intelligence, machine learning, and other relevant topics.

Inclusion criteria for studies were those that focused on describing risk factors, prognosis, or predicting mortality from healthcare-associated infections (HAIs). Studies that fulfilled the inclusion criteria were incorporated in the investigation.

Exclusion criteria included studies that were unable to fulfill the inclusion criteria or whose primary objectives or outcomes related to the implantation of AI programs for HAIs, drug-resistant infections, treatment, or AI solutions for molecular diagnosis or genome detection of microbial pathogens of nosocomial infections.

The search was conducted for articles published after the COVID-19 pandemic (December 2019). The search was conducted through 14, 2023, the last date for which sources were searched.

The search strategy was as follows:

"cross infection" OR "hospital infection" OR "nosocomial infection" OR "healthcare associated infection") AND ("prediction model" OR "artificial intelligent*"OR "artificial intelligence" OR "artificial learning" OR "deep learning OR "learning" OR "machine learning" OR "knowledge representation" OR "neural network" OR "automated monitoring system" OR "probabilistic network*" OR "statistical learning" OR "support vector machine*" OR "generalized linear model*" OR "naive Bayes*" OR "ensemble method*" OR "decision tree*").

3.1. Eligibility screening and data collection

Data were collected by the original author (AS) and monitored by a second author (ZE) using a spreadsheet file. A random sample of 10 included papers (one-third of the total included papers) that developing AI in HIAs was selected to test data extraction using a separate spreadsheet file.

3.2. Data synthesis

The data extraction checklist was modified as need during the study. Data extraction was performed for the included paper: Journals ranking based on their SJR (Scientific Journal Rankings), authors, publication year, country, SJR ranking, study period, source of data, software, hospital complication, AI algorithms, AI algorithms values, and sample size.

4.Results

After conducting a systematic search on the PubMed database on April 7, 2023, a total of 177 papers were identified. These papers underwent a rigorous screening and selection process, as shown in Figure 1, with resulted in 29 studies were incorporated into our systemic review.

4.1. Characteristics of the Included Studies

4.1.1. Country

The included studies were conducted in 10 various countries. The majority of studies (n = 11, 38%) were conducted in China (11-21), followed by the United States (n = 7, 24.1%) (10, 22-24). Three studies (10.3%) were conducted in Japan (25-27), while the remaining eight studies were conducted in Denmark, France, Iran, Italy, Naples, the United Kingdom, and Israel (in collaboration with the United States).

4.1.2. Publication years

Regarding publication years, of 29 included studies, 6 (20.7%) were published in 2019 (12, 22, 26-29), 7 (24.1%) in 2020 (10, 21, 23, 25, 30-32), 5 (17.2%) in 2021 (7, 13, 24, 33-35), 7 (24.1%) in 2022 (14, 15, 19, 20, 36, 37), and 4 (13.8%) in 2023 (11, 16-18). This indicates that the most studies were published in recent years, with the highest number of publications in 2022.

4.1.3. Regarding the scientific journal rankings (SJR)

The 29 studies, nine (31.0%) in Q1 ranked journals (7, 12, 21, 23, 27, 30, 33, 34, 36), ten (34.5%) were published in Q2 ranked journals (10, 13, 15, 19, 20, 26, 28, 29, 31, 32), six (20.7%) were published in Q3 ranked journals (16, 18, 22, 24, 25, 35), and one (3.4%) was published in Q4 ranked journal (14). Three studies (10.3%) were not ranked in the SJR ranking. This distribution suggests that a significant proportion of studies were published in journals with high SJR rankings, with the majority falling into the Q1 and Q2 categories.

The range of years in which data were collected in the included studies ranged from 1 to 18 years. Among the total 29 studies, 17.2% had a data collection period of 1 year (15, 26, 29, 31, 32, 35), 13.8% had a study period of 2 years (13, 25, 36), 6.9% had a study period of 3 years (16, 25, 28),

