The predictive effect of different machine learning algorithms for pressure injuries in hospitalized patients: A network meta-analyses

Chaoran Qu, Weixiang Luo, Zhixiong Zeng, Xiaoxu Lin, Xuemei Gong, Xiujuan Wang, Yu Zhang, Yun Li

ARTICLE INFO

Keywords:
- Machine learning algorithms
- Network meta-analysis
- Wound management
- Information management
- Information technology

ABSTRACT

Background: Pressure injury has always been a focus and difficulty of nursing. With the development of nursing informatization, a large amount of structured and unstructured data has been generated, and it is difficult for traditional methods to utilize these data. With the intersection of artificial intelligence and nursing, it has become a new trend to apply machine learning algorithms to build pressure injury prediction models to manage pressure injuries. However, there is no evidence on the effectiveness of the method and which of a large number of algorithms for machine learning is more applicable to pressure injuries.

Objective: This review aims to systematically synthesize existing evidence to determine the effectiveness of applying machine learning algorithms for pressure injury management, to further evaluate and compare pressure injury prediction models constructed by numerous machine learning algorithms, and to derive evidence for the best algorithms for predicting and managing pressure injuries.

Design: Systematic review and network meta-analysis.

Methods: A systematic electronic search was conducted in the EBSCO, Embase, PubMed, and Web of Science databases. We included all retrospective diagnostic accuracy trials and prospective diagnostic accuracy trials constructing a predictive model by machine learning for pressure injuries up to December 2021. Two review authors independently selected relevant studies and extracted data using the Cochrane handbook for systematic reviews of diagnostic test accuracy. The network meta-analysis was conducted using statistical software R and STATA. The certainty of the evidence was rated using the QUADAS-2 tool.

Result: Twenty-five clinical diagnostic trials with a total of 237397 participants were identified in this review. The results of our study revealed that pressure injury machine learning models can effectively predict these injuries. Combining the algorithms separately yields the main results: decision trees (sensitivity: 0.66, 95% CI: 0.42 to 0.84, specificity: 0.90, 95% CI: 0.78 to 0.96, diagnostic odds ratio [DOR]: 18, 95% CI: 7 to 49, AUC: 0.88, 95% CI: 0.85 to 0.91), logistic regression (sensitivity: 0.71, 95% CI: 0.60 to 0.80, specificity: 0.83, 95% CI: 0.75 to 0.89, DOR: 12, 95% CI: 9 to 17, AUC: 0.84, 95% CI: 0.81 to 0.87), neural networks (sensitivity: 0.73, 95% CI: 0.55 to 0.86, specificity: 0.78, 95% CI: 0.65 to 0.87, DOR: 9, 95% CI: 5 to 19, AUC: 0.82, 95% CI: 0.79 to 0.85), random forests (sensitivity: 0.72, 95% CI: 0.26 to 0.95, specificity: 0.96, 95% CI: 0.80 to 0.99, DOR: 56, 95% CI: 3 to 1258, AUC: 0.95, 95% CI: 0.93 to 0.97), support vector machines (sensitivity: 0.81, 95% CI: 0.69 to 0.90, specificity: 0.81, 95% CI: 0.59 to 0.93, DOR: 19, 95% CI: 6 to 54, AUC: 0.88, 95% CI: 0.85 to 0.90). According to the analysis of ROC and AUC values, random forest is the best algorithm for the prediction model of pressure injury.

Conclusions: This review revealed that machine learning algorithms are generally effective in predicting pressure injuries, and after data merging, the random forest algorithm is the best algorithm for pressure injury prediction.

Further well-designed diagnostic controlled trials are recommended to strengthen the current evidence.

Registration number (PROSPERO): CRD42021276993.

* Corresponding author.
E-mail address: Corran0501@outlook.com (Y. Li).

https://doi.org/10.1016/j.heliyon.2022.e11361
Received 15 July 2022; Received in revised form 21 September 2022; Accepted 27 October 2022
2405-8440/© 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
What is already known

- Pressure injuries are small areas of skin or underlying tissue damage caused by pressure and/or shear, usually over a bony prominence.
- With the development of nursing informatization, pressure injury has a large amount of structured and unstructured data.
- Immobility is a major risk factor for the development of pressure injury and is an important component of risk assessment.

What this paper adds

- This review demonstrates that the pressure injury prediction model constructed by a machine learning algorithm has a better effect in diagnosing and predicting pressure injury.
- Among a large number of machine learning algorithms, the random forest algorithm is the best algorithm for building a pressure injury prediction model.
- This study provides evidence for the cross-fertilization of the field of artificial intelligence in nursing.

1. Introduction

Pressure injury (PI) has been classified as a skin disease, and many of factors contribute to the formation of PI (Gillespie et al., 2021). The factors include external pressures causing inadequate or obstructed blood flow to stressed tissues and generating ischemia and hypoxia in the stressed tissues, oxidative stress, and inflammatory responses that occur following hypoxic reperfusion (Alshahrani et al., 2021). Furthermore, for more than 20 years, the global incidence of PI has remained high. PI have been found to have a prevalence of up to 30% in adult patients in evidence-based research (Gillespie et al., 2021). Hospital-acquired pressure injuries are linked to a higher risk of death, longer hospital stays for patients, and significantly higher hospitalization costs (Lovegrove et al., 2021). The Braden tool, Waterlow tool, and Ramstadius tool are three extensively used organized and systematic PI risk assessment methodologies. Unfortunately, studies have demonstrated that their efficiency in preventing the occurrence of PI or assisting in the reduction of the prevalence of PI is inadequate (Healey, 2006; Moore and Patton, 2019).

