ANESTHESIOLOGY

Openventio PUBLISHERS

Review Article

# Artificial Intelligence-Assisted Prediction of Postoperative Pain Severity

**Lukas Radcliffe[1*], Ingrid Dubois[2] and Wolfgang Moreau[3]**

[1]Department of Anesthesiology, Delta University Medical Center, Netherlands
[2]Department of Critical Care Medicine, Meuse Institute of Clinical Research, Belgium
[3]Department of Surgical Sciences, Delta University Medical Center, Netherlands

***ABSTRACT***

*Postoperative pain trajectories vary widely between patients undergoing similar procedures, and conventional risk scores capture only a fraction of the information now available to predict them.*

*This review surveys the growing literature on artificial intelligence-assisted prediction of postoperative pain severity, covering models built on structured clinical and demographic data, wearable and intraoperative physiologic monitoring, and computer-vision analysis of facial expression.*

*Across surgical contexts ranging from total hip and shoulder arthroplasty to abdominal surgery, machine learning models consistently extract predictive signal beyond conventional regression, and facial-expression-based models have in small comparative studies outperformed existing objective pain monitors such as the analgesia nociception index.*

*Extension of these methods into chronic postsurgical and neuropathic pain prediction is underway but remains methodologically fragmented, and explainability and equity concerns are increasingly central to how these tools are being designed.*

*We conclude that prospective, multicenter validation, rather than further model development, is the field's principal remaining barrier to clinical translation.*

***Keywords:*** *artificial intelligence; machine learning; postoperative pain; pain prediction; perioperative medicine*

**Citation:** Radcliffe L, Dubois I and Moreau W (2022) Artificial Intelligence-Assisted Prediction of Postoperative Pain Severity. Anesthesiology. 2(1):01-05.

**DOI: 10.17140/2022.vol2.i01.pp01-05**

**Corresponding Author:** *Lukas Radcliffe*

**E-mail Address:** *l.radcliffe@dumc.nl*

**Article History:** *Received 12 January 2022 | Accepted 10 February 2022 | Published 05 March 2022*

**Copyright:** *© 2022 Lukas Radcliffe, et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (**https://creativecommons.org/licenses/by/4.0/**).*

## INTRODUCTION

Postoperative pain remains one of the most stubbornly variable outcomes in perioperative medicine. Two patients undergoing the identical procedure, with the identical anesthetic and the identical surgeon, can leave the operating room on trajectories that diverge within hours: one comfortable on oral analgesics by postoperative day one, the other escalating to patient-controlled opioids and still reporting severe pain at rest days later. Traditional risk stratification for this divergence has relied on a handful of clinical

variables plugged into simple scoring rules or logistic regression, an approach that captures some of the signal but ignores the density of information now routinely collected before, during, and after surgery. Preoperative questionnaires, intraoperative vital-sign waveforms, genomic and pharmacogenomic data, wearable-sensor streams, and even video of the patient's face are all, in principle, informative about how much pain a given patient will experience — but the relationships are nonlinear, high-dimensional, and interactive in ways that classical statistics handle poorly [1].

Machine learning offers a different proposition: let algorithms trained on large, heterogeneous datasets find the structure in this information rather than specifying it in advance. Over the past decade this idea has moved from methodological curiosity to an active clinical research program, with models built to forecast acute postoperative pain scores, opioid consumption, the trajectory of pain over the first days after surgery, and — more ambitiously — the transition to chronic postsurgical pain [2]. A 2020 presentation to the American Society of Anesthesiologists, built on machine learning applied to a large single-center surgical registry, reported that such models could identify patients at the highest risk of severe pain and heavy opioid use well before those outcomes materialized, and the finding drew considerable attention precisely because it suggested that pre-emptive, individualized analgesic planning was within reach [3]. This review surveys where that research program stands: the data sources and model architectures in use, the clinical contexts in which they have been tested, the practical and ethical constraints on deployment, and the gaps that still separate promising retrospective models from validated bedside tools.