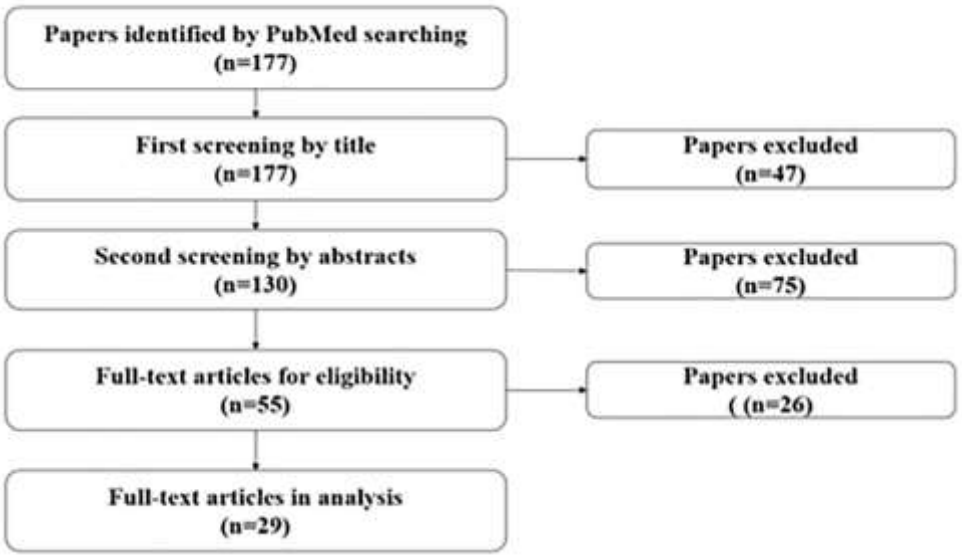

Figure 1. Process of selecting studies

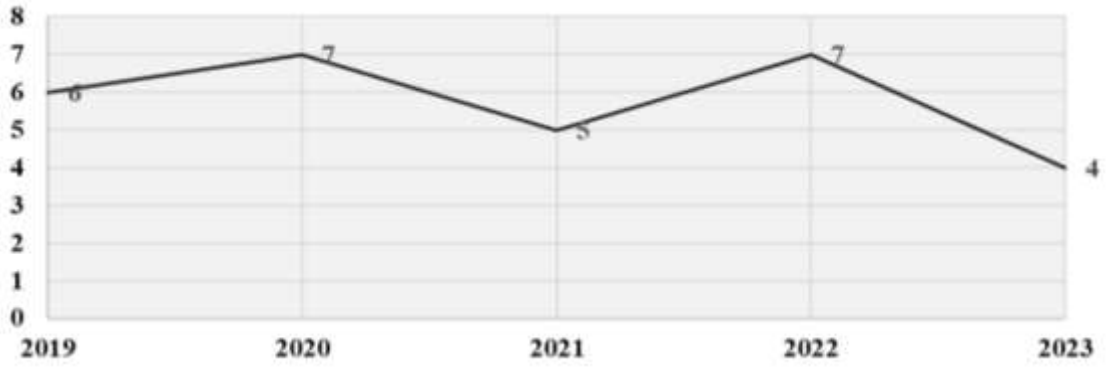

Figure 2. The distribution of the identified studies based on the year of publication

13.8% had a study period of 4 years (11, 14, 19, 23), 13.8% had a study period of 5 years (12, 18, 30), 6.9% had a study period of 6 years (10, 27), 3.4% had a study period of 7 years (20), 3.4% had a study period of 11 years (21), 10.3% had a study period of 12 years (22, 24, 37), 6.9% had a study period of 14 years (33, 34), and 3.4% had a study period of 18 years (7). These results indicate that the study periods varied across the included studies, ranging from 1 to 18 years.

4.2. Study setting

Of included studies, 34.5% (n = 10) were conducted in intensive care units (ICU) (13, 16, 22, 24-26, 28-30, 33). In addition, 17.2% (n = 5) were focused on elderly patients (12, 18-20, 37), and another 17.2% were conducted pediatric patients (17, 21, 31, 34). Some examined patients with burn (14), schizophrenic (27), acute respiratory distress syndrome (23), elective abdominal surgery (11), and trauma patients (32). The remaining studies were performed on all hospitalized patients inpatients (7, 10, 15, 35, 36).

4.3. Sample size and data source

The findings of the present study indicate that the majority of studies (n= 14, 48.3%) relied on data from electronic medical records or hospital information systems to develop artificial intelligence models (7, 10-15, 19, 22, 24, 30, 34, 35, 37). This show the importance of these data sources for research in this field. In addition, 17.2% of studies used data from medical records data to develop predictive models using artificial intelligence techniques (16, 20, 21, 27, 29). In addition, five studies (17.2%) used data from research databases (18, 23, 26, 33, 36), while three studies (10.3%) utilized data from national databases (25, 31, 32). These findings highlight the different data sources used in the studies and the importance of electronic medical records and hospital information systems for research on this topic.

The results show variability in sample sizes among the identified studies, with some having smaller sample sizes (< 200) (13, 16, 17, 21, 25-27), some having medium sample sizes (>= 200 to <1000) (12, 15, 20, 28, 32, 34), and other studies

having larger sample sizes (>= 1000). The range of sample size was from 24 (26) to 897344 (10).

4.4. Validation methods

The most frequently used validation methods in the incorporated studies were, in descending order, were logistic regression models (n=20), random forest (RF) (n=11), support vector machines (SVM) (n=12), decision tree classification (n=12), Multi-Layer Perceptron (MLP) (n=2), calibration plot (n=1), Neural networks (n=10), eXtreme Gradient Boosting [XGBoost) (n=9), decision curve analysis (DCA) (n=3), classification and regression tree (CART) (n=2), Dense Encoder (n=1), k-nearest neighbors (KNN) (n=2), concordance index (c-index) (n=2), clinical impact curve analysis (CICA) (n=1), Ranger Forest Classifier (RFC) (n=2), least absolute shrinkage and selection operator (LASSO) (n=4), Bayes search method (n=1), and finally Deep Averaging Network (DAN) (n=1). These methods were used to analyze and predict infection rates in hospital.

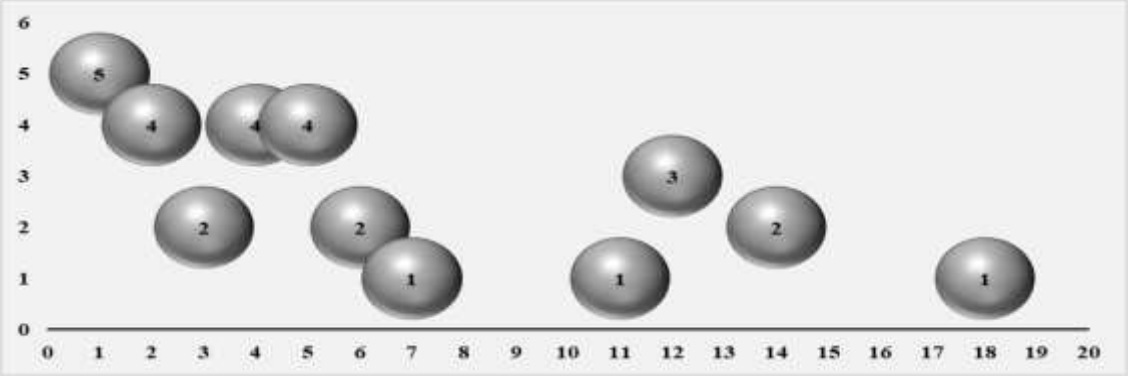

Figure 3. The distribution of the identified studies based on the years of data gathering

4.5. Hospital acquired infection (HAIs)

Seven studies were conducted with the aim of applying machine learning-based risk prediction models to prognosticate the occurrence of nosocomial infections by selected infection types (13, 15, 16, 26, 28, 36), including COVID-19 (two studies) (15, 36), carbapenem resistant Enterobacteriaceae (one study) (13), Clostridium difficile infection (one study) (28), Acinetobacter baumannii infections (one study) (16), and Pseudomonas aeruginosa (one study) (26).