A variety of information management platforms and information-based medical devices have emerged as a result of advancements in medical information technology. These tools generate large amounts of data in a variety of formats. As a result, these systems store a large amount of structured and unstructured data related to PI. The current phase of research in this field is directed at how to use the data for PI prevention. In recent years, artificial intelligence (AI) has been increasingly utilized to prevent PI. The use of AI to analyze and collect PI-related data has become popular (Jiang et al., 2017). Machine learning (ML) is the most important part of AI technology, and there has been much research on constructing PI prediction models using various ML algorithms (Alderden et al., 2018; Goldstein et al., 2017). Furthermore, preventing PI with AI techniques is more objective and efficient than using the Braden, Waterlow, and Ramstadius tools (Jin et al., 2017; Song et al., 2021a, 2021b). However, machine learning involves a large number of algorithms, and current studies have not yet agreed on the predictive effectiveness of different algorithms. Furthermore, the significance of identifying the best algorithm not only facilitates PI data management but also provides direction and a theoretical basis for future developments in the field. Researchers must select algorithms based on solid scientific evidence.

We determined the comparative efficacy of the ML algorithm for the prevention of pressure injury for inpatients. Network meta-analysis (NMA) was used to fill this critical knowledge gap created by a paucity of directly comparing the effects of different algorithms for constructing predictive models' trials.

2. Materials and methods

2.1. Design

We registered our protocol, which contains details of the literature search strategy. This systematic review and network meta-analysis protocol were registered prospectively in PROSPERO (registration number: CRD42021276993). Our systematic review and NMA manuscript are written in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) extension statement for reporting systematic reviews incorporating NMA (Hutton et al., 2015).

2.2. Inclusion/exclusion criteria

Publications identified in the search of the four databases were combined, and duplicates were removed. Diagnostic trials, crossover trials, and cluster-controlled trials were included, and quasi-randomized studies were excluded as were all other designs.

The population, intervention, comparison, and outcomes (PICO) criteria were applied to all studies to determine their eligibility. The population includes adult patients admitted to any hospital without pressure injury at baseline. Studies were included in which the systematic differences between reported and unreported findings, and varying machine learning algorithms.

The primary outcome measure is an evaluation metric for the effectiveness of applying ML algorithms to construct predictive models (i.e., Receiver Operating Characteristic Curve (ROC) and Area Under the ROC Curve (AUC)) in new pressure injuries, as a primary or secondary outcome measure in the study. The pressure injury stage was defined according to published criteria (Munoz and Posthauer, 2021), or as defined by the study authors. Secondary outcomes included accuracy, precision, negative precision, recall, sensitivity, specificity, and F1 score. Comparative full and partial evaluations were conducted within the framework of eligible diagnostic trials.

2.3. Search strategy

A systematic electronic search will be conducted in the following databases: EBSCO (2010 to December 2021), Embase (2003 to December 2021), PubMed (1985 to December 2021), and Web of Science (1993 to December 2021). Furthermore, expert opinions, the reference lists of the selected studies, and previous systematic reviews will also be reviewed. Finally, all studies will be published in English.

The reference lists of included trials, relevant systematic reviews, meta-analyses, and technology assessment reports were also searched to identify other potentially eligible trials. Search filters were applied to databases, including the Ovid Embase filter, the PubMed Medline filter, and the EBSCO Medline filter. The major search key combination terms were “Pressure ulcer” OR “pressure injury” OR “pressure sore” OR “pressure damage” OR “decubitus ulcer” OR “decubitus sore” OR “Bed sore” OR “pressure sore” OR “Algorithms” OR “Artificial Intelligence” OR “Machine Learning” OR “Deep Learning” OR “Supervised Machine Learning” OR “Support Vector Machine” OR “Unsupervised Machine Learning” OR “Decision Tree” OR “K-means” OR “Bayesian” OR “Ordinary Least Squares Regression” OR “Logistic Regression” OR “Ensemble methods” OR “Clustering Algorithms” OR “Clustering.”
2.4. Study selection

Eligible studies for inclusion were diagnostic trials of different algorithms used to construct predictive models for inpatients with a pressure injury. Pressure injury (e.g., the oxidative stress and inflammatory response that occurs after pressure) causes poor or blocked blood flow to the pressure site tissues, causing ischemia and hypoxia or hypoxic reperfusion to the tissues under pressure. Machine learning algorithms are the most important part of AI technology. In this study, we included the main machine learning algorithms for constructing predictive models: Logistic Regression (LR), Decision Trees (DT), Support Vector Machines (SVM), Random Forests (RF), Bayesian Networks (BN), and Neural Networks (NN).

Two review authors (CRQ and WXL) independently assessed whether all titles and abstracts of retrieved citations met the eligibility criteria by following these steps: (1) reading the title and abstracts and then (2) reading the full texts. Full reports of all potentially relevant trials were retrieved to further assess eligibility. Discrepancies between the reviewers were resolved first by a consensus meeting and then arbitration by a third reviewer (YL) if consensus could not be reached.