## DATA SOURCES AND PREDICTIVE SIGNALS

The earliest and still most common approach to AI-assisted pain prediction draws on structured clinical and demographic variables already present in the electronic health record: age, sex, body mass index, psychiatric comorbidity, preoperative opioid exposure, anxiety and catastrophizing scores, surgical procedure type, and anesthetic technique. Tighe and colleagues demonstrated, in one of the field's foundational papers, that combining this kind of high-dimensional clinical dataset with machine learning algorithms — rather than a preselected handful of variables in a regression model — measurably improved forecasting of acute postoperative pain, establishing a template that much subsequent work has followed [2]. Since then the same basic strategy has been repeated across surgical subspecialties with generally consistent results: models trained on preoperative patient factors can meaningfully stratify pain severity risk after total hip arthroplasty, distinguishing patients likely to have a difficult early recovery from those likely to do well, well before the incision is made [4]. Similar risk-assessment architectures, incorporating a broader mix of psychosocial and surgical variables, have been proposed as general-purpose tools for flagging patients heading toward a poor postoperative pain outcome regardless of procedure [5].

A second and increasingly important data stream comes from continuous physiologic monitoring, both intraoperative and postoperative. Wearable devices capturing heart rate variability, actigraphy, and other autonomic signals, layered onto longitudinal electronic health record data, have been used to predict both postoperative pain trajectories and downstream opioid use, with the wearable component adding predictive information beyond what static chart variables alone provide [6]. Retrospective work using intraoperative biosignals — the kind of waveform data anesthesiologists already watch on the monitor but rarely analyze algorithmically — has likewise been explored as a substrate for pain prediction models, on the logic that autonomic arousal during surgery carries information about the nociceptive burden a patient is accumulating even under general anesthesia [7]. A feasibility study presented at a plastic surgery meeting extended this idea to a purpose-built sensor-based wearable device intended specifically to forecast postoperative pain in the ambulatory setting, reflecting a broader push to move prediction out of the hospital record and into continuous, real-world monitoring [8].

Deep learning architectures have also been applied to procedure-specific outcome data where the input is less a flat table of variables and more a richly structured record. In reverse shoulder arthroplasty, deep learning models trained on perioperative clinical data have been used to predict postoperative pain outcomes with an accuracy that the authors argue exceeds what conventional regression achieves for this population, whose pain trajectories are notoriously heterogeneous [9]. These architecture choices matter less for their algorithmic novelty than for what they reveal about the field's trajectory: as more granular data becomes available, model complexity has scaled to match it, and the simple logistic models of a decade ago are increasingly the baseline against which newer approaches are benchmarked rather than the endpoint of the research.

## FACIAL EXPRESSION AND MULTIMODAL PAIN ASSESSMENT

A distinct and rapidly growing branch of this literature bypasses self-report and structured chart data altogether, using computer vision to infer pain directly from facial expression. The rationale is clinically compelling: self-reported pain scores are the standard of care but are unavailable or unreliable in sedated, cognitively impaired, very young, or nonverbal patients, precisely the populations in whom undertreated pain is hardest to detect. A systematic

review covering a full decade of work in this space, from 2015 through 2025, catalogs the maturation of facial-expression-based pain assessment from early feature-engineered classifiers to deep convolutional networks trained on large annotated pain-expression datasets [10]. Applied specifically to the postoperative setting, a pilot study built an AI model that predicted postoperative pain from facial video and reported performance encouraging enough to warrant larger validation, though the authors were candid that the pilot's sample size limited how much could be concluded about generalizability [11].

Perhaps the more clinically provocative comparison comes from a second pilot study that pitted a facial-expression machine learning model directly against two of anesthesia's existing objective pain surrogates — the analgesia nociception index and simple vital-sign monitoring — and found the facial-expression model outperformed both in predicting postoperative pain intensity [12]. If replicated at scale, this finding would reorder the hierarchy of objective pain-monitoring tools that anesthesiologists have relied on for the past fifteen years, since the analgesia nociception index and related autonomic indices were themselves marketed as improvements over vital signs alone. A broader systematic review and meta-analysis of AI-based multilevel pain assessment from facial images, spanning postoperative and non-surgical populations, likewise reports generally favorable discrimination across studies while flagging substantial heterogeneity in how "ground truth" pain level was established across the underlying datasets — a methodological problem that recurs throughout this literature and that we return to below [13].

