Author/s:
Zachreson, C;Schofield, R;Marshall, C;Kainer, M;Buising, K;Monty, J;Sullivan, S;Subbarao, K;Geard, N

Title:
Modelling the joint effects of single occupancy and N95 respirators on COVID-19 outbreaks in hospital wards

Date:
2025-12-01

Citation:
Zachreson, C., Schofield, R., Marshall, C., Kainer, M., Buising, K., Monty, J., Sullivan, S., Subbarao, K. & Geard, N. (2025). Modelling the joint effects of single occupancy and N95 respirators on COVID-19 outbreaks in hospital wards. Journal of Hospital Infection, 166, pp.190-197. https://doi.org/10.1016/j.jhin.2025.06.014.

Persistent Link:
https://hdl.handle.net/11343/363133

Journal of Hospital Infection 162 (2025) 53–61

Available online at www.sciencedirect.com

Journal of Hospital Infection

journal homepage: www.elsevier.com/locate/jhin

Review

Clinical prediction models to guide treatment of periprosthetic joint infections: a systematic review and meta-analysis

E. Naufal [a], C. Shadbolt [a], M. Wouthuyzen-Bakker [b], S. Rele [a], S. Sahebjada [c], S. Thuraisingam [a], S. Babazadeh [a,d], P.F. Choong [a,d], M.M. Dowsey [a,*]

[a] Department of Surgery, University of Melbourne, St Vincent's Hospital Melbourne, Fitzroy, VIC, Australia
[b] Department of Medical Microbiology and Infection Prevention, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands
[c] Corneal Research Unit, Centre for Eye Research Australia, East Melbourne, VIC, Australia
[d] Department of Orthopaedics, St Vincent's Hospital Melbourne, Fitzroy, VIC, Australia

ARTICLE INFO

Article history:
Received 10 February 2025
Accepted 29 April 2025
Available online 19 May 2025

Keywords:
Systematic review
Clinical predictive model
Prosthetic joint infection

SUMMARY

Background: Several clinical prediction models that aim to guide decisions about the management of periprosthetic joint infections (PJIs) have been developed. While some models have been recommended for use in clinical settings, their suitability remains uncertain.

Methods: We systematically reviewed and critically appraised all multi-variable prediction models for the treatment of PJI. We searched MEDLINE, EMBASE, Web of Science, and Google Scholar from inception until 1st March 2024 and included studies that developed or validated models that predict the outcome of PJI. We used PROBAST (Prediction model Risk Of Bias ASsessment Tool) to assess the risk of bias and applicability. Model performance estimates were pooled via random effect meta-analysis.

Results: Thirteen predictive models and seven external validations were identified. Methodological issues were identified in all studies. Pooled estimates indicated that the KLIC (Kidney, Liver, Index surgery, Cemented prosthesis, C-reactive protein) score had fair discriminative performance (pooled c-statistic 0.62, 95% CI 0.55–0.69). Both the τ^2 (0.02) and I^2 (33.4) estimates indicated that between-study heterogeneity was minimal. Meta-analysis indicated Shohat *et al.*'s model had good discriminative performance (pooled c-statistic 0.74, 95% CI 0.57–0.85). Both the τ^2 (0.0) and I^2 (0.0) indicated that between study heterogeneity was minimal.

Conclusions: Clinicians should be aware of limitations in the methods used to develop available models to predict outcomes of PJI. As no models have consistently demonstrated adequate performance across external validation studies, it remains unclear whether any available models would provide reliable information if used to guide clinical decision making.

* Corresponding author. Address: Level 2, Clinical Sciences Building, 29 Regent Street, Fitzroy, Victoria 3065, Australia. Tel.: +61 3 9231 3955.
E-mail address: mmdowsey@unimelb.edu.au (M.M. Dowsey).

https://doi.org/10.1016/j.jhin.2025.04.035

Introduction

Periprosthetic joint infections (PJIs) are associated with troubling rates of morbidity and mortality [1]. While only 1—2% of patients undergoing primary total joint arthroplasty will develop a PJI, the increasing volume of arthroplasty procedures being performed means that tens of thousands of these complications occur each year [2]. Unfortunately, prognosis for these patients has not improved meaningfully over the past two decades [3]. Selecting the appropriate approach to surgical management of PJI must balance restoration of function and the likelihood of eradicating the underlying infection [4]. Failure rates of surgical debridement for acute PJIs and one- or two-stage exchanges for chronic PJIs greatly differ depending on the host, type of implant and causative micro-organism [5], and require a tailored surgical approach.

Several predictive models that aim to guide patients and clinicians in decisions about the management of PJI have been developed. While some of these models have been recommended for use in clinical settings [6], their suitability for clinical use remains uncertain. To address this uncertainty, we conducted a systematic review to identify studies developing or validating multi-variable models that provide individual-level predictions of treatment outcomes in patients with a PJI. We aimed to describe and appraise these studies, while summarizing the discrimination and calibration of each model.

Methods

This systematic review is reported in line with the transparent reporting of multi-variable prediction models for individual prognosis or diagnosis checklist for systematic reviews and meta-analyses (TRIPOD-SRMA) [7]. The protocol was prospectively registered on PROSPERO (CRD42021281125).