2.5. Data collection process and extraction

Data were independently extracted from the included studies by three review authors. A specifically designed data collection tool was used to extract information (e.g., year of publication, author, title, journal title, country, algorithm, eligibility criteria, sample size, confusion matrix, primary and secondary outcome measures). If data were missing from reports, attempts were made to contact the authors. Data were entered into Review Manager 5 and Microsoft Office Excel software by one author and a data check for accuracy was performed by two review authors.
### Table 1. Characteristics of included diagnostic trials (listed in reverse chronological order).

| Author, year | Setting and participants | ML algorithm | Confusion matrices |
|--------------|--------------------------|--------------|--------------------|
| Song et al., 2021a, 2021b | Type of participants: hospital patients  
Number of participants: n = 5814  
Number of features: n = 19  
Include the features: Age, Weight, row, Diarrhea, Bed rest, Restraint bands, Surgery, Braden, Passive Turnover, Nutritional Score, Incontinence Score, Activity Score, Delirium Score, Total Intake, Total Output, Body Temperature, Systolic Blood Pressure, Blood Sugar, Diabetes, Fractures | SVM | TP 770  
FP 41  
FN 50  
TN 2070 |
| Song et al., 2021a, 2021b | Type of participants: hospital patients  
Number of participants: n = 10,915  
Number of features: n = 22  
Include the features: Pressure injury, Race, gender, age, Glasgow coma scale, level of consciousness, gait/ transferring, activity, Pain score, diabetes, peripheral vascular disease, spinal cord injury, stroke, anemia, Albumin, blood urea nitrogen, chloride, potassium, sodium, creatinine, hemoglobin, white blood cell count, platelet blood count | LR  
SVM  
RF  
NN | 1467  
1378  
814  
648 | 1830  
1647  
1  
172 | 300  
389  
1  
95 | 7318  
7501  
2075  
9016 |
| Cai et al., 2021 | Type of participants: surgical patients  
Number of participants: n = 149  
Number of features: n = 9  
Include the features: patient age, gender, disease category, weight, duration of surgery, duration of cardiopulmonary bypass procedure, perioperative corticosteroid administration, use of intraoperative vasoactive agents, use of postoperative vasoactive agents. | DT | 3  
0  
34  
112 |
| Nakagami et al., 2021 | Type of participants: hospital patients  
Number of participants: n = 75,353  
Number of features: n = 46  
Include the features: Age, Gender, Ward type (Internal medicine, Surgery department, Intensive care unit, Obst.), Anorexia, Restricted diet, Denture use, Dysphagia, Urination route (Continent, Incontinent, Catheterized, Fistulized), Urinary incontinence care, Fecal incontinence care, Glasses or contact lens use, Healing aid use, Difficulty in speaking, Paralysis, Pain, Japan Coma Scale (Alert, Dizzy, Somnolent, Comatose), Difficulty in repositioning, Difficulty in sitting up, Difficulty in keeping a sitting position, Difficulty in standing up, Difficulty in keeping a standing position, Difficulty in transferring, Difficulty in moving around, Difficulty in going up and down stairs, Difficulty in hygiene, Cold skin, Hot skin, Wet skin, Dry skin, Jaundice, Oedema, Bony prominence, Contracture, Nasal breathing, Sensory perception, Cough, Sputum, Wheezing, Forced breathing, Cyanosis due to respiratory disorder, Peripheral cold sensation due to respiratory disorder, Cyanosis due to cardiac disorder, Arrhythmia, Palpitation, Peripheral cold sensation due to cardiac disorder | LR  
RF  
SVM  
DT | 288  
300  
284  
308 | 20,308  
18,462  
17,491  
18,948 | 107  
95  
111  
87 | 54,650  
56,496  
57,467  
56,010 |
| Hu et al., 2020 | Type of participants: hospital patients  
Number of participants: n = 11,838  
Number of features: n = 12  
Include the features: Skin integrity, Systolic pressure, Expression ability, Capillary refill time, Level of consciousness, Eye-opening, Level of mobility, Emotional responses, Diastolic pressure, Skin properties, Color in the peripheral limbs, Pulse rate. | DT  
LR  
RF | 129  
112  
140 | 3141  
2791  
3258 | 32  
49  
21 | 8586  
8886  
8419 |
| Li et al., 2020 | Type of participants: hospital patients  
Number of participants: n = 554  
Number of features: n = 8  
Include the features: Department Category, BMI, Skin Type, Incontinence, Poor Eating/Lack of Appetite, Feeding Restricted, Total Assessment Score, Activity Score | SVM  
NN | 160  
158 | 42  
45 | 13  
14 | 62  
60 |
| Mireia et al., 2020 | Type of participants: intensive care unit patients  
Number of participants: n = 1769  
Number of features: n = 23  
Include the features: Medical service, Days of oral antidiabetic agent or insulin therapy, Ability to eat, Number of red blood cell units transfused, Hemoglobin range, Pressure injury present on admission, Illness severity (total APACHE II score), Admission diagnosis, Parenteral or enteral nutrition, Ability to control urination, Cardiac drug treatments, Days of cardiac treatment, Mobility type, History of chronic obstructive pulmonary disease, Admission service, Type of activity, Patient age range, Treatment with sedatives or anesthetics, Physical condition, Type of incontinence, History of cancer, History of dementia, History of diabetes | BN  
DT  
RF  
SVM  
LR  
NN | 3  
44  
20  
65  
62  
56 | 119  
714  
238  
1497  
1497  
1412 | 65  
24  
48  
3  
6  
12 | 1582  
987  
1463  
204  
204  
289 |
| Chen et al., 2019 | Type of participants: cardiovascular disease patients  
Number of participants: n = 1,163  
Number of features: n = 9  
Include the features: Preoperative hemoglobin value, blood sodium value, preoperative albumin, intraoperative mean body temperature, lowest mean arterial pressure, serum potassium value, smoking frequency, history of hypertension, age≥70. | LR | 44  
64  
44  
64 | 78  
71  
23  
16 | 298  
249 |
| Yang et al., 2019 | Type of participants: tumor patients  
Number of participants: n = 611  
Number of features: n = 5  
Include the features: Braden, inability to turn over, existing/potential damage to the skin, special circumstances. | DT | 39  
64 | 128  
71 | 7  
16 | 437  
249 |
| Park et al., 2019 | Type of participants: hospital patients  
Number of participants: n = 400  
Number of features: n = 11  
Include the features: Need for assistance with hygiene, Decreased consciousness, Foley catheter, Cardiac | LR | 64  
64 | 71  
71 | 16  
16 | 249  
249 |