The identified studies focused on the development of artificial intelligent models for different AHI (14, 25, 29, 32, 33, 35). For example, Karajizadeh et al. conducted a study aimed at developing a model to anticipate in-hospital death due to HAIs in trauma patients. The researchers used an unbalanced dataset that differnt types of HAIs such as upper respiratory tract infections, urinary tract infections (UTI), surgical site infections skin infection, bloodstream infection, pneumonia, central nervous system infections, and meningitis (32). Risk prediction models for ventilator-Associated Pneumonia (VAP), hospital-acquired pneumonia (HAP), UTI, surgical site infection (SSI), bloodstream infection (BSI), and tracheobronchitis were developed using random forest (RF), logistic regression, and convolutional neural networks (CNN). The positive predictive value (PPV), and negative predictive value (NPV) (30).

VAP (n=8) (12, 16, 18, 23, 24, 26, 35, 37) and HAP (n= 5) (16, 19, 22, 27, 32) were the most frequent HAIs in the included studies. Wu et al. developed a nomogram to VAP in patients with acute respiratory distress syndrome (ARDS) using decision curve analysis (DCA) calibration plot, and C-index. This nomogram can be used after ICU admission and utilizes readily available variables (23).

Xu et al. investigated the occurrence of VAP and related risk factors in elderly patients receiving mechanical ventilation using a logistic regression model. They developed a model with an ROC of 0.722 (95% CI, 0.679 to 0.765) specifically to

Table 1. The Features of the Included Studies

Row	Authors	Publication year	Country	SJR ranking	Study period	Source of Data	Software
1	Kuo et al. (27)	2019	Japan	Q1	2013 to 2018	PMRs	-[c]
2	Zhang et al. (11)	2023	China	-[a]	2018 to 2021	EHRs	R
3	Zachariah et al. (10)	2020	USA	Q2	2009 to 2014	EHRs	-[c]
4	Wu et al. (23)	2020	USA	Q1	2008 to 2011	Research dataset	R
5	Li et al. (21)	2020	China	Q1	2009 to 2019	PMRs	R
6	Xu et al. (12)	2019	China	Q1	2011 to 2015	EHRs	SPSS
7	Moller et al. (7)	2021	Denmark	Q1	2001 to 2018	EHRs	SAS
8	Li et al. (18)	2023	China	Q3	2015 to 2019	Research data set	SPSS

9	Giang et al. (24)	2021	USA	Q3	2001to 2012	EHRs	-[c]
10	Goodwin and Demner-Fushman (22)	2019	USA	Q3	2001to 2012	EHRs	-[c]
11	Pei et al. (16)	2023	China	Q3	2019 to 2021	PMRs	EpiData \ R
12	Chen et al. (19)	2022	China	Q2	2017 to 2020	EHRs	R
13	Liang et al. (37)	2022	USA	Q2	2001 to 2012	EHRs	Python
14	Karajizadeh et al. (32)	2020	Iran	Q2	2017 to 2018	National database	SPSS/Python
15	Ai et al. (20)	2022	China	-[a]	2016 to 2022	PMRs	R \ Python
16	Myall et al. (36)	2022	UK	Q1	2020 to 2021	Research dataset	R
17	Wang et al. (13)	2021	China	Q2	2018 to 2019	EHRs	-[c]
18	Wang et al. (15)	2022	China	Q2	2020	EHRs	-[c]
19	Barchitta et al. (33)	2021	Italy	Q1	2006 to 2019	Research dataset	SPSS
20	Nistal-Nuño et al. (25)	2020	Japan	Q3	2002 -2004	National database	MATLAB
21	Walker et al. (34)	2021	USA	Q1	2005 to 2018	EHRs	Python
22	Dos Santos et al. (35)	2021	Brazil	Q3	2017	EHRs	Python
23	Tilton and Johnson (28)	2019	USA	Q2	2015 to 2017	-[b]	-[c]
24	Jiang et al. (17)	2023	China	-[a]	2021 to 2022	-[b]	R
25	Rabhi et al. (29)	2019	French	Q2	2009 to 2010	PMRs	-[c]
26	Liao et al. (26)	2019	Japan	Q2	2015	Research dataset	MATLAB
27	Roimi et al. (30)	2020	USA-Israel	Q1	2013 to 2017	EHRs	Python
28	Wang et al. (14)	2022	China	Q4	2016 to 2019	EHRs	R
29	Montella et al. (31)	2020	Naples	Q2	2016 to 2020	National database	Python

-a: Journal not index in Scopus
-b: Dataset was not reported
-c : Software was not reported

predict the happening of VAP and recognize high-risk patients (12).

Li et al. also used a logistic regression model to create a VAP prediction model and constructed a nomogram according to the baseline clinical features of elderly ICU patients on mechanical ventilation. Giang et al. examined the suitability of machine learning methods for predicting of VAP and used logistic regression, multilayer perceptron, random forest, support vector machine, XGBoost, CURB-65, and predisposition methods (24).