Data sources and study selection

We searched MEDLINE, EMBASE, Web of Science, and Google Scholar from inception until 1st March 2024, for studies reporting multi-variable models predicting the outcome of surgically managed PJIs. Detailed search queries for each database are outlined in Supplementary Tables S1—S5. We manually reviewed the reference lists of eligible studies, and reviewed studies that have referenced all eligible studies since their publication. All citations were uploaded to Covidence (Veritas Health Innovation) to identify duplicates and facilitate screening. Each study record was independently assessed for inclusion by two reviewers (E.N., C.S.) against the eligibility criteria.

Study eligibility criteria

We included all studies that reported on the development or validation of a multi-variable model that aimed to predict the outcome of surgically managed PJIs. To ensure that all included studies aimed to guide patients and clinicians in decisions about the optimal strategy for surgical management of an infection, we excluded models that were designed to be implemented after commencing a specific approach to surgical management. We excluded studies that examined specific predictive factors, rather than estimating the probability of an outcome for an individual patient. Studies that aimed to validate an available model without reporting any statistical measures of performance (e.g., calibration or discrimination) were excluded. Finally, we excluded conference abstracts, pre-prints, and studies published in languages other than English.

Data extraction and risk of bias

Data extraction was conducted by one reviewer (E.N.) and cross-checked by a second reviewer (C.S.). Risk of bias assessment of each eligible study were both undertaken independently by two reviewers (E.N. and S.R.). Data extraction was conducted using a standardized form based on the Checklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies (CHARMS) [8]. When a study described the development of multiple models in different sub-populations or using different statistical approaches, data was extracted separately for each model. When a study validated the same model in the same population across several follow up periods, we extracted data relating to the longest follow up period. For external validations using data from multiple sites, the combined estimates from all sites were extracted.

Risk of bias was assessed using the Prediction model Risk Of Bias Assessment Tool (PROBAST) [9]. We considered studies to be at an overall low risk of bias when all domains were considered low risk, and at an overall high risk of bias when any domain was identified as high risk. When a study reported on the development or validation of more than one eligible model, risk of bias was assessed separately for each model.

Data analyses

We reported descriptive statistics to summarize the key characteristics of all eligible models. Random effect meta-analysis was used to provide summary estimates when performance measures were available from at least two independent datasets. This included both performance measures derived external validations and those derived from internal validations Estimates derived from development datasets were not included in meta-analyses. When confidence intervals (CIs) were not reported, the number of events and number of participants was used to approximate 95% CIs [10]. Heterogeneity was quantified using τ^2 and I^2 statistics. Analyses were conducted using the *metamisc* package in R version 4.3.2.

Results

Of the 2845 studies screened, 16 studies were eligible for inclusion in this review [11—26]. (Figure 1). These studies reported on the development of 13 unique predictive models,