(continued on next page)
### Table 1 (continued)

| Author, year | Setting and participants | ML algorithm | Confusion matrices |
|--------------|--------------------------|--------------|--------------------|
|              |                          | TP          | FP            | FN          | TN          |
| (Cramer et al., 2019) | Type of participants: intensive care unit patients | LR | 777 | 5701 | 913 | 43,460 |
|              | Number of participants: n = 50,851 | SVM | 744 | 5453 | 946 | 43,708 |
|              | Number of features: n = 10 | RF | 473 | 2156 | 1217 | 47,005 |
|              | Include the features: Stage 1 PU within the first 24h, GCS, BUN, pO2, Cardiac Surg, Recovery Unit, Albumin, Medical ICU, Pressure reduction device, Mechanical ventilation, Mean arterial pressure | NN | 828 | 6700 | 862 | 42,461 |
| (Li et al., 2019) | Type of participants: hospital patients | DT | 816 | 261 | 261 | 775 |
|              | Number of participants: n = 2062 | NN | 835 | 261 | 261 | 775 |
|              | Number of features: n = 11 | SVM | 831 | 220 | 220 | 816 |
|              | Include the features: Gender, age, differential diagnosis (ICD-9-CM), history of pressure injuries, length of hospitalization, mental status, excretion, activity/mobility, local skin sensation, skin condition/circulation, nutrition | LR | 489 | 3421 | 246 | 8498 |
| (Hyun et al., 2019) | Type of participants: hospital patients | LR | 1368 |
|              | Number of participants: n = 12,654 | LR | 42 | 614 | 6 | 1368 |
|              | Number of features: n = 10 | LR | 42 | 614 | 6 | 1368 |
| (Gao et al., 2018) | Type of participants: hospital patients | DT | 3349 | 819 | 725 | 3025 |
|              | Number of participants: n = 1940 | NN | 10 | 16 | 4 | 79 |
| (Moon and Lee, 2017) | Type of participants: Long-Term Care patients | DT | 3349 | 819 | 725 | 3025 |
|              | Number of participants: n = 15,856 | NN | 10 | 16 | 4 | 79 |
|              | Number of features: n = 8 | NN | 10 | 16 | 4 | 79 |
| (Chen et al., 2018) | Type of participants: cardiovascular disease patients | LR | 99 | 84 | 491 | 7043 |
|              | Number of participants: n = 100 | BN | 379 | 1359 | 211 | 5768 |
|              | Number of features: n = 13 | DT | 143 | 469 | 447 | 6658 |
|              | Include the features: Gender, BMI, Combined diabetes, Fever one day before operation (body temperature: >37.5 °C, Operative position, Tilt of operating bed, Application of external force, Wet bed sheet, Hypotension, Hypothermia, Emergency operation, Cardiopulmonary bypass, Age, Preoperative score of daily activity ability, Preoperative skin feeling score, Preoperative hemoglobin, Preoperative serum albumin, Operation time, Preoperative waiting time, Intraoperative blood loss, | NN | 6 | 7 | 584 | 7120 |
| (Kaeprag et al., 2017) | Type of participants: intensive care unit patients | DT | 307 | 301 |
|              | Number of participants: n = 7717 | RF | 52 | 44 | 538 | 7083 |
|              | Number of features: n = 12 | SVM | 438 | 1957 | 152 | 5170 |
| (Deng et al., 2017) | Type of participants: intensive care unit patients | DT | 76 | 111 | 18 | 263 |
|              | Number of participants: n = 417 | LR | 36 | 73 | 7 | 301 |
|              | Number of features: n = 7 | DT | 37 | 67 | 6 | 307 |
| (Deng et al., 2016) | Type of participants: intensive care unit patients | LR | 18 | 263 |
|              | Number of participants: n = 468 | DT | 76 | 111 | 18 | 263 |
|              | Number of features: n = 5 | DT | 76 | 111 | 18 | 263 |
| (Setoguchi et al., 2016) | Type of participants: surgical patients | LR | 65 | 1658 | 17 | 4264 |
|              | Number of participants: n = 12,008 | LR | 65 | 1658 | 17 | 4264 |
|              | Number of features: n = 6 | LR | 65 | 1658 | 17 | 4264 |
| (Kim and Lang, 2006) | Type of participants: hospital patients | SVM | 5 | 4 | 3 | 156 |
|              | Number of participants: n = 826 | DT | 4 | 3 | 4 | 157 |
|              | Number of features: n = 8 | LR | 5 | 2 | 3 | 158 |
| (Su et al., 2012) | Type of participants: surgical patients | LR | 12,572 |
|              | Number of participants: n = 168 | LR | 5 | 2 | 3 | 158 |
|              | Number of features: n = 12 | LR | 5 | 2 | 3 | 158 |
| (Cho and Chung, 2011) | Type of participants: hospital patients | LR | 12,807 |
|              | Number of participants: n = 21,114 | LR | 2549 | 5149 | 799 | 12,572 |
|              | Number of features: n = 24 | NN | 2744 | 4959 | 604 | 12,807 |

(continued on next page)
2.6. Risk of bias assessment

Two review authors independently assessed the risk of bias in eligible trials using the QUADAS-2 tool (quality assessment of diagnostic accuracy studies) (Whiting et al., 2011) and PROBAST (prediction model risk of bias assessment tool) (Wolff et al., 2019). The selection of cases, trials to be evaluated, gold standard, case flow, and progression were assessed as low risk of bias, high risk of bias, or unclear risk of bias. The risk of bias was assessed as low if all landmark questions in a range were answered “yes” and high if one of the answers to all questions was “no”.