Liang et al. compared the use of the random forest technique with the Clinical Pulmonary Infection Score ((CPIS)-based model for VAP prediction. They found that the VAP prediction model had excellent performance and outperformed the CPIS model in accurately predicting VAP (37). Liao et al. collected outperforming data from patients and built Ensemble Neural Network (ENN) and SVM prediction models to foretell whether patients were infected with VAP (26).

Table 2. The application of AI in HAIs

Row	Authors	Hospital complication	AI Algorithms	AI Algorithms values	Sample size
1	Kuo et al. [27]	HAP	CART,C5.0, KNN,NB, RF, SVM, LRG	Train: CART: 0.851, C5.0: 0.971, KNN: 0.696, NB: 0.798, RF: 0.971, SVM: 0.936, LRG: 0.762, Model: CART: 0.880, C5.0: 0.993, KNN: 0.701, NB: 0.831, RF: 0.994, SVM: 0.953, LRG: 0.823	185
2	Zhang et al. [11]	SSIs	Multivariate LRG	LRG Model:0.926	3018
3	Zachariah et al. [10]	UTIs	ANN, DT, LRG	LRG Model:0.63, ANN:0.77, Decision Tree:0.78	897344
4	Wu et al. [23]	HAP, VAP	C-index/ / DCA	Model: 0.744	1000
5	Li et al. [21]	SOFA,Kp-BSI	Multivariate Cox regression	SOFA score: 0.79 Klebsiella pneumoniae (KP) specific SOFA:0.85	146
6	Xu et al. [12]	HAP- VAP	LRG	LRG Model:0.722 (95% CI, 0.679 to 0.765)	901
7	Moller et al. [7]	HA,UTIs	ANN, XBG, Regression, DT.	HA-UTI model: 0 hours after admission=0.82 to 0.84\ 48 hours after admission=0.71 to 0.77	17768
8	Li et al. [18]	VAP	LRG	LRG Training: 0.859 LRG Model: 0.813	1219
9	Giang et al. [24]	VAP	LRG, Multilayer perceptron, RF, SVM, XGB	Prediction Model of VAP in ICU 6 hours LRG: 0.744, Multilayer perceptron: 0.731, RF: 0.771, SVM: 0.765, XGB: 0.799, CURB-65: 0.503, predisposition: 0.565. Prediction of VAP in ICU 48 hours LRG: 0776, Multilayer perceptron: 0.741, RF: 0.777, SVM: 0.775, XGB: 0.791	6126
10	Goodwin and Demner-Fushman [22]	HAP	RNNs, CNNs, DAN, Dense Encoder	Training: RNNs: 0.64, CNNs: 0.65, DAN: 0.64, Dense Encoder: 080, Sparse: 0.68. Model: RNNs: 0.65, CNNs: 0.60, DAN: 0.66, Dense Encoder: 0.55, Sparse: 0.61.	1467
11	Pei et al. [16]	HAP	LASSO regression	The C-index: Training: 0.922 (95% CI: 0.873–0.970) Validation:0.823 (95%CI: 0.706–0.941) 90-day mortality	164

				Training: 0.922 (95% CI: 0.873, 0.971) Validation: 0.823 (95% CI:0.703, 0.943)	
12	Chen et al. [19]	Non-ventilator HAP	C-index, DCA, LASSO	Model: 0.821 Test: 0.813	15420
13	Liang et al. [37]	VAP	RF	: 59% ± 2% Random Forest Model: 84% ± 2%	38515
14	Karajizadeh et al. [32]	URI, UTI, SSI, SKIN, BSI, HAP, CNS, surgery took place	C5.0 tree	C5.0 tree Model: 0.619	549
15	Ai et al. [20]	UTI	RFC, SVM, XGB, ANN, DT	Training: RFC: 0.925 (95% CI, 0.868-0.982), SVM: 0.787 (95% CI, 0.730-0.844), DT: 0.776 (95% CI, 0.719-0.833), ANN: 0.879 (95% CI, 0.822-0.936), XGboost 0.797 (95% CI, 0.740-0.854). Test: RFC 0.918 (95% CI, 0.861-0.975), SVM 0.779 (95% CI, 0.722-0.836), DT 0.769 (95% CI, 0.712-0.826), ANN 0.854 (95% CI, 0.797-0.911),	674
16	Myall et al. [36]	Hospital-onset COVID-19	XGB	XGB Model: 0·89 (95% CI 0·88–0·90)	51157
17	Wang et al. [13]	BSI	Multivariate analysis	LRG Model: 0.921	42
18	Wang et al. [15]	Nosocomial SARS-CoV-2 Infection	LASSO, LRG, SVM, DT, RF, DCA, CICA	Test: 0.863 (95% CI: 0.834–0.892). Model: 0.813 (95% CI: 0.760–0.866)	857
19	Barchitta et al. [33]	HAIs	SVM	Traditional statistical analysis 0.612 (95% CI = 0.60-0.63), SVM with SAPS II 0.90 (95% CI = 0.88-0.91), SVM Without SAPS II along 0.66 (95% CI = 0.65-0.68)	7827
20	Nistal-Nuño et al. [25]	HAIs	ANN	Backpropagation with the for xor.README file values test: 0.0, model: 0.02760293 (MSE) Backpropagation with momentum with the for xor.README file values test: 0.0, model: 0.0 (MSE)	-
21	Walker et al. [34]	CLABSIs	Regularized LR RF, SVM, XGB	Best model: 14- day Infection Recurrence 0.83, 91-day Infection Recurrence 0.77, 14-day CVC Removal 0.66, 365-d CVC Removal 0.76.	969
22	Dos Santos et al. [35]	Pneumonia VAP, UTI, SSI, BSI, Tracheobronchitis	RF, LRG, CNN,	All infections 90.27% (SD ± 0.15)	5105
23	Tilton and Johnson [28]	CDI	LRG	Risk factors for CDI identified and incorporated into the model included age ≥70 years (adjusted odds ratio,1.89;95% confidence interval1.05-3.43; P =.0326) and recent hospitalization in the past 90 days (adjusted oddsratio,3.55;95% confidence interval1.90-6.83; P < .0001).	200
24	Jiang et al. [17]	Incubator infection	XGB, RF, SVM, DT	XG-Boost: 0.93, RF: 0.91, SVM: 0.91, and DT: 0.89	76
25	Rabhi et al. [29]	HAIs	CNN	CNN: 0.98	1531
26	Liao et al. [26]	VAP	ANN/ SVM	ENN AUC=0.9879 SVM AUC=0.9508	24
27	Roimi et al. [30]	BSI	Multivariate LRG	Hospital1 cross-validation 1: 0.87 ± 0.02, Internal validation: 0.89 ± 0.01 Hospital2 cross-validation 2: 0.93 ± 0.03, Internal validation: 0.92 ± 0.02	2351
28	Wang et al. [14]	HAIs	Multivariate LRG	LRG Model: 0.97 (95% , CI: 0.95-0.99)	3475
29	Montella et al. [31]	BSI	SVC, CATBOOST, XGB, RFC, MLP, RF, LR	SVC: 0.5357, CATBOOST: 0.5670, XGB: 0.5313, RFC: 0.5335, MLP: 0.6027, RF: 0.5335, LR: 0.6027	1203