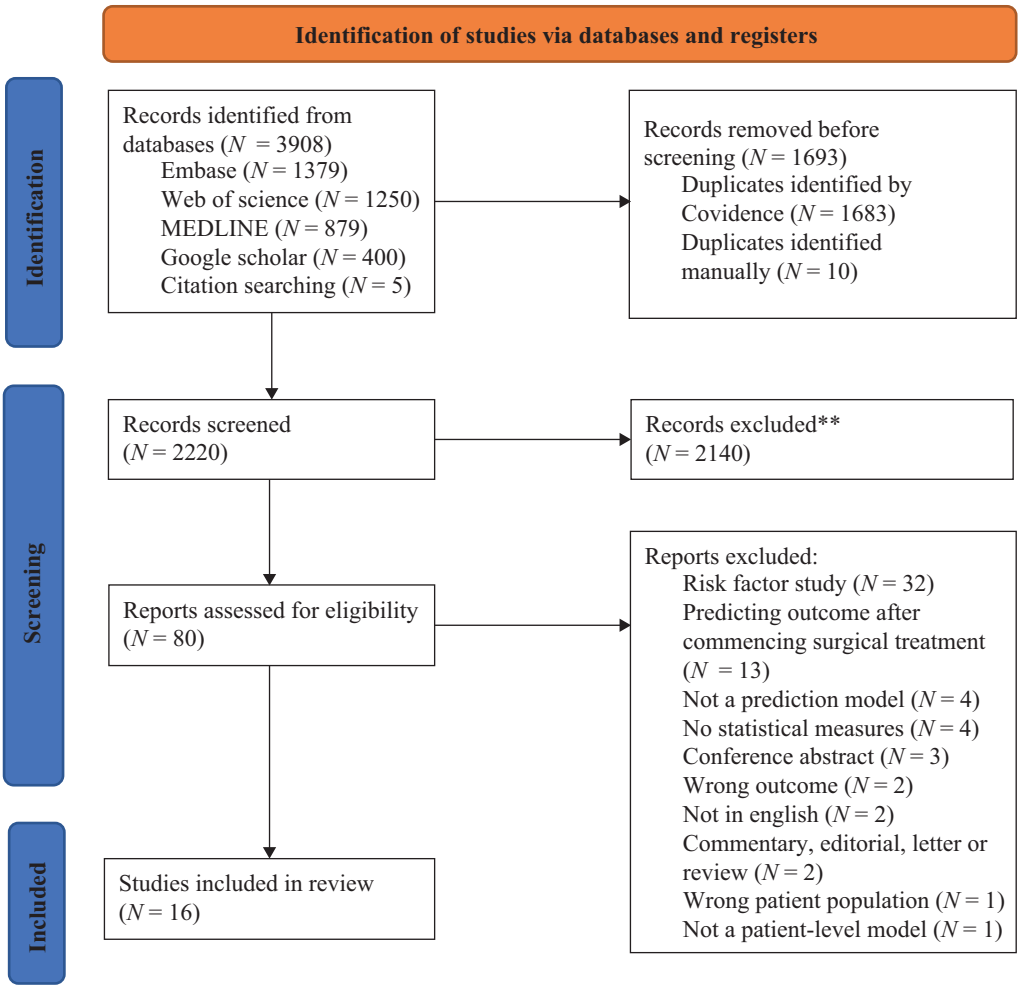

Figure 1. PRISMA flow diagram.

and the external validation of four of these models. Most of these tools were developed using data from single institutions ($N = 10$, 77%) (Table I). Three models (23%) were developed using data from multiple centres, two (15%) of which were trained on data from sites across multiple countries. While most models were developed in the USA ($N = 8$, 62%), a small number of models were developed from centres based in Spain ($N = 1$, 8%), Russia ($N = 1$, 8%) and Belgium ($N = 1$, 8%). All included models were designed using data extracted from hospital medical records or institutional records.

Among models developed using standard statistical modelling techniques, five (63%) were derived from logistic regression models, while three (37%) were derived from Cox models. Among models that employed machine learning methods, the methods employed included neural network ($N = 1$, 20%), random forest ($N = 2$, 40%) and elastic net regularization ($N = 1$, 20%), and classification trees ($N = 1$, 20%). All models aimed to predict a composite endpoint relating to successful treatment or infection eradication, though these endpoints were defined heterogeneously across studies. Most looked at an outcome period of at least one year ($N = 11$, 85%), however, one model included at follow up periods of only 60 days ($N = 1$, 8%). For derivation cohorts, the median sample size was 618 (range: 48–1438) and the median number of events was 165 (range: 14–543). The median number of events per candidate predictor (EPV) was 4 (range: 1–7). None of the model derivation studies reported a sample size calculation.

Target populations

Most models identified in this review were designed to predict treatment outcomes in patients with infections of either hip or knee prostheses ($N = 7$, 54%), though some were designed exclusively for infections of knee prostheses ($N = 4$, 31%) or hip prostheses ($N = 2$, 15%). Most models were designed for all type of PJIs, though one model was developed specifically for acute infections, and another was designed specifically for patients with a late acute haematogenous infection. Six (46%) of the included models were designed to be implemented regardless of the planned approach to surgical management, while five (39%) models specifically predicted outcomes following debridement, antibiotics, irrigation and retention (DAIR) and two (15%) specifically predicted outcomes following two-stage revision. No model was designed to specifically predict outcomes of infection treated with a one-stage revision procedure.

Predictive variables

Across eligible models, the most commonly used variables were body mass index (BMI) ($N = 11$, 85%), age ($N = 8$, 62%),

Table I
Characteristics of studies reporting on the development of models predicting treatment success in patients with prosthetic joint infection

Author (year)	Model name	Follow up	Type of surgery	Infection type	Joint	Modelling method	Centres	Sample size	Events	Candidate predictors	EPP	Internal validation	Reporting on model performance	
													Discrimination	Calibration
Buller *et al.* (2012)	Buller	34 months[a]	DAIR with liner exchange	All	Both	Cox hazard model	1	309	149	13	6	Boot strapping	Yes	No
Kheir *et al.* (2018)	Kheir	≥12 months	All	Unclear	Both	Logistic regression	2	1438	543	29	5	None	Yes	Yes
Klemt *et al.* (2022)	Klemt neural	≥36 months	All	Unclear	Knees	Neural network	1	618	165	77	7	Cross-validation	Yes	Yes
Klemt *et al.* (2022)	Klemt forest	≥36 months	All	Unclear	Knees	Random forest	1	618	165	43	3	Cross-validation	Yes	Yes
Klemt *et al.* (2022)	Klemt elastic	≥36 months	All	Unclear	Knees	Elastic net	1	618	165	43	3	Cross-validation	Yes	Yes
Klemt *et al.* (2021)	Klemt hip	≥24 months	All	Unclear	Both	Logistic regression	1	1081	293	43	3	Random split	Yes	Yes
Klemt *et al.* (2021)	Klemt Knee	≥24 months	All	Unclear	Both	Logistic regression	1	1081	293	56	5	Random split	Yes	Yes
Morcillo *et al.* (2020)	Morcillo	≥24 months	DAIR	All	Hips	Logistic regression	1	48	25	56	5	None	Yes	Yes
Sabry *et al.* (2014)	Sabry	40 months[a]	Two-stage	All	Knees	Cox hazard model	1	314	105	16	1	Boot strapping	Yes	No
Shohat *et al.* (2020)	Shohat	≥12 months	DAIR	Acute and late acute	Both	Random forest	27	1174	405	27	3	Cross-validation	Yes	No
Tornero *et al.* (2015)	KLIC	2 months	DAIR	Acute infection	Both	Logistic regression	1	222	52	52	7	None	Yes	No
TIkhilove *et al.* (2015)	Tikhilov	10 weeks	First step of two-stage	All	Hips	Classification trees	1	217	78	50	1	None	No	No
Wouthuzen-Bakker *et al.* (2019)	CRIME-80	≥24 months	DAIR	Late acute	Both	Cox hazard model	27	340	153	34	4	None	No	No