The “risk of bias” summary figure, which details reviewers’ judgments in a cross-tabulation of studies, provides an assessment of the risk of bias. These trials were classified as having an unknown risk of bias because the authors did not report any validity criteria. Any disagreements between review authors were resolved by consensus or by referral to another review author during study selection, data extraction, and risk of bias assessment.

2.7. Data analysis

Machine learning models are evaluated mainly by confusion matrices, which include True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN). Statistical analysis was conducted using Stata 16 and RStudio. We used a fixed-effects model to summarize the results of the studies with nonsignificant heterogeneity; otherwise, we used the random-effects model. If there was great heterogeneity within the studies (I² > 70%), which would not allow a meta-analysis to be performed, a narrative synthesis of the available data would be conducted. Quantitative data were analyzed using Review Manager 5. We used ANOVA model in R software to implement Bayesian network meta-analysis of diagnostic test accuracy (DTA-NMA).

3. Result

3.1. Study characteristics

The study flow diagram is shown in Figure 1. This research yielded a total of 5082 records retrieved based on the search strategy, and 4308 records were obtained by removing 756 duplicate records and 18 non-English records. A total of 189 records remained after removing 4119 records based on reading abstract and title; then, we removed 112 records because they lacked a prediction model, 24 records owing to discrepancy outcome, 6 records with unavailable full text and 9 records with the image as the data type of the model. Finally, a total of 25 records were included.

Figure 2 shows the network of machine learning algorithm comparisons in available trials. The trial sample size ranged from 100 to 75,353. This research concludes with studies related to the construction of predictive models based on the included studies, mainly applying five ML algorithms: Logistic Regression (LR), Decision Trees (DT), Support Vector Machines (SVM), Random Forests (RF), and Neural Networks (NN). The included studies all contained confusion matrices. Further details regarding the characteristics of included studies are outlined in Table 1.

3.2. Risk of bias assessment

Figure 3 summarizes the risk of bias across the included studies. Some of the trials were assessed as being at high risk of bias in at least one domain (Wade et al., 2013; Whiting et al., 2011). There were domains of the study in which the risk of bias was unclear because they were not mentioned in the text or were poorly described.

3.3. Summary performance estimation

Table 2 shows the summary performance estimates of machine learning algorithms for the meta-analysis of the predictive effect of pressure injury in hospitalized patients for each of the five algorithms. Sensitivity and specificity in machine learning usually combine these two outcomes to evaluate the effectiveness of a model. In the results obtained in this study, the algorithm with the highest sensitivity is the SVM and the algorithm with the highest specificity is the RF. However, for the same algorithm, the sensitivity and specificity within the model are potentially against each other. In this study, a model with higher sensitivity would diagnose as many patients as possible with pressure injuries but would misdiagnose some patients with nonpressure injuries. In addition, models with higher specificity will try to avoid misdiagnosing patients with nonpressure injuries in the results as much as possible, making the diagnosis results more biased toward patients with pressure injuries. In the results of this study, the highest positive likelihood ratio is RF, and the lowest negative likelihood ratio is SVM. Further analysis showed that the outcome of the positive likelihood ratio is the ratio of the algorithm’s correct judgment of the positive to the wrongly judged positive. The larger the ratio is, the greater the probability of true positives when the test result is positive. The result of the negative likelihood ratio is the ratio of the algorithm’s wrong judgment of negative to the true negatives.
Figure 3. Risk of bias summary for the 25 included studies. (Colour coding: Green = low risk; Yellow = unclear; Red = high risk.)

| Study                        | Risk of Bias | Applicability Concerns |
|------------------------------|--------------|------------------------|
| chaoton su 2012              |              |                        |
| cho, in sook 2011             |              |                        |
| eric m, cramer 2019          |              |                        |
| gojiro nakagami 2021         |              |                        |
| honglin chen 2017            |              |                        |
| hsiu–lan li 2019             |              |                        |
| jie song 2021                |              |                        |
| ji–yu, cai 2021              |              |                        |
| lili hou 2010                |              |                        |
| ling gao 2018                |              |                        |
| mikyung moon 2017            |              |                        |
| mireia ladios–martin 2020    |              |                        |
| pacharmon kaewprag 2017      |              |                        |
| qing li 2020                 |              |                        |
| qing yang 2019               |              |                        |
| seul ki park 2019            |              |                        |
| sookyung hyun 2019           |              |                        |
| tae youn kim 2006            |              |                        |
| tara borlawsky, ma 2015      |              |                        |
| wenyu, song 2021             |              |                        |
| xiaohong, deng 2017          |              |                        |
| xiaohong deng 2016           |              |                        |
| yahan hu 2020                |              |                        |
| yoko, se 2016                |              |                        |
| yuan chen 2019               |              |                        |
Table 2. Summary performance estimation of machine learning algorithms for meta-analysis of the predictive effect of pressure injuries in hospitalized patients.