Dos Santos et al. performed healthcare-associated infection surveillance, including VAP as well as pneumonia, UTI, SSI, and BSI, using random forest, logistic regression, CNN, PPV, and NPV methods (35).

Chen et al. tried to create and validate a simple nomogram and dynamic web-based calculator for anticipating the risk of nonventilated hospital-acquired pneumonia (NV-HAP) in elderly hospitalized patients. They used the C-index, decision curve analysis, and least absolute shrinkage and selection operator (LASSO) methods (19).

In another study, a predictive model for hospital-acquired pneumonia in schizophrenic patients was developed using CART, C5.0, KNN, NB, RF, SVM, and logistic regression (LRG). This model can act as a useful tool for clinician physicians curing schizophrenic patients (27).

Goodwin et al. created a deep learning system to foretell the risk of impending pneumonia in the future using clinical findings recorded in ICU files for a risky population (22).

In the study by Pei et al. prognostic factors for HAP and VAP triggered by Acinetobacter baumannii were validated regarding 90-day death in patients with in the respiratory ICU and a predictive nomogram was created to individually prognosticate the probability of 90-day death in patients with HAP and VAP triggered by AB. In the respiratory ICU. This nomogram model showed excellent performance in treating patients with - HAP and VAP caused by Acinetobacter baumannii (16).

We identified four models to predict and mortality s for BSI (13, 30-32, 35). Two studies the used multivariable regression models (13, 30). In Montella's study a prognosticative analysis of healthcare-associated blood stream infections in the neonatal intensive care unit was performed (31).

Other AHI in the identified studies were as follows: UTIs (n=5) (7, 10, 20, 32, 35), SSIs) (n=3) (11, 23, 35), Klebsiella pneumonia bloodstream infection (Kp-BSI) (21), incubator infection (17), skin infection (20), central nervous system infection (20), meningitis (20), CLABSIs (34), and tracheobronchitis (n=1) (35).

5. Discussion

The purpose of the current study was to review and investigate the studies in which machine learning-based risk prediction models were used to predict the occurrence of nosocomial infections caused by different types of infections. The identified studies highlight the capability of machine learning-based risk prediction models to improve the prediction, identification, and management of nosocomial infections. These models are valuable tools for healthcare professionals to identify high-risk patients, take preventive measures, and optimize patient outcomes.

Our results suggest that human-like artificial intelligence is indeed experiencing rapid growth in healthcare, similar to other arears of healthcare domains such as dentistry (38), gastric cancer (39), and nutrition (40). Most studies in HAI have been published in recent years, with 2022 having the highest number of publications. This temporal distribution indicates an increasing interest in HAI and a greater focus on HAI in healthcare. It indicates that researchers and practitioners are recognizing the potential benefits and applications of HAI in improving healthcare outcomes.

Among the studies identified in the systematic review, VAP and HAP were the most commonly studied HAIs. Various machine learning techniques were used to develop predictive models for VAP, such as logistic regression, random forest, SVM, and neural networks. These models showed excellent performance in accurately predicting the occurrence of VAP. in Addition, nomograms were developed as useful tools for identifying risky patients and supporting VAP prediction. Despite, these advancements, the systematic study by Frondelius et al. highlights current limitations in the development and application of these predictive models. The authors emphasize the need for further research and collaboration between clinical studies to effectively translate these tools from the laboratory to healthcare practice, ultimately improving the diagnosis and prognosis of VAP and related outcomes (41).

Healthcare-associated infections such as BSIs, UTIs, SSIs, Klebsiella Kp-BSI, incubator infections, skin infections, central nervous system infections, meningitis, CLABSIs, and tracheobronchitis have also been investigated, although to a lesser extent. These studies focused primarily on ICU, elderly, and neonate settings. Among hospital units, the highest rates of these infections were observed in transplant units, neonatal units, and ICUs. Bacteremia, bloodstream infections, gastrointestinal infection, pneumonia, and respiratory tract infection were the most commonly reported infections in the study conducted by Raoofi et al. (42). These particular areas of focus may have been selected due to the high incidence and death rates related to these infections in hospitals.