DAIR, debridement antibiotics irrigation retention; EPP, events per predictor parameter.
[a] Mean follow up in the study sample.

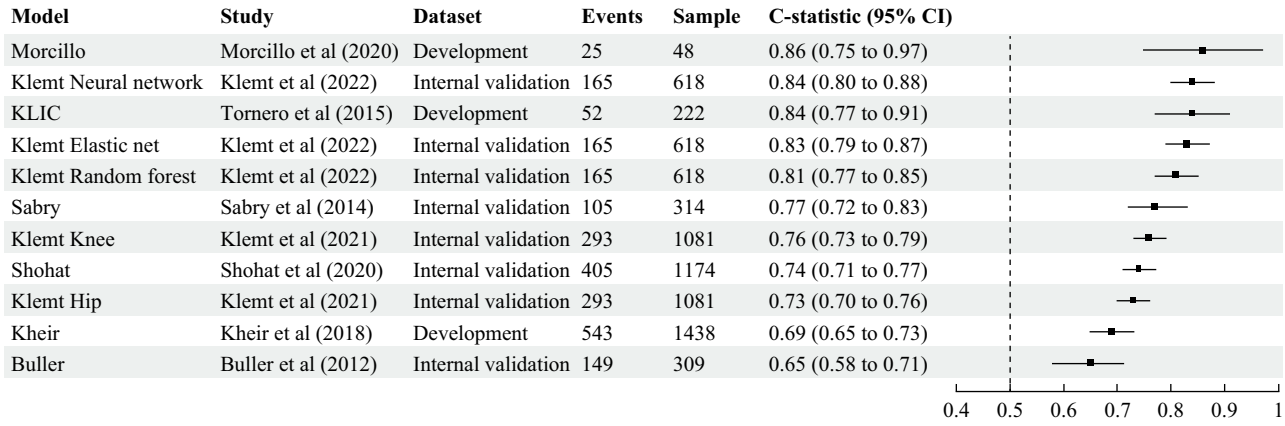

Figure 2. Forest plot of model discrimination for model development studies.

smoking status ($N = 8$, 62%), the exchange of mobile components during DAIR ($N = 8$, 62%), male gender ($N = 8$, 62%), number of previous surgeries ($N = 7$, 54%), renal failure ($N = 7$, 54%), alcohol use ($N = 6$, 46%), presence of enterococcus species ($N = 6$, 46%) and the level of c-reactive protein (CRP) ($N = 6$, 46%) (Supplementary Table S5). Predictors relating to the microbiology of a given infection were included in 11 (85%) models, though the variables used to include microbiological findings varied substantially between models. The median number of candidate predictor parameters included in each model was 43 (range: 13–77). Continuous variables were categorized or dichotomized in most models ($N = 8$, 62%), and only one (8%) model was developed through processes that explicitly allowed for non-linearity when modelling continuous predictors.

Internal validation

Eight (62%) of the models were internally validated. Another three (23%) of the models provided performance measures only from the development dataset, and two (15%) models did not provide any appropriate statistical measures of performance. Internal validations were performed using a random split ($N = 2$, 15%), bootstrapping ($N = 2$, 15%), and cross-validation ($N = 4$, 31%). Among studies that were internally validated, all provided a mean c-statistic as a measure of discrimination, none of these studies reported corresponding CIs. The performance of these model in the internal validations ranged from a mean c-statistic of 0.65–0.84 (Figure 2). Calibration was assessed in seven (54%) internal validation datasets.

External validation

Nine studies reported on the external validation of at least one predictive model, though only four of the unique models were validated across these studies (Table II). The KLIC (Kidney, Liver, Index surgery, Cemented prosthesis, CRP) model, which was developed by Torerno *et al.* [18], was externally validated in six distinct datasets. No other model was validated in more than one external dataset. Cohorts used for external validation had a median sample size of 159 (range: 48–386), the median number of events was 52 (range: 9–164). Seven (78%) of these validations were conducted in cohorts from single centres and two (22%) were conducted in data from more than one site. All external validations provided a mean c-statistic as a measure of discrimination, but only three (33%) of these studies reported corresponding confidence intervals. Model discrimination in external datasets ranged from a mean

Table II
Characteristics of external validations of models predicting treatment success in patients with prosthetic joint infection