| ML algorithm | Sensitivity (95% CI) | Specificity (95% CI) | +LR (95% CI) | -LR (95% CI) | DOR (95% CI) | AUC (95% CI) |
|--------------|---------------------|----------------------|--------------|--------------|-------------|-------------|
| DT           | 0.66 (0.42, 0.84)   | 0.90 (0.78, 0.96)   | 6.9 (3.2, 14.7) | 0.37 (0.20, 0.69) | 18 (7, 49)   | 0.88 (0.85, 0.91) |
| LR           | 0.71 (0.60, 0.80)   | 0.83 (0.75, 0.89)   | 4.3 (3.1, 5.9)   | 0.35 (0.26, 0.46) | 12 (9, 17)   | 0.84 (0.81, 0.87) |
| NN           | 0.73 (0.55, 0.86)   | 0.78 (0.65, 0.87)   | 3.3 (2.1, 5.0)   | 0.35 (0.21, 0.59) | 9 (5, 19)    | 0.82 (0.79, 0.85) |
| RF           | 0.72 (0.26, 0.95)   | 0.96 (0.80, 0.99)   | 16.3 (2.4, 108.9) | 0.29 (0.07, 1.29) | 56 (3, 1258) | 0.95 (0.93, 0.97) |
| SVM          | 0.81 (0.69, 0.90)   | 0.81 (0.59, 0.93)   | 4.3 (1.8, 9.9)   | 0.23 (0.13, 0.39) | 19 (6, 54)   | 0.88 (0.85, 0.90) |

Abbreviations: +LR, Positive Likelihood Ratio; -LR, Negative Likelihood Ratio; DOR, Diagnostic Odds Ratio; AUC, Area Under Receiver Operating Characteristic Curve; LR, Logistic Regression; DT, Decision Trees; SVM, Support Vector Machines; RF, Random Forests; NN, Neural Networks.

3.4. Outcome of SROC and AUC

Figure 4 summarizes the receiver operating characteristic curves (SROC) of five machine learning algorithms. For the diagnostic test of the classification model, the ROC curve of each model is closer to the upper left corner of the figure, reflecting the better the effect of the model. In the results of this study, the classification effect of the RF prediction model is the best, followed by the SVM model. In this study, the AUC value was used to further rank the models constructed by each algorithm. The results showed that the AUC value of RF was 0.95, the AUC value of SVM and DT was 0.88, the AUC value of LR was 0.84, and the AUC value of NN was 0.82.

4. Discussion

The objective of this systematic review and network meta-analysis is to summarize and compare the prediction effects of pressure injury prediction models constructed by various machine learning algorithms and to obtain the prediction model with the best effect. The literature does not conclude which of the many algorithms is more suitable for predicting pressure injury and does not combine the effect values of the algorithms (Jiang et al., 2021; Nakagami et al., 2021). This study included data from twenty-five clinical diagnostic trials with a total of 237,397 participants. The confusion matrix results of each machine learning prediction model were combined and evaluated, and the prediction effects of five machine learning algorithms in predicting pressure injuries were evaluated and compared according to sensitivity, specificity, positive likelihood ratio, and negative likelihood ratio.

The SROC of each prediction model was plotted to further compare and analyze the prediction effects of the algorithms, and the prediction effects of the algorithms were ranked according to the AUC values. It was concluded that the prediction model was constructed by the LR algorithm had the best prediction effect.

4.1. Limitations of included studies

Overall, the limitations of the included studies diminished the completeness and applicability of the evidence. Assessment of risk of bias identified limitations relative to the reference standard. The inclusion of some trials (Hou and Yao, 2010; Hyun et al., 2019; Mireia et al., 2020; Park et al., 2019; Setoguchi et al., 2016) in this study does not indicate the diagnostic method or diagnostic criteria for pressure injuries during the construction of the model. The threshold effect in the meta-analysis of diagnostic tests is the main source of heterogeneity, and the threshold effect is caused by using different diagnostic cutoff values in a single diagnostic test. However, there was no threshold effect in this study, and further combined effect values were performed.

The limitations of the study data mainly stem from the fact that some of the included studies did not directly give the specifics of the confusion matrix, which was finally calculated based on the data already provided in the study.

4.2. Strengths and limitations of this review

We conducted a rigorous and comprehensive systematic literature search that was reproducible. This review was guided by clearly defined, prespecified procedures to prevent potential bias in the review process and all evidence that could be obtained in the review was considered. Nevertheless, we may have missed trials published in journals that were outside our search strategy. We only merged the five commonly used algorithms, and for the data type also selected structured data. We exclude individual publications using unstructured data for pressure injury prediction, which is different from other literature in this study using structured data for prediction. Explained from the perspective of the algorithm, different data types also affect the effectiveness of the prediction model constructed by the same algorithm.

4.3. Implications for clinical practice

There is no literature on the combined analysis of machine learning prediction models for pressure injuries, and it is not clear which of a large number of machine learning algorithms best predicts and classifies pressure injuries. With the development and advancement of information technology and medical information management, a large amount of complex data has been generated, and traditional analysis methods cannot process and analyze the large amount and complex structure of data. It is necessary to clarify the best algorithm for predicting pressure injuries and to lay the foundation for data, dynamics, and automation of
clinical pressure injury management. Using a large amount of structured and unstructured data in the clinic will improve the accuracy, objectivity, and convenience of pressure injury management.

4.4. Implications for research

To address the methodological limitations identified in the included trials, researchers must ensure transparency of the research process and adhere to the Cochrane handbook for systematic reviews of Diagnostic Test Accuracy (McGrath et al., 2017; Subsoontorn et al., 2020). To minimize sources of bias, researchers need to ensure rigorous processes in study design, type of data for model building, feature engineering, blinding, and the gold standard for diagnosis (McGrath et al., 2017). In addition, due to the specificity of machine learning algorithms, the goodness of fit of the model depends on the source of data and the amount of the data. For supervised learning, feature engineering prior to model construction is a particularly important procedure, and therefore researchers should evaluate the data sources in the literature along with the feature selection methods used to construct the models.