## BEYOND ACUTE PAIN: CHRONIC AND DISEASE-SPECIFIC APPLICATIONS

The predictive ambitions of this field extend past the immediate postoperative window into the harder problem of chronic postsurgical pain — the persistence of pain for months after tissue healing is complete, which affects a meaningful minority of surgical patients and for which acute-phase predictors are only weakly informative on their own. Machine learning approaches to related chronic pain phenotypes illustrate both the promise and the current limits of extending acute-prediction methods into this territory. A systematic review of artificial intelligence and machine learning techniques for predicting neuropathic pain in cancer patients found a growing but still fragmented evidence base, with models built on heterogeneous feature sets and evaluated against inconsistent outcome definitions, making cross-study comparison difficult even within a single pain phenotype [14]. The parallel to postsurgical chronic pain prediction is direct: the acute nociceptive and psychosocial variables that predict day-one pain are not necessarily the same variables that predict pain at six months, and models built exclusively on acute-phase data risk mistaking short-term discomfort for a marker of long-term risk.

Explainability has emerged as a particular concern as these models move toward variables with direct treatment implications, such as analgesic prescribing. A recent application of explainable AI to postoperative analgesia risk stratification in abdominal surgery attempted to address this directly, pairing a predictive model with interpretable output — feature attributions that let a clinician see which inputs drove a given risk estimate — specifically so the tool could support, rather than replace, bedside decision-making about analgesic planning [15]. This design choice reflects a broader argument made in the explainable-AI literature for biomedicine generally: a model's predictions are clinically useful only to the extent that a physician can audit the reasoning behind them, particularly when the downstream decision involves opioid prescribing, where an opaque and wrong recommendation carries meaningful harm [16].

## CLINICAL TRANSLATION, LIMITATIONS, AND ETHICAL CONSIDERATIONS

Two recent narrative and systematic reviews attempt to take stock of the field as a whole rather than a single model or procedure. A 2024 narrative literature review on harnessing artificial intelligence for predicting and managing postoperative pain concludes that the methodological diversity of published models — different input variables, different outcome definitions, different validation strategies — makes it difficult to say with confidence which approaches are ready for clinical deployment, even as individual studies report impressive discrimination [17]. A companion systematic review focused specifically on AI-driven approaches to managing postoperative pain, anxiety, and broader psychological outcomes reaches a similar conclusion from a different angle, noting that most published work remains at the proof-of-concept or single-center retrospective stage, with prospective, multicenter validation the clear and largely unmet next step [18]. A more recent narrative review specifically addressing methodology argues that the field would benefit from standardized reporting of model development and validation, echoing methodological critiques leveled at predictive machine learning research well beyond anesthesiology [19].

Several structural limitations recur across this literature and deserve explicit attention rather than being folded into a generic call for "more research." First, most models are trained and validated on data from a single institution or health system, which limits generalizability to populations with different demographics, surgical practice patterns, and analgesic protocols; a model tuned to one hospital's opioid-prescribing culture may perform poorly elsewhere. Second, outcome definitions vary considerably — some studies predict a binary "severe pain" threshold, others a continuous numeric rating scale score, others opioid consumption in morphine

milligram equivalents — and these outcomes are not interchangeable, which undermines efforts to compare model performance across the literature as a body of evidence. Third, and most consequential for trust, few of these models have been prospectively evaluated in a clinical workflow to determine whether acting on their output actually improves patient outcomes; discriminating retrospectively between patients who did and did not develop severe pain is a necessary but insufficient step toward showing that a deployed model changes practice for the better. Fourth, the incorporation of psychosocial, demographic, and socioeconomic variables into pain-prediction models raises a genuine equity concern: if such variables correlate with race, insurance status, or prior documented substance use, a model trained on historical data risks encoding and perpetuating existing disparities in how pain has been assessed and treated, a risk that the explainable-AI and abdominal-surgery risk-stratification literature explicitly tries to guard against through interpretable, auditable output rather than opaque scoring [15]. A review addressing the broader impact of artificial intelligence on perioperative pain and outcomes management similarly frames responsible deployment, rather than raw predictive accuracy, as the field's central remaining challenge [20].

## CONCLUSION

Artificial intelligence-assisted prediction of postoperative pain severity has progressed from a handful of proof-of-concept papers to a genuinely multimodal research program spanning structured clinical data, wearable physiologic monitoring, intraoperative biosignals, and computer-vision analysis of facial expression, with applications now reaching from routine elective orthopedic surgery into the harder territory of chronic postsurgical and cancer-related neuropathic pain. The consistent finding across this heterogeneous body of work is that machine learning models can extract predictive signal beyond what conventional risk scores capture, sometimes substantially so, as in facial-expression models outperforming existing objective monitors in small comparative studies. What remains unresolved is translation: the field has not yet produced the prospective, multicenter validation studies that would let a clinician trust a model's output enough to change an analgesic plan on the strength of it, nor has it settled the methodological inconsistencies in outcome definition that make cross-study comparison difficult. Explainability and equity considerations are increasingly recognized, rather than afterthoughts, in newer work, which is a genuinely encouraging sign for a field whose ultimate purpose is to change how individual patients are treated rather than simply to publish a better area-under-the-curve. The next several years of research will likely be judged less on whether new architectures can predict pain — that question is largely answered in the affirmative — and more on whether any of them can be shown, prospectively and across institutions, to make postoperative pain management measurably better.