Model name	Author (year)	Follow up	Type of surgery	Type of infection	Joint	Sample size	Events	Reporting on model performance		
								Discrimination	Calibration	Other
CRIME-80	Chalmers (2021)	≥24 months	DAIR	Late acute	Hip and knee	256	52	Yes	No	No
Kheir	Monarrez (2021)	≥12 months	All	All	Hip and knee	380	164	Yes	No	No
KLIC	Chalmers (2021)	≥24 months	DAIR	Acute	Hip and knee	256	52	Yes	No	No
KLIC	Jimenez-Garrido (2019)	2 months	DAIR	Acute	Hip and knee	30	9	Yes	No	No
KLIC	Liukkonen (2023)	≥12 months	One-stage	Acute	Hip and knee	123	29	Yes	Yes	No
KLIC	Liukkonen (2023)	≥12 months	DAIR	Acute	Hip and knee	159	58	Yes	Yes	No
KLIC	Lowik (2019)	2 months	DAIR	Acute	Hip and knee	386	148	Yes	No	No
KLIC	Morcillo (2020)	≥24 months	DAIR	All	Hip	48	25	Yes	No	No
Shohat	Sancho (2022)	≥24 months	DAIR	Acute and late acute	Hip and knee	64	25	Yes	Yes	No

DAIR, debridement antibiotics irrigation retention; KLIC, kidney, liver, index surgery.

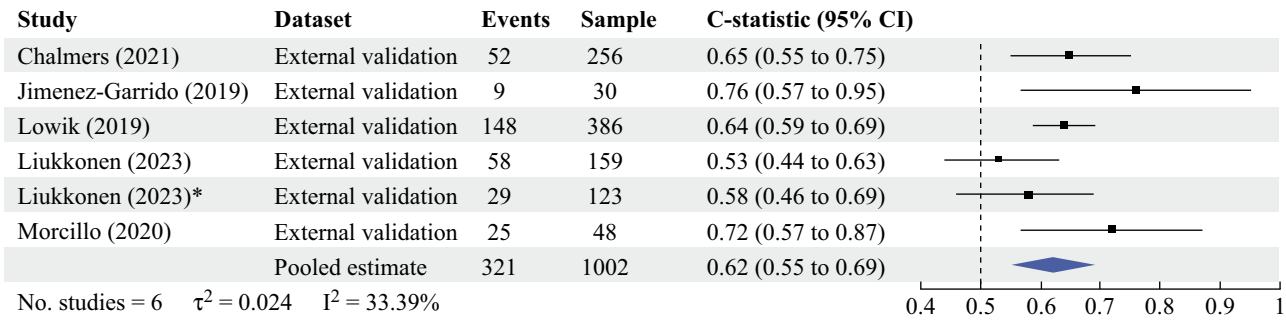

Figure 3. Meta-analysis of model discrimination for KLIC score.

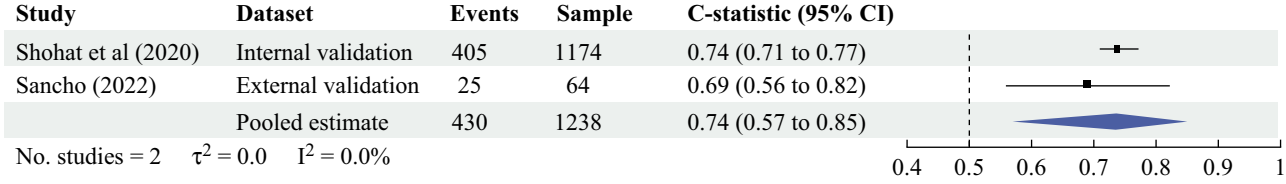

Figure 4. Meta-analysis of model discrimination of model developed by Shohat et al.

c-statistic of 0.53–0.77 (median: 0.65), which was generally lower than that reported in internal validation studies. Only three (33%) models had measures for model calibration in an external validation dataset reported.

Meta-analysis of predictive models

Two models had been validated in multiple independent datasets. Our meta-analysis of the KLIC score used pooled estimates of model discrimination from six external validation studies involving 1002 patients (Figure 3). Measures reported in the initial model development paper were excluded from this analysis as these were derived from the original training dataset. Pooled estimates indicated that the KLIC had fair discriminative performance (pooled c-statistic 0.62, 95% CI 0.55–0.69). Both the τ^2 (0.024) and I^2 (33.39) estimates indicated that between-study heterogeneity was minimal. Meta-analysis of the machine learning model developed by Shohat *et al.* [17] used pooled estimates of model discrimination from one external validation and one internal validation, involving a total of 1238 patients (Figure 4). Although the model has good discriminative performance and the I^2 (0%) showed minimal between-study heterogeneity, pooled estimates were imprecise (pooled c-statistic 0.74, 95% CI 0.57–0.85), and the $\tau^2 = 0.0$. No measures of model calibration could be meta-analysed, as they were not reported in sufficient detail across the available validation studies.