Further research is needed relative to the classification of the model algorithm, which is further refined and then combined and analyzed. The subgroup analysis is performed after classification according to the underlying disease types of the research subjects, and the research on the use of unstructured data to build models is combined. Based on this study, the resulting optimal algorithm builds the decision-making system; applying machine learning algorithms to calculate the economic cost of pressure injury.

5. Conclusion

This literature shows that the application of a machine-learning algorithm to build a model to predict the occurrence of pressure injury is effective. In this study, the combined analysis concluded that the RF algorithm is the best pressure injury prediction model. It is feasible to manage pressure injuries through artificial intelligence methods, which can promote the management methods of pressure injuries to be more objective, data-based, and automated.

Declarations

Author contribution statement

All authors listed have significantly contributed to the development and the writing of this article.

Funding statement

Mr Chaoran Qu was supported by Shenzhen People's Hospital Nursing Research Fund for Young and Middle-aged Nursing Projects [SYHL2022-N0004].

Data availability statement

Data included in article/supp. material/referenced in article.

Declaration of interest’s statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

References

Alderden, J., Pepper, G.A., Wilson, A., Whitney, J.D., Richardson, S., Butcher, R., Joy, Y., Cummins, M.R., 2018. Predicting pressure injury in critical care patients: a machine-learning model. Am. J. Crit. Care 27 (6), 461–468.
Alshahrani, B., Sim, J., Middleton, R., 2021. Nursing interventions for pressure injury prevention among critically ill patients: a systematic review. J. Clin. Nurs. 30 (15-16), 2531–2568.
Cai, J.-Y., Zha, M.-L., Song, Y.-P., Chen, H.-L., 2021. Predicting the development of surgery-related pressure injury using a machine learning algorithm model. J. Nurs. Res. 29 (1).
Chen, H.-L., Yu, S.-J., Xu, Y., Yu, S.-Q., Zhang, J.-Q., Zhao, J.-Y., Liu, P., Zhu, B., 2018. Artificial neural network: a method for prediction of surgery-related pressure injury in cardiovascular surgical patients. J. Wound, Ostomy Cont. Nurs. 45 (1).
Chen, Y., Wu, B., Qian, Q., Fang, Q., Wang, W., 2019. Development and validation of an intraoperatively acquired pressure ulcer risk prediction model for adults receiving cardiovascular surgery. J. Nurs. Sci. 34 (10), 455.
Cho, I.S., Chung, E., 2011. Predictive Bayesian network model using electronic patient record for prevention of hospital-acquired pressure ulcers. J Korean Acad. Nurs. 41 (3), 423–431.
Cramer, E.M., Seneviratne, M.G., Shariff, H., Ozturk, A., Hernandez-Boussard, T., 2019. Predicting the incidence of pressure ulcers in the intensive care unit using machine learning. EGEMS (Washington, DC) 7 (1).
Deng, X., Wang, Q., Li, M., Hu, A., 2016. Predicting the risk of hospital-acquired pressure ulcers in intensive care unit patients based on decision tree. Chin. J. Prac. Nurs. 32 (5), 485–489.
Deng, X., Yu, T., Hu, A., 2017. Predicting the risk for hospital-acquired pressure ulcers in critical care patients. Crit. Care Nurse 37 (4), e1–e11.
Gao, L., Yang, L., Li, X., Chen, J., Du, J., Bai, X., Yang, X., 2018. The use of a logistic regression model to develop a risk assessment of intraoperatively acquired pressure ulcer. J. Clin. Nurs. 27 (15-16).
Gillespie, B.M., Walker, R.M., Latimer, S.L., Thalib, L., Whitby, J.A., McNelis, E., Lockwood, I., Chaboyer, W.P., 2021. Repositioning for pressure injury prevention in adults: an abridged Cochrane systematic review and meta-analysis. Int. J. Nurs. Stud. 120, 103976.
Goldstein, B.A., Novar, A.M., Pencina, M.J., Ioannidis, J.P., 2017. Opportunities and challenges in developing risk prediction models with electronic health records data: a systematic review. J. Am. Med. Inf. Assoc. 24 (1), 198–208.
Healey, F., 2000. Risk assessment tools in the prevention of pressure ulcers. J. Tissue Viability 10 (1), 34–35.
Hou, L., Yao, L., 2010. Logistic regression analysis of risk factors of pressure sore and building of discriminant analysis model. Shanghai Nurs. (1).
Hu, Y., Lee, Y., Kang, M., Lee, P., 2020. Constructing inpatient pressure injury prediction models using machine learning techniques. Comput., Inf., Nurs.: CIN 38 (8).
Hutton, B., Salmanti, G., Caldwell, D.M., Chaimani, A., Schmid, C.H., Cameron, C., Ioannidis, J.P., Straus, S., Thoresen, K., Jansen, J.P., Mulrow, C., Catalá-López, F., Getzsche, P.C., Dickersin, K., Boutilier, I., Altman, D.G., Mohr, D., 2015. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. Ann. Intern. Med. 162 (11), 777–784.
Hyun, S., Moffatt-Bruce, S., Cooper, C., Hixon, B., Kaeawprag, P., 2019. Prediction model for hospital-acquired pressure ulcer development: retrospective cohort study. JIMIR Med. Inf. 7 (3).
Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., Wang, Y., 2017. Artificial intelligence in healthcare: past, present and future. Stroke Vasc. Neuroul. 2 (4), 230–243.
Jiang, Y., Ma, Y., Guo, S., Jin, L., Lv, L., Han, L., An, N., 2021. Using machine learning technologies in pressure injury management: systematic review. JIMIR Med. Inf. 9 (3), e25704.
Jin, Y., Jin, T., Lee, S.M., 2017. Automated pressure injury risk assessment system incorporated into an electronic health record system. Nurs. Res. 66 (6), 462–472.
Kaeawprag, P., Newton, C., Vermillion, B., Hyun, S., Huang, K., Machiraju, R., 2017. Predictive models for pressure ulcers from intensive care unit electronic health records using Bayesian networks. BMC Med. Inf. Decis. Making 17 (Suppl 2).
Kim, T.Y., Lang, N., 2006. Predictive modeling for the prevention of hospital-acquired pressure ulcers. In: AMIA Annual Symposium Proceedings. AMIA Symposium.
Li, H.L., Lin, S.W., Hwang, Y.T., 2019. Using nursing information and data mining to explore the factors that predict pressure injuries for patients at the end of life. Comput. Inf. Nurs. 37 (3), 133–141.
Li, Q., Su, Q., Lin, Y., Deng, G., 2020. Pressure injury analysis and prediction based on machine learning methods. J. Tongji Univ. (Natl. Sci.) 48 (10), 1530–1536.
Lovegrove, J., Ven, S., Miles, S.J., Fulbrook, P., 2021. Comparison of pressure injury risk assessment outcomes using a structured assessment tool versus clinical judgement: a systematic review. J. Clin. Nurs.
McGrath, T.A., Alabasui, M., Skidmore, B., Kurevaa, D.A., Bosvitt, P.M.M., Moher, D., Thombs, B., McNelis, M.D.E., 2017. Recommendations for reporting of systematic reviews and meta-analyses of diagnostic test accuracy: a systematic review. Syst. Rev. 6 (1), 194.
Mirea, L.-M., José, F.-d.-M., Francisco-Javier, B.-L., Adrián, B.-G., Manuel, M.-A., José, C.-M.M., 2020. Predictive modeling of pressure injury risk in patients admitted to an intensive care unit. Am. J. Crit. Care 29 (4).
Moon, M., Lee, S.K., 2017. Applying of decision tree analysis to risk factors associated with pressure ulcers in long-term care facilities. Healthc. Inf. Res. 23 (1), 43–52.
Moore, Z.E., Patton, D., 2019. Risk assessment tools for the prevention of pressure ulcers. Cochrane Database Syst. Rev. 1 (1), Cd006471.

