EBJIS–MSIS ANNUAL MEETING 2026 · PORTO, PORTUGAL
ABSTRACT ID 456

A Machine Learning Approach to Preoperative Diagnosis of Culture-Positive Periprosthetic Joint Infection

External validation and comparison with EBJIS criteria

Wessling M[1,2], Hanusrichter Y[1,2], Gebert C[1], von Eisenhart-Rothe R[3], Mandraka F[4], Hinterwimmer F[5], Strobl R[6], Weber L[1,7], Gravius S[7], Lazic I[3]

[1] Department of Tumour Orthopaedics and Revision Arthroplasty, Orthopaedic Hospital Volmarstein, Wetter, Germany · [2] Center for Musculoskeletal Surgery, Essen, Germany · [3] Department of Orthopaedics and Sports Orthopaedics, School of Medicine and Health, TUM University Hospital, Technical University of Munich, Munich, Germany · [4] Department of Clinical Infectious Diseases, Wisplinghoff Laboratory, Cologne, Germany · [5] Institute for AI and Informatics in Medicine, School of Medicine and Health, TUM University Hospital, Technical University of Munich, Munich, Germany · [6] IBE — Institute for Medical Information Processing, Biometry and Epidemiology, LMU Munich, Germany · [7] Department of Orthopaedics and Trauma Surgery (OUZ), Heidelberg University Hospital – UK Mannheim GmbH, Mannheim, Germany

0.960
AUC — **development cohort** (n = 1,325)

91.5 / 87.8 %
sensitivity / specificity — development cohort

0.834
AUC — **external cohort**, model applied unchanged (n = 925)

55.2 / 90.2 %
sensitivity / specificity — external cohort

STUDY FLOW

reference standard: identical microorganisms from two independent samples

4,855
revision arthroplasties
2013–2022
culture-positive · 15.3 %

›

1,949
successful aspiration
EBJIS criteria applied
culture-positive · 19.8 %

›

1,325
complete cases
model development
culture-positive · 17.4 %

›

925
external cohort
model unchanged
culture-positive · 64.9 %

INTRODUCTION & AIM

No single currently available test can diagnose periprosthetic joint infection (PJI) with sufficient accuracy across all clinical settings. Every established definition combines several criteria, among them microbiological and histological findings, **some of which** become available only days after the surgical decision has been made.

The synovial thresholds at the centre of these criteria also remain contested: the rule-in cut-offs of 3000 cells/µL and 75 % PMN proposed by the 2025 task-force meta-analysis were reported, on independent reanalysis of the same 74 studies, to reach positive predictive values of only **67–73 % at a prevalence of 20 %**.

For surgical planning, a timely preoperative classification of PJI (yes / no) would be helpful. The objective of this study is the development of a preoperative classification tool. Our AI-based approach required a standardised training cohort with confirmed PJI; we therefore used **culture-positive PJI, defined as identical microorganisms isolated from two independent samples**, as a reference standard that is unambiguous and identical across centres. We aimed **(1)** to analyse the diagnostic performance of a machine-learning model capable of differentiating between PJI and aseptic cases in our own cohort, and **(2)** to validate this model in an external cohort.

METHODS

Design	Retrospective cohort of revision arthroplasties performed at a single centre between 2013 and 2022 (n = 4,855), of which 742 were culture-positive.
Reference standard	Culture-positive PJI, defined as the isolation of identical microorganisms from two independent samples.
Development cohort	First presentations only, one case per patient. Qualifying intervals ensured that antibiotic pre-treatment was unlikely or documented to be absent; referrals under prior antibiotic therapy were excluded. Complete cases: n = 1,325.
Models	Random Forest and XGBoost, using systemic laboratory parameters, synovial fluid cytology and clinical variables. Hyperparameters were tuned by ten-fold cross-validation; the decision threshold was set by the Youden index.
EBJIS comparison	The categories “confirmed” and “likely” were combined. The criteria were evaluated at two timepoints: with preoperative data only, and with the data available up to the start of surgery.
External validation	Independent tertiary referral cohort. The model and its threshold were applied **unchanged**, without refitting or recalibration. Missing predictors were left missing and were not imputed.

DECISION TIMEPOINT AND AVOIDANCE OF CIRCULARITY

“Decision timepoint” denotes the point at which the final surgical strategy is chosen, including synovial fluid obtained immediately before or at the start of the intervention. **No microbiological, histological or sonication result entered the model or the EBJIS classification**, because their inclusion would introduce circularity when using a culture-based reference standard.

MODEL DEVELOPMENT

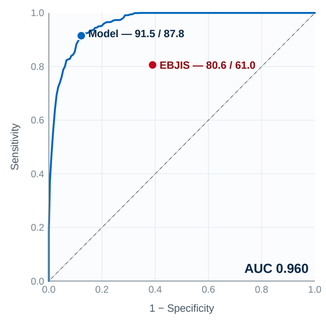
development cohort

CLASSIFIER	N	SENS.	SPEC.	ACC.	AUC
EBJIS criteria preoperative data only	1,777	68.0 %	66.8 %	67.0 %	—
EBJIS criteria up to the start of surgery	1,949	80.6 %	61.0 %	64.9 %	—
XGBoost up to the start of surgery	**1,325**	**91.5 %**	**87.8 %**	**88.6 %**	**0.960**

These figures describe performance on the development data and are therefore optimistic; the external cohort provides the validated estimate. No AUC is given for the EBJIS criteria, as a dichotomised classification yields a single operating point rather than a curve.