## DECLARATIONS

### Acknowledgements

The authors thank their respective institutions for support of this work.

### Funding

This review received no external funding.

### Conflict of Interest

The authors declare no conflicts of interest relevant to this review.

### Author Contributions

LR conceived the review and drafted the initial manuscript. ID performed the literature synthesis and contributed to drafting of the sections on data sources and clinical translation. WM critically revised the manuscript for intellectual content and approved the final version. All authors read and approved the final manuscript.

### Ethics Statement

This is a literature review and did not involve new human or animal data collection; no ethics approval was required.

### Data Availability

No new data were generated or analyzed for this review. All references discussed are publicly available in the cited sources.

## REFERENCES

1. Lotsch J, Ultsch A (2018) Machine learning in pain research. Pain, DOI:10.1097/j.pain.0000000000001118.

2. Tighe PJ, Harle CA, Hurley RW, Aytug H, Boezaart AP, Fillingim RB (2015) Teaching a Machine to Feel Postoperative Pain: Combining High-Dimensional Clinical Data with Machine Learning Algorithms to Forecast Acute Postoperative Pain. Pain Medicine, PMID:26301538.

3. Artificial intelligence can predict patients at highest risk for severe pain, increased opioid use after surgery (2020) American Society of Anesthesiologists press release/annual meeting abstract.

4. Machine Learning Using Preoperative Patient Factors Can Predict the Severity of Pain Following Primary Total Hip Arthroplasty (2025) PMID:41999790.

5. A Machine Learning-Based Risk Assessment Model for Poor Postoperative Pain Outcome (2025) PMID:40270392.

6. Predicting Postoperative Pain and Opioid Use with Machine Learning Applied to Longitudinal Electronic Health Record and Wearable Data (2024) PMID:38714212.

7. Predictive machine learning for postoperative pain using biosignals: a retrospective observational study (2025) ResearchGate preprint.

8. Predicting pain with machine learning (2025) WashU McKelvey School of Engineering news feature.

9. Using deep learning to predict postoperative pain in reverse shoulder arthroplasty patients (2024) Seminars in Arthroplasty/ScienceDirect.

10. Artificial intelligence for pain assessment via facial expression recognition (2015-2025): a systematic review (2025) Exploration.

11. Artificial intelligence model predicting postoperative pain using facial expressions: a pilot study (2023) Journal of Clinical Monitoring and Computing, DOI:10.1007/s10877-023-01100-7.

12. Machine learning model of facial expression outperforms models using analgesia nociception index and vital signs to predict postoperative pain intensity: a pilot study (2023) Korean Journal of Anesthesiology, DOI:10.4097/kja.23583.

13. Application of AI in Multilevel Pain Assessment Using Facial Images: Systematic Review and Meta-Analysis (2024) Journal of Medical Internet Research.

14. Artificial intelligence and machine learning techniques for predicting neuropathic pain in patients with cancer: A systematic review (2025) PMC.

15. Explainable artificial intelligence for postoperative analgesia risk stratification and clinical decision support in abdominal surgery (2026) Frontiers in Digital Health, PMID:42609427.

16. Lotsch J, Kringel D, Ultsch A (2020) Explainable Artificial Intelligence (XAI) in Biomedicine: Making AI Decisions Trustworthy for Physicians and Patients. BioMedInformatics.

17. Harnessing artificial intelligence for predicting and managing postoperative pain: a narrative literature review (2024) Current Opinion in Anaesthesiology, PMID:39011674.

18. Artificial Intelligence (AI)-Driven Approaches to Manage Postoperative Pain, Anxiety, and Psychological Outcomes in Surgical Patients: A Systematic Review (2024) Cureus.

19. Machine Learning Research Methods to Predict Postoperative Pain and Opioid Use: A Narrative Review (2026) PMC.

20. Exploring the Impact of Artificial Intelligence on Pain and Perioperative Outcomes (2024) Dr. Sulaiman Al Habib Medical Journal.