Risk of bias and applicability

All model development studies, and all external validations were assessed as being at high risk of bias (Tables III and IV). All studies were at risk of bias due to reliance on retrospectively

Table III
Risk of bias summary — model development

Model	Risk of bias				Applicability		
	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome
Buller	High	Unclear	High	High	Low	Low	Low
Kheir	High	High	Unclear	High	Low	High	Low
Klemt neural	High	Unclear	Unclear	High	Low	Low	Low
Klemt forest	High	Unclear	Unclear	High	Low	Low	Low
Klemt elastic	High	Unclear	Unclear	High	Low	Low	Low
Klemt hip	High	Unclear	Unclear	High	Unclear	Low	Low
Klemt Knee	High	Unclear	Unclear	High	Unclear	Low	Low
Morcillo	High	Unclear	High	High	Low	Unclear	High
Sabry	High	Unclear	High	High	High	Unclear	Low
Shohat	High	Unclear	High	High	Low	Unclear	Low
KLIC	High	Unclear	High	High	Low	Low	High
Tikhilov	High	High	High	High	Unclear	High	High
CRIME-80	High	Unclear	High	High	Low	Low	Low

KLIC, kidney, liver, index surgery.

Table IV
Risk of bias summary — external validation

Author	Model	Risk of bias				Applicability		
		Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome
Chalmers	KLIC	High	High	Unclear	High	Low	High	Low
Chalmers	CRIME-80	High	High	Unclear	High	Low	High	Low
Jimenez-Garrido	KLIC	High	Unclear	High	High	Low	Low	High
Liukkonen - one stage	KLIC	High	Unclear	High	High	Low	Low	Low
Liukkonen - DAIR	KLIC	High	Unclear	High	High	Low	Low	Low
Lowik	KLIC	High	Unclear	High	High	Low	Low	High
Monarrez	Kheir	High	Unclear	Unclear	High	Low	Low	Low
Morcillo	KLIC	High	Unclear	High	High	Low	Low	High
Sancho	Shohat	High	High	High	High	Low	High	Low

collected data to identify participants. When assessing bias due to inappropriate methods of analysis during model development, no models adequately accounted for complexities in their study data (e.g., accounting for competing risks) and 10 (77%) models handled continuous predictors inappropriately. Furthermore, only two (15%) of the models and none of the external validations appropriately accounted for missing data. Risk of bias assessments also highlighted substantial issues with reporting, with all studies not reporting enough information to respond to at least one signalling question. In total, five models (38%) had at least some concerns about applicability.

Common reason for concerns about applicability included reliance (or potential reliance) on intra-operative microbiological data as predictors ($N = 5$, 38%) or predicting outcomes measured over a period that was too short to adequately guide treatment decisions ($N = 2$, 15%).

Discussion

Management of PJI aims to maximize each patient's chance of eradicating the infection while minimizing the burden associated with invasive surgical intervention. Ideally, decisions would be made with access to reliable information about each individual patient's chance of successful treatment. This systematic review identified 13 predictive models that aimed to provide clinicians and patients with individualized information to guide decisions about the optimal management of PJI. Although these models have been made widely accessible, this review identified important limitations with how each of these models were developed and highlighted that, to date, none of these models has demonstrated strong predictive performance across external validation studies.

The most common predictive variables used in these models were BMI, age, smoking status, the exchange of mobile components during DAIR and male gender. This aligns with previous systematic reviews examining individual risk factors for treatment failure among patients with PJI [27,28]. While most of these variables would be routinely available to treating clinicians if these tools were implemented in practice, half of the tools identified relied upon (or potentially relied upon) intra-operative microbiological data. As intra-operative cultures are not available to clinicians when making decisions about the most appropriate surgical management for a particular patient, this limits the clinical applicability of many of the available predictive models. Even when all included predictive variables would be routinely available, most models categorized continuous predictors or assumed a linear relationship with the outcome, rather than attempting to appropriately specify the functional form of these relationships. This is known to negatively impact model performance [29]. Future efforts to develop improved models should ensure appropriate selection and handling of predictive variables to improve the performance and clinical applicability of these models.

Another common concern was the reliance on inappropriately small samples during model development, with no models being developed in cohorts with more than seven events per candidate predictor parameter. Reliance on such small samples often results in models that are overfit to the development data, which may result in clinicians overestimating how reliable a model is when applied to new patients [30]. This concern is supported by our finding that estimates of discrimination reported in development datasets were generally higher than those reported in internal or external validations. Given the heterogenous nature of patients treated for PJI across centres and countries, external validation is particularly important to ensure that models perform well outside of the context in which they were developed. Unfortunately, only four of the 13 models identified had been externally validated, and all external validation studies were assessed as being of high risk of bias. Several studies purporting to externally validate predictive tools in this area were excluded in the screening process as they did not provide any measures of model performance [31—34].

To date, only one model — the KLIC score [18] — has been externally validated across multiple external cohorts. Our pooled estimates indicate that this score has moderate discriminative performance. Until new models are developed, or existing models can be externally validated across cohorts, this indicates that the clinical value of available tools remains highly uncertain. Without access to reliable and precise predictions of treatment outcomes for individual patients, decision making is likely to be guided by broader, population-level information about outcomes of particular treatment strategies.