SAME DATA, SAME MOMENT

development cohort



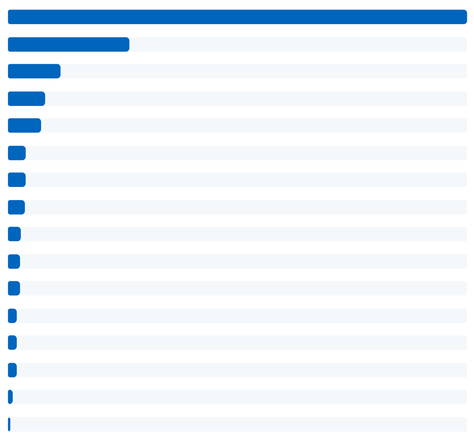
The EBJIS criteria provide a **single operating point**. The model provides a curve, from which the operating point can be chosen.

The EBJIS point lies **below the curve**: within the development cohort, the model achieved higher sensitivity at comparable specificity.

No AUC is reported for the EBJIS criteria: a dichotomised classification has one point, not a curve. Development data; operating point at the Youden cut-off of 0.503.

PREDICTORS

XGBoost variable importance (gain)



Relative contribution of each predictor. The pre-/intraoperative **synovial leukocyte count alone accounts for more than half** of the total, with CRP for about seventy per cent. Both aspirate variables rank above their preoperative-only counterparts. The ranking reflects the size of the contribution, not the direction of the effect.

EXTERNAL VALIDATION

independent tertiary referral cohort · model and threshold unchanged · no imputation

CLASSIFIER	N	PREV.	SENS.	SPEC.	ACC.	AUC	PPV	NPV
XGBoost, unchanged	**925**	**64.9 %**	**55.2 %**	**90.2 %**	**67.5 %**	**0.834**	**91 %**	**52 %**
EBJIS criteria	1,027	64.3 %	91.1 %	88.6 %	90.2 %	—	93.5 %	84.6 %

In this cohort the EBJIS criteria performed better in sensitivity and accuracy, while specificity was comparable.

PREDICTIVE VALUES — RULE-IN RATHER THAN RULE-OUT

Externally, **91 % of positive results were true infections, whereas only 52 % of negative results were truly aseptic**: in this external cohort the model performed predominantly as a **rule-in rather than a rule-out** tool. In the development cohort, at a far lower prevalence, the same model had a PPV of 61 % and an NPV of 98 %. Predictive values follow prevalence; they describe the setting as much as the model.

Potential reasons for reduced sensitivity with preserved specificity

1. **The strongest predictor was not available in its informative form.** Aspiration at the start of surgery is not routine at the external site, and these values were left missing rather than imputed — the external result therefore reflects degraded rather than ideal information.
2. **The prevalence of culture-positive PJI increased from 17.4 % to 64.9 %**, reflecting a substantially different case mix in the external referral cohort. Differences in disease spectrum and predictor distributions may have contributed to calibration drift and to the reduced sensitivity observed at the unchanged development threshold.
3. **Antibiotic pre-treatment**, excluded by design during development, may be more frequent in a referral population and may attenuate the markers on which the model relies.

LIMITATIONS

- The reference standard was a positive microbiological finding: identical organisms in two independent samples. Infections without organism growth lie outside this target by design and count against the EBJIS criteria in this comparison.
- The development cohort was restricted to first presentations without prior antibiotic therapy, whereas the external cohort was a referral population with a far higher infection rate and no routine aspirate at the start of surgery.
- The development figures describe model fit; the external cohort provides the validated estimate.
- Prospective evaluation is required before clinical use.

CONCLUSIONS

- **Higher discrimination in our own cohort.** Applied to the same data at the same moment, the model reached higher sensitivity and specificity for culture-positive PJI than the EBJIS criteria (91.5 / 87.8 % versus 80.6 / 61.0 %) — for this endpoint and this cohort, not as a general statement about the criteria.
- **External application requires recalibration and validation in additional independent cohorts.** Discrimination was retained in the external cohort (AUC 0.834); the decision threshold from the development cohort was not transferable.
- **The model may complement established definitions; it is not intended to replace them.** Externally it behaved as a rule-in test: a positive result was correct in nine out of ten cases, a negative result was not.
- **The two approaches showed potentially complementary diagnostic profiles in this cohort:** the EBJIS criteria had a higher NPV (84.6 %), whereas the model achieved a PPV of 91 %. Whether combining them improves diagnostic performance, and how discordant results should be handled — repeat aspiration or biopsy — remains to be tested.

TAKE-HOME MESSAGE

An additional tool at the moment of decision. From routinely available blood values and aspirate cytology, the model quantifies the probability of culture-positive PJI — information the categorical criteria do not provide.

www.pjai.app
Structured pathway for revision arthroplasty and PJI
SCAN FOR FURTHER INFORMATION

References. McNally M, Sousa R, Wouthuyzen-Bakker M, et al. *Bone Joint J* 2021;103-B:18–25. · Sabater-Martos M, Clauss M, Ribau A, Sousa R, et al. *J Bone Joint Infect* 2025;10:165–184. · Parthasarathi A, Setoguchi S, DeConti DK, et al. *Clin Orthop Relat Res* 2026, epub, doi:10.1097/CORR.0000000000004002. · Tsai SW, Mu W, Parvizi J. *J Clin Orthop Trauma* 2024;52:102430. · Sax FH, Hoyka M, Blersch BP, Fink B. *Antibiotics* 2024;13:351.

Correspondence. Dr. Martin Wessling, Orthopaedic Hospital Volmarstein, Wetter, Germany · **wesslingm@esv.de** · Prof. Dr. med. Igor Lazic, TUM University Hospital, Technical University of Munich.
Keywords. PJI diagnostics · machine learning · EBJIS criteria · external validation · transportability