The results of our systematic review highlight the variability in sample sizes among the identified studies. Smaller sample sizes may have limited statistical power and are more susceptible to sampling error, while larger sample sizes are generally more representative of the population and provide greater statistical power. It is crucial to take into account the sample size when interpreting the results of each study and evaluating the strength of the findings.

In a related systematic review by Akazawa and Hashimoto on artificial intelligence in gynecologic cancers, they also reported several studies with relatively small data set. This is due to the fact that many studies used datasets from single institutions, resulting in a limited number of included patients. They emphasize the importance of further validation with larger databases to determine the accuracy and reliability of the suggested algorithms. This indicates that studies in this area need to include larger and more diverse datasets to enhance the generalizability and

robustness of the findings (43).

6. Conclusion

The review includes studies that have used AI techniques such as machine learning algorithms, deep neural networks, and other AI methods. We will explore the different types of data sources used, including health data, electronic health records, genetic data, and microbiological data, to train and develop predictive models. In addition, we will investigate the performance and accuracy of these AI models in predicting HAIs and their associated outcomes.

By providing a comprehensive state of art of the current landscape of AI solutions in HAI research, this review aims to facilitate knowledge sharing, foster further research collaborations, and ultimately contribute to the creation of operative strategies for stopping and management HAIs.

Acknowledgments

None

Conflicts of interest

The authors of the article declared no conflict of interest.

References

1. Zingg, W., A. Holmes, M. Dettenkofer, T. Goetting, F. Secci, L. Clack, et al., Hospital organisation, management, and structure for prevention of health-care-associated infection: a systematic review and expert consensus. *Lancet Infect Dis*, 2015. 15(2): p. 212-224. doi: 10.1016/S1473-3099(14)70854-0. [PubMed:25467650).
2. Haque, M., M. Sartelli, J. McKimm and M.A. Bakar, Health care-associated infections–an overview. *Infect Drug Resist*, 2018: p. 2321-2333. doi: 10.2147/IDR.S177247. [PubMed: 6245375).
3. Despotović, A., N. Milić, A. Cirković, B. Milošević, S. Jovanović, V. Mioljević, et al., Incremental costs of hospital-acquired infections in COVID-19 patients in an adult intensive care unit of a tertiary hospital from a low-resource setting. *Antimicrob. Resist. Infect. Control*, 2023. 12(1): p. 39. doi: 10.1186/s13756-023-01240-0. [PubMed: 10120483).
4. Sundermann, A.J., J. Chen, J.K. Miller, M.I. Saul, K.A. Shutt, M.P. Griffith, et al., Outbreak of Pseudomonas aeruginosa Infections from a Contaminated Gastroscope Detected by Whole Genome Sequencing Surveillance. *Clin. Infect. Dis*, 2021. 73(3): p. e638-e642. doi: 10.1093/cid/ciaa1887 . [PubMed: 33367518).
5. Mancini, A., L. Vito, E. Marcelli, M. Piangerelli, R. De Leone, S. Pucciarelli, et al., Machine learning models predicting multidrug resistant urinary tract infections using "DsaaS". *BMC bioinformatics*, 2020. 21(Suppl 10): p. 347. doi: 10.1186/s12859-020-03566-7. [PubMed: 7446147).
6. Ghaffar Nia, N., E. Kaplanoglu and A. Nasab, Evaluation of artificial intelligence techniques in disease diagnosis and prediction. *Discover Artificial Intelligence*, 2023. 3(1): p. 5. doi: 10.1007/s44163-023-00049-5. [PubMed: 9885935).
7. Møller, J.K., M. Sørensen and C. Hardahl, Prediction of risk of acquiring urinary tract infection during hospital stay based on machine-learning: A retrospective cohort study. *PloS one*, 2021. 16(3): p. e0248636. doi: 10.1371/journal.pone.0248636. [PubMed: 8011767).