There are limitations of this study. We only identified peer reviewed studies written in English. Furthermore, even when meta-analysis for measures of discrimination was possible, it was necessary to impute CIs for most studies. Finally, due to the limited number of studies eligible for meta-analysis, we were unable to explore sources of heterogeneity in the performance of these models through methods such as meta-regression.

In conclusion, clinicians should be aware of important methodological issues with all available studies reporting on the development or validation of predictive models that aim to guide the treatment of PJI. None of the 13 models identified in this systematic review have consistently demonstrated strong predictive performance across multiple external validation studies. Until additional external validations are conducted, the reliability of predictions made by available models is uncertain if they were to be used to guide clinical decisions making. Future research is needed to address these limitations by developing new models in line with methodological best practice and by ensuring that both new and existing models are rigorously externally validated prior to being considered for implementation in clinical settings.

Conflict of interest statement

Members of the team have received grants or contracts from Medacta, Eli Lilly, Medibank Private, HCF foundation, National Health, Medical Research Foundation, Medical Research Future Fund. Members of the team receive royalties from DePuy and Kulwer. Members of the team receive consulting fees from DePuy, Surgeon advisory board, Stryker Corporation, Johnson and Johnson and Medacta. Members have received payment from Zimmerbiomet, Hereaus, University of Otago and Biomerieux for lectures. An author is a member of Osteoarthritis Clinical Research Group and University of Sydney Data Safety Monitoring Board, another author is the board director for Australian Orthopaedic Association Research Foundation. Authors are on the following editorial boards; *EFFORT Reviews*, *Journal of Clinical Medicine* and *JAAOS International*.

Funding sources

No direct funding was received. Elise Naufal acknowledges the support of the Australian Commonwealth Government through a Research Training Scheme Scholarship and Stipend. Michelle Maree Dowsey is the recipient of a University of Melbourne Dame Kate Campbell Fellowship.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhin.2025.04.035.

References

[1] Lum ZC, Natsuhara KM, Shelton TJ, Giordani M, Pereira GC, Meehan JP. Mortality during total knee periprosthetic joint infection. J Arthroplasty 2018;33:3783—8.
[2] Springer BD, Cahue S, Etkin CD, Lewallen DG, McGrory BJ. Infection burden in total hip and knee arthroplasties: an international registry-based perspective. Arthroplast Today 2017;3:137—40.
[3] Xu C, Goswami K, Li WT, Tan TL, Yayac M, Wang SH, et al. Is treatment of periprosthetic joint infection improving over time? J Arthroplasty 2020;35:1696—702e1.
[4] Patel R. Periprosthetic joint infection. N Engl J Med 2023;388:251—62.
[5] Lee J, Kang CI, Lee JH, Joung M, Moon S, Wi YM, Chung DR, et al. Risk factors for treatment failure in patients with prosthetic joint infections. J Hosp Infect 2010;75:273—6.
[6] Argenson JN, Arndt M, Babis G, Battenberg A, Budhiparama N, Catani F, et al. Hip and knee section, treatment, debridement and retention of implant: Proceedings of International Consensus on Orthopedic Infections. J Arthroplasty 2019;34(2):S399—419.
[7] Snell KI, Levis B, Damen JA, Dhiman P, Debray TP, Hooft L, et al. Transparent reporting of multivariable prediction models for individual prognosis or diagnosis: checklist for systematic reviews and meta-analyses (TRIPOD-SRMA). BMJ 2023;381:e073538.
[8] Moons KG, de Groot JA, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. PLoS Med 2014;11:e1001744.
[9] Moons KG, Wolff RF, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. Ann Intern Med 2019;170:W1—33.
[10] Debray TP, Damen JA, Riley RD, Snell K, Reitsma JB, Hooft L, Collins GS, et al. A framework for meta-analysis of prediction model studies with binary and time-to-event outcomes. Stat Methods Med Res 2019;28:2768—86.
[11] Buller LT, Sabry FY, Easton RW, Klika AK. The preoperative prediction of success following irrigation and debridement with polyethylene exchange for hip and knee prosthetic joint infections. J Arthroplasty 2012;27:857—64.
[12] Kheir MM, Tan TL, George J, Higuera CA, Maltenfort MG, Parvizi J. Development and evaluation of a prognostic calculator for the surgical treatment of periprosthetic joint infection. J Arthroplasty 2018;33:2986. 92e1.
[13] Klemt C, Laurencin S, Uzosike AC, Burns JC. Machine learning models accurately predict recurrent infection following revision total knee arthroplasty for periprosthetic joint infection. Knee Surg Sports Traumatol Arthrosc 2022;30:2582—90.
[14] Klemt C, Tirumala V, Smith EJ, Padmanabha A. Development of a preoperative risk calculator for reinfection following revision surgery for periprosthetic joint infection. J Arthroplasty 2021;36:693—9.
[15] Morcillo D, Detrembleur C, Poilvache H, Van Cauter M, Yombi JC, Cornu O. Debridement, antibiotics, irrigation and retention in prosthetic joint infection: predictive tools of failure. Acta Orthop Belg 2020;86:636—43.
[16] Sabry FY, Buller L, Ahmed S, Klika AK. Preoperative prediction of failure following two-stage revision for knee prosthetic joint infections. J Arthroplasty 2014;29:115—21.
[17] Shohat N, Goswami K, Tan TL, Yayac M, Soriano A, Sousa R, et al. ESCMID Study Group of Implant Associated Infections (ESGIAI) and the Northern Infection Network of Joint Arthroplasty (NINJA). 2020 Frank Stinchfield Award: Identifying who will fail following irrigation and debridement for prosthetic joint infection: a machine learning-based validated tool. Bone Joint J 2020;102:11—9.
[18] Tornero E, Morata L, Martinez-Pastor JC, Bori G, Climent C, García-Velez DM, et al. KLIC-score for predicting early failure in prosthetic joint infections treated with debridement, implant retention and antibiotics. Clin Microbiol Infect 2015;21:786-e9. e17.
[19] Wouthuyzen-Bakker M, Sebillotte M, Lomas J, Taylor A, Palomares EB, Murillo O, Parvizi J, et al. Clinical outcome and risk factors for failure in late acute prosthetic joint infections treated with debridement and implant retention. J Infect 2019;78:40—7.
[20] Chalmers BP, Berbari EF, Osmon DR, Hanssen AD, Berry DJ, Abdel MP. Elevated infection and complication rates in patients undergoing a primary THA with a history of a PJI in a prior hip or knee arthroplasty: a matched cohort study. J Arthroplasty 2020;35:1928—32.
[21] Monárrez R, Maltenfort MG, Figoni A, Szapary HJ. External validation demonstrates limited clinical utility of a preoperative prognostic calculator for periprosthetic joint infection. J Arthroplasty 2021;36:2541—5.
[22] Jimenez-Garrido C, Gomez-Palomo JM, Rodriguez-Delourme I, Duran-Garrido FJ, Nuno-Alvarez E, Montanez-Heredia E. The Kidney, Liver, Index surgery and C reactive protein score is a