Munoz, N., Posthauer, M.E., 2021. Nutrition strategies for pressure injury management: implementing the 2019 international clinical practice guideline. Nutr. Clin. Pract.

Nakagami, G., Yokota, S., Kitamura, A., Tahashii, T., Morita, K., Noguchi, H., Ohe, K., Sanada, H., 2021. Supervised machine learning-based prediction for in-hospital pressure injury development using electronic health records: a retrospective observational cohort study in a university hospital in Japan. Int. J. Nurs. Stud. 119.

Park, S., Park, H., Hwang, H., 2019. Development and comparison of predictive models for pressure injuries in surgical patients: a retrospective case-control study. J. Wound, Ostomy Cont. Nurs. 46 (4).

Setoguchi, Y., Ghaibeh, A.A., Mitani, K., Abe, Y., Hashimoto, I., Moriguchi, H., 2016. Predictability of pressure ulcers based on operation duration, transfer activity, and body mass index through the use of an alternating decision tree. J. Med. Invest. : JMI 63 (3-4).

Song, J., Gao, Y., Yin, P., Li, Y., Li, Y., Zhang, J., Su, Q., Fu, X., Pi, H., 2021a. The random forest model has the best accuracy among the four pressure ulcer prediction models using machine learning algorithms. Risk Manag. Healthc. Pol. 14, 1175-1187.

Song, W., Kang, M.J., Zhang, L., Jung, W., Song, J., Bates, D.W., Dykes, P.C., 2021b. Predicting pressure injury using nursing assessment phenotypes and machine learning methods. J. Am. Med. Inf. Assoc. 28 (4), 759-765.

Su, C., Wang, P., Chen, Y., Chen, L., 2012. Data mining techniques for assisting the diagnosis of pressure ulcer development in surgical patients. J. Med. Syst. 36 (4).

Subsoonporn, P., Lohitnavy, M., Kongkaew, C., 2020. The diagnostic accuracy of isothermal nucleic acid point-of-care tests for human coronaviruses: a systematic review and meta-analysis. Sci. Rep. 10 (1), 22349.

Wade, R., Corbett, M., Eastwood, A., 2013. Quality assessment of comparative diagnostic accuracy studies: our experience using a modified version of the QUADAS-2 tool. Res. Synth. Methods 4 (3), 280-286.

Whiting, P.F., Rutjes, A.W., Westwood, M.E., Mallett, S., Deeks, J.J., Reitsma, J.B., Leeflang, M.M., Sterne, J.A., Bossuyt, P.M., 2011. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann. Intern. Med. 155 (8), 529-536.

Wolff, R.F., Moons, K.G.M., Riley, R.D., Whiting, P.F., Westwood, M., Collins, G.S., Reitsma, J.B., Kleijnen, J., Mallett, S., 2019. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. Ann. Intern. Med. 170 (1), 51-58.

Yang, Q., Wang, G., Jiang, B., Zhang, H., Lu, X., 2019. Study on risk prediction model of unavoidable pressure ulcers in cancer patients based on decision tre. J. Nurs. Sci. 34 (13), 4-7.