8. Han, X., W. Wu, H. Zhao and S. Wang, Developing and validating a prediction model for in-hospital mortality in patients with ventilator-associated pneumonia in the ICU. *Ann Palliat Med*, 2022. 11(5): p. 1799-1810. doi: 10.21037/apm-22-502. [PubMed: 35672896).
9. Atkinson, A., B. Ellenberger, V. Piezzi, T. Kaspar, L. Salazar-Vizcaya, O. Endrich, et al., Extending outbreak investigation with machine learning and graph theory: Benefits of new tools with application to a nosocomial outbreak of a multidrug-resistant organism. *Infect Control Hosp Epidemiol*, 2023. 44(2): p. 246-252. doi: 10.1017/ice.2022.66. [PubMed: 36111457).
10. Zachariah, P., E. Sanabria, J. Liu, B. Cohen, D. Yao and E. Larson, Novel Strategies for Predicting Healthcare-Associated Infections at Admission: Implications for Nursing Care. *Nurs Res*, 2020. 69(5): p. 399-403.10.1097. doi: 10.1097/NNR.0000000000000449. [PubMed: 32604154).
11. Zhang, J., F. Xue, S.D. Liu, D. Liu, Y.H. Wu, D. Zhao, et al., Risk factors and prediction model for inpatient surgical site infection after elective abdominal surgery. *World J Gastrointest Surg*, 2023. 15(3): p. 387-397. doi: 10.4240/wjgs.v15.i3.387. [PubMed: 10080607).
12. Xu, Y., C. Lai, G. Xu, W. Meng, J. Zhang, H. Hou, et al., Risk factors of ventilator-associated pneumonia in elderly patients receiving mechanical ventilation. *Clin Interv Aging*, 2019. 14: p. 1027-1038. doi: 10.2147/cia.s197146. [PubMed: 6566835).
13. Wang, Y., Q. Lin, Z. Chen, H. Hou, N. Shen, Z. Wang, et al., Construction of a Risk Prediction Model for Subsequent Bloodstream Infection in Intestinal Carriers of Carbapenem-Resistant Enterobacteriaceae: A Retrospective Study in Hematology Department and Intensive Care Unit. *Infect Drug Resist*, 2021. 14: p. 815-824. doi: 10.2147/idr.s286401 . [PubMed: 7936666).
14. Wang, H., P. Zhao, D. Sun, X. Wu, Q.F. Yuan and K.W. Wang, [Epidemiological characteristics and the establishment and evaluation of a risk prediction model for nosocomial infection in burn patients). *Zhonghua Shao Shang Za Zhi*, 2022. 38(12): p. 1170-1178. doi: 10.3760/cma.j.cn501225-20220214-00025. [PubMed: 36594148).
15. Wang, C., C. Peng, L. Ning, X. Qiu, K. Wu, N. Yang, et al., Development and Validation of a Nomogram to Assist Monitoring Nosocomial SARS-CoV-2 Infection of Hospitalized Patients. *J Inflamm Res*, 2022. 15: p. 1471-1481. doi: 10.2147/jir.s351509.Pmc8896769. [PubMed: 35250295).
16. Pei, Y., Y. Huang, X. Pan, Z. Yao, C. Chen, A. Zhong, et al., Nomogram for predicting 90-day mortality in patients with Acinetobacter baumannii-caused hospital-acquired and ventilator-associated pneumonia in the respiratory intensive care unit. *J Int Med Res*, 2023. 51(3): p. 3000605231161481. doi: 10.1016/s0045-6535(98)00362-2 . [PubMed: 10028662).
17. Jiang, L., J. Ma, F. Li and N. Qin, Association between incubator standards and newborn nosocomial infection with machine-learning prediction. *Transl Pediatr*, 2023. 12(4): p. 655-662. doi: 10.21037/tp-23-171. [PubMed: 10167382).
18. Li, S., L. Shang, L. Yuan, W. Li, H. Kang, W. Zhao, et al., Construction and Validation of a Predictive Model for the Risk of Ventilator-Associated Pneumonia in Elderly ICU Patients. *Can Respir J*, 2023. 2023: p. 7665184. doi: 10.1210/jcem.83.12.5329 . [PubMed: 9851783).
19. Chen, Z., Z. Xu, H. Wu, S. Gao, H. Wang, J. Jiang, et al., Derivation and validation of a nomogram for predicting nonventilator hospital-acquired pneumonia among older hospitalized patients. *BMC Pulm Med*, 2022. 22(1): p. 144.doi: 10.1186/s12890-022-01941-z . [PubMed: 35428276).
20. Ai, J., Y. Hu, F.F. Zhou, Y.X. Liao and T. Yang, Machine learning-assisted ensemble analysis for the prediction of urinary tract infection in elderly patients with ovarian cancer after cytoreductive surgery. *World J Clin Oncol*, 2022. 13(12): p. 967-979. doi: 10.5306/wjco.v13.i12.967 . [PubMed: 36618079).
21. Li, S., J. Liu, F. Chen, K. Cai, J. Tan, W. Xie, et al., A risk score based on pediatric sequential organ failure assessment