predictor of treatment response in acute prosthetic joint infection. Int Orthop 2018;42:33—8.
[23] Liukkonen R, Honkanen M, Eskelinen A, Reito A. KLIC-score does not predict failure after early prosthetic joint infection: an external validation with 153 knees and 130 hips. J Arthroplasty 2024;39:1563. 8.e2.
[24] Löwik CA, Jutte PC, Tornero E, Ploegmakers JJ, Knobben BA, de Vries AJ, et al. Predicting failure in early acute prosthetic joint infection treated with debridement, antibiotics, and implant retention: external validation of the KLIC score. J Arthroplasty 2018;33:2582—7.
[25] Sancho I, Otermin-Maya I, Gutiérrez-Dubois J, Aláez I. Accuracy of a novel preoperative failure risk model for Debridement Antibiotics and Implant Retention (DAIR) in acute prosthetic joint infection. Diagnostics 2022;12:2097.
[26] Tikhilov R, Bozhkova S, Denisov A, Labutin D, Shubnyakov I, Razorenov V, et al. Risk factors and a prognostic model of hip periprosthetic infection recurrence after surgical treatment using articulating and non-articulating spacers. Int Orthop 2016;40:1381—7.
[27] Kunutsor SK, Beswick AD, Whitehouse MR, Wylde V, Blom AW. Debridement, antibiotics and implant retention for periprosthetic joint infections: a systematic review and meta-analysis of treatment outcomes. J Infect 2018;77:479—88.
[28] Bhanushali A, Tran L, Nairne-Nagy J, Bereza S, Callary SA, Atkins GJ, et al. Patient-related predictors of treatment failure after two-stage total hip arthroplasty revision for periprosthetic joint infection: a systematic review and meta-analysis. J Arthroplasty 2024;39:2395—402.
[29] Collins GS, Ogundimu EO, Cook JA, Manach YL, Altman DG. Quantifying the impact of different approaches for handling continuous predictors on the performance of a prognostic model. Stat Med. 2016;35:4124—35.
[30] Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. Stat Med. 2016;35:214—26.
[31] Duffy SDX, Ahearn N, Darley ESR. Analysis of the KLIC-score; an outcome predictor tool for prosthetic joint infections treated with debridement, antibiotics and implant retention. Bone Joint Infect 2018;3:150—5.
[32] Sabater Martos M, Hernandez Hermoso JA, Garcia Oltra E, Molinos S, Martínez-Pastor JC. Validity of the KLIC and CRIME80 scores in predicting failure in late acute infection treated by debridement and implant retention. Rev Esp Cir Ortop Traumatol 2020;64:415—20.
[33] Bernaus M, Auñón-Rubio Á, Monfort-Mira M, Arteagoitia-Colino I, Martínez-Ros J, Castellanos J, et al. Risk factors of DAIR failure and validation of the KLIC score: a multicenter study of four hundred fifty-five patients. Surg Infect 2022;23:280—7.
[34] Renz N, Trampuz A, Perka C, Rakow A. Outcome and failure analysis of 132 episodes of hematogenous periprosthetic joint infections — a cohort study. Open Forum Infect Dis 2022;9:ofac094.