predicts 90-day mortality in children with Klebsiella pneumoniae bloodstream infection. *BMC Infect Dis*, 2020. 20(1): p. 916. doi: 10.1186/s12879-020-05644-w . [PubMed: 7709332).
22. Goodwin, T.R. and D. Demner-Fushman, Deep Learning from Incomplete Data: Detecting Imminent Risk of Hospital-acquired Pneumonia in ICU Patients. MIA ... Annual Symposium proceedings / AMIA Symposium, 2019. 2019: p. 467-476. [PubMed: 32308840).
23. Wu, Z., Y. Liu, J. Xu, J. Xie, S. Zhang, L. Huang, et al., A Ventilator-associated Pneumonia Prediction Model in Patients With Acute Respiratory Distress Syndrome. *Clin Infect Dis*, 2020. 71(Suppl 4): p. S400-s408.10.1093/cid/ciaa1518. doi: 10.1093/cid/ciaa1518 . [PubMed: 33367575).
24. Giang, C., J. Calvert, K. Rahmani, G. Barnes, A. Siefkas, A. Green-Saxena, et al., Predicting ventilator-associated pneumonia with machine learning. *Medicine*, 2021. 100(23): p. e26246. doi: 10.1097/MD.0000000000026246. [PubMed: 34115013).
25. Nistal-Nuño, B., A neural network for prediction of risk of nosocomial infection at intensive care units: a didactic preliminary model. Einstein (Sao Paulo, Brazil), 2020. 18: p. eAO5480. doi: 10.31744/einstein_journal/2020AO5480 . [PubMed: 33237246).
26. Liao, Y.H., Z.C. Wang, F.G. Zhang, M.F. Abbod, C.H. Shih and J.S. Shieh, Machine Learning Methods Applied to Predict Ventilator-Associated Pneumonia with Pseudomonas aeruginosa Infection via Sensor Array of Electronic Nose in Intensive Care Unit. *Sensors (Basel)*, 2019. 19(8). doi: 10.3390/s19081866. [PubMed: 31003541).
27. Kuo, K.M., P.C. Talley, C.H. Huang and L.C. Cheng, Predicting hospital-acquired pneumonia among schizophrenic patients: a machine learning approach. *BMC Med Inform Decis Mak*, 2019. 19(1): p. 42. doi: 10.1186/s12911-019-0792-1. [PubMed: 30866913).
28. Tilton, C.S. and S.W. Johnson, Development of a risk prediction model for hospital-onset Clostridium difficile infection in patients receiving systemic antibiotics. *Am J Infect Control*, 2019. 47(3): p. 280-284.10.1016/j.ajic.2018.08.021. doi: 10.1016/j.ajic.2018.08.021 . [PubMed: 30318399).
29. Rabhi, S., J. Jakubowicz and M.H. Metzger, Deep Learning versus Conventional Machine Learning for Detection of Healthcare-Associated Infections in French Clinical Narratives. *Am J Infect Control*, 2019. 58(1): p. 31-41.10.1055/s-0039-1677692. doi: 10.1055/s-0039-1677692 . [PubMed: 30877683).
30. Roimi, M., A. Neuberger, A. Shrot, M. Paul, Y. Geffen and Y. Bar-Lavie, Early diagnosis of bloodstream infections in the intensive care unit using machine-learning algorithms. *Intensive Care Med*, 2020. 46(3): p. 454-462. doi: 10.1007/s00134-019-05876-8 . [PubMed: 31912208).
31. Montella, E., A. Ferraro, G. Sperlì, M. Triassi, S. Santini and G. Improta, Predictive Analysis of Healthcare-Associated Blood Stream Infections in the Neonatal Intensive Care Unit Using Artificial Intelligence: A Single Center Study. *Int J Environ Res Public Health*, 2022. 19(5). doi: 10.3390/ijerph19052498. [PubMed: 35270190).
32. Karajizadeh, M., M. Nasiri, M. Yadollahi, A.H. Zolfaghari and A. Pakdam, Mortality Prediction from Hospital-Acquired Infections in Trauma Patients Using an Unbalanced Dataset. *Healthc Inform Res*, 2020. 26(4): p. 284-294. doi:10.4258/hir.2020.26.4.284. [PubMed: 33190462).
33. Barchitta, M., A. Maugeri, G. Favara, P.M. Riela, G. Gallo, I. Mura, et al., A machine learning approach to predict healthcare-associated infections at intensive care unit admission: findings from the SPIN-UTI project. *J Hosp Infect*, 2021. 112: p. 77-86. doi: 10.1016/j.jhin.2021.02.025 . [PubMed: 33676936).
34. Walker, L.W., A.J. Nowalk and S. Visweswaran, Predicting outcomes in central venous catheter salvage in pediatric central line-associated bloodstream infection. *J Am Med Inform Assoc*, 2021. 28(4): p. 862-867.10.1093/jamia/ocaa328.Pmc7973452. doi: 10.1093/jamia/ocaa328. [PubMed: 33463685).
35. Dos Santos, R.P., D. Silva, A. Menezes, S. Lukasewicz, C.H. Dalmora, O. Carvalho, et al., Automated healthcare-associated infection surveillance using an artificial intelligence algorithm. *Infect Prev Pract*, 2021. 3(3): p. 100167. doi: 10.1016/j.infpip.2021.100167. [PubMed: 34471868).
36. Myall, A., J.R. Price, R.L. Peach, M. Abbas, S. Mookerjee, N. Zhu, et al., Prediction of hospital-onset COVID-19 infections using dynamic networks of patient contact: an international retrospective cohort study. *The Lancet*. Digital health, 2022. 4(8): p. e573-e583. doi: 10.1016/S2589-7500(22)00093-0 . [PubMed: 35868812).
37. Liang, Y., C. Zhu, C. Tian, Q. Lin, Z. Li, Z. Li, et al., Early prediction of ventilator-associated pneumonia in critical care patients: a machine learning model. *BMC Pulm Med*, 2022. 22(1): p. 250. doi: 10.1186/s12890-022-02031-w . [PubMed: 35752818).
38. Thurzo, A., W. Urbanová, B. Novák, L. Czako, T. Siebert, P. Stano, et al., Where Is the Artificial Intelligence Applied in Dentistry? Systematic Review and Literature Analysis. Healthcare (Basel, Switzerland), 2022. 10(7). doi: 10.3390/healthcare10071269 . [PubMed: 35885796).
39. Jin, P., X. Ji, W. Kang, Y. Li, H. Liu, F. Ma, et al., Artificial intelligence in gastric cancer: a systematic review. J Cancer *Res Clin Oncol* , 2020. 146(9): p. 2339-2350.10.1007/s00432-020-03304-9. doi: 10.1007/s00432-020-03304-9. [PubMed: 32613386).
40. Kirk, D., C. Catal and B. Tekinerdogan, Precision nutrition: A systematic literature review. *Comput Biol Med*, 2021. 133: p. 104365. doi: 10.1016/j.compbiomed.2021.104365 . [PubMed: 33866251).
41. Frondelius, T., I. Atkova, J. Miettunen, J. Rello and M.M. Jansson, Diagnostic and prognostic prediction models in ventilator-associated pneumonia: Systematic review and meta-analysis of prediction modelling studies. *J Crit Care*, 2022. 67: p. 44-56. doi: 10.1016/j.jcrc.2021.10.001 . [PubMed: 34673331).
42. Raoofi, S., F. Pashazadeh Kan, S. Rafiei, Z. Hosseinipalangi, Z. Noorani Mejareh, S. Khani, et al., Global prevalence of nosocomial infection: A systematic review and meta-analysis. *PloS one*, 2023. 18(1): p. e0274248. doi: 10.1371/journal.pone.0274248 . [PubMed: 36706112).
43. Akazawa, M. and K. Hashimoto, Artificial intelligence in gynecologic cancers: Current status and future challenges - A systematic review. *Artif Intell Med*, 2021. 120: p. 102164. doi: 10.1016/j.artmed.2021.102164 [PubMed: 34629152).