

The Future of Neonatal and Infant Pain Assessment: A Systematic Review of Discriminative AI Approaches

INPLASY2026100090

doi: 10.37766/inplasy2026.10.0090

Received: 20 October 2026

Published: 20 October 2026

Hamwi, L; Khosrojerdi, B; Zahir, S, Ebrahimi, M; Kassiri, H; Riddell, RP.

Corresponding author:

Rebecca Pillai Riddell

rpr@yorku.ca

Author Affiliation:

Department of Psychology, Faculty of Health, York University, Toronto, Ontario, Canada.

ADMINISTRATIVE INFORMATION

Support - This research is funded by operating grants from the Canadian Institutes of Health Research (CIHR) and Natural Sciences and Engineering Research Council (NSERC) to Dr. Pillai Riddell. Dr. Pillai Riddell is supported by salary support from the York University Research Chair Program. Ms. Hamwi is supported by salary support or awards from CIHR, York University, and the Meighen Wright Maternal-Child Health Scholarship.

Review Stage at time of this submission - Formal screening of search results against eligibility criteria.

Conflicts of interest - None declared.

INPLASY registration number: INPLASY2026100090

Amendments - This protocol was registered with the International Platform of Registered Systematic Review and Meta-Analysis Protocols (INPLASY) on 20 October 2026 and was last updated on 20 October 2026.

INTRODUCTION

Review question / Objective The objective of this systematic review is to characterize the current state of approaches to neonatal and infant pain assessment using artificial intelligence (AI). To this end, the proposed systematic review will address the following questions: What machine learning approaches have been developed or evaluated for the assessment of pain in neonates and infants? Which populations, pain contexts, infant pain response data modalities, and algorithms have been used? What model-performance results have been reported? What methodological limitations and recommendations for future research to support the development of AI for infant pain assessments?

Rationale Recent systematic reviews examined artificial intelligence applications for pediatric pain assessment and management across children (Zamzmi et al., 2018; Crispin et al., 2026; Zhou et al, 2026; Wang et al., 2026). However, their broad pediatric scope combined neonatal and infant studies with research involving older children and adolescents and included both pain assessment and AI-supported pain-management interventions. The proposed review will focus specifically on AI algorithms for pain assessment in neonates and infants up to 12 months of age, enabling a more detailed examination of developmental representation, pain contexts and reference standards, behavioural, physiological and neurophysiological modalities, and reuse of datasets across studies. The review will also include approaches using physiological and cortical signals and will assess methodological

quality using the PROBAST framework. This focused synthesis is intended to identify methodological and developmental gaps that may go unnoticed within broader reviews of pediatric AI applications.

Condition being studied The review will examine infant pain occurring in clearly defined painful or noxious clinical contexts, including acute procedural pain and postoperative pain. The review focuses on AI methods used to detect, classify, or predict pain from behavioural, physiological, neurophysiological, or multimodal data.

METHODS

Search strategy A comprehensive electronic search strategy was developed in consultation with an academic librarian at York University. Searches were conducted in MEDLINE (via Ovid), PubMed, PsycINFO, CINAHL, IEEE Xplore, ACM Digital Library, Scopus, and Web of Science. Because the review spans health sciences and computer science/engineering literature, databases from both disciplinary areas were included.

The search strategy combined three primary concepts: (1) neonates and infants, (2) pain or nociception, and (3) artificial intelligence, machine learning, or other automated/computational methods. Population terms included neonat*, newborn*, infant*, baby/babies, preterm infant*, and premature infant*. Pain-related terms included pain*, nocicept*, noxious, pain assessment, and pain measurement. Computational terms included artificial intelligence, machine learning, deep learning, supervised machine learning, support vector machine/SVM, convolutional neural network/CNN, neural network*, classifier*, automated, detection, and algorithm*. Additional modality-related terms were identified to capture the range of approaches used in infant pain assessment, including EEG and cortical activity; physiological measures such as heart rate, heart-rate variability, oxygen saturation, respiration, and skin conductance; behavioural measures such as facial expression, crying, vocalization, and body movement; and multimodal approaches.

Boolean operators were used to combine synonymous terms with OR within each concept and to combine the population, pain, and computational concepts using AND. Search syntax and controlled vocabulary were adapted for each database.

Participant or population Human neonates and infants from preterm birth up to 12 months of age will be included. Eligible samples may include

preterm infants, term neonates, older infants, or mixed neonatal/infant samples.

Intervention The method of interest is the application of AI to infant pain assessment, including machine learning, deep learning, computational models, computer vision, signal processing, and other algorithmic classification or prediction methods. Input infant pain response modalities may include facial expression or video, cry, electroencephalography, heart rate or electrocardiography, oxygen saturation, respiration, movement, skin conductance, other physiological signals, or multimodal combinations. Studies must apply an AI method to pain detection, classification, intensity estimation, or monitoring. Pain stimulus must be clearly identifiable.

Comparator Eligible discrimination comparisons may include pain versus no-pain conditions; pre-procedure versus intra-procedure or post-procedure periods; automated-model predictions versus validated observational pain scales, clinician or expert ratings, procedure-based labels, physiological thresholds, or another model.

Study designs to be included Any original empirical article or conference proceedings will be considered eligible for inclusion. Review articles, commentaries, case studies, and dissertations will not be included. Publications with insufficient methodological or outcome information after reasonable attempts to obtain clarification will also be excluded from analysis but listed.

Eligibility criteria Studies are eligible if they reported original empirical data; included human neonates or infants from preterm birth up to 12 months of age; examined pain in the context of a clearly defined painful or noxious clinical event or procedure; and applied an AI method to classify, detect, estimate, or predict pain. Eligible methods include machine learning, deep learning, computer vision, automated signal-processing approaches, and other algorithmic classification or prediction methods. Peer-reviewed journal articles and full conference proceedings are included. Review articles, commentaries, editorials, protocols, conference abstracts without a full manuscript, studies without a clearly identifiable pain-related condition, and studies without an automated pain-assessment method are excluded. Studies containing participants older than 12 months are eligible only when data for the eligible infant population could be separated or when the relevant pain-assessment sample fell within the eligible age range. No restrictions are placed on

publication year or country. Studies published in languages other than English are considered where the full text could be translated and adequately assessed.

Information sources The following databases were searched: Medline (Ovid), CINAHL, IEEE Xplore, Web of Science, PubMed, Scopus, ACM Digital Library, and PsycINFO. The reference lists of relevant articles are also manually inspected to ensure completeness.

Main outcome(s) The primary outcomes are the reported performance and validation characteristics of automated models for detecting, classifying, estimating, predicting, or continuously monitoring pain in neonates and infants. Performance measures will be extracted as appropriate to the modelling task and may include area under the receiver operating characteristic curve (ROC-AUC), area under the precision-recall curve (PR-AUC), accuracy, sensitivity/recall, specificity, precision/positive predictive value, negative predictive value, F1 score, confusion-matrix measures, regression or agreement metrics, and event-detection measures. The evaluation context associated with each performance estimate will also be extracted, including the validation strategy, unit of data splitting, whether testing was subject-independent or sample/window-independent, internal versus external validation, and confidence intervals or variability across folds or repeated runs where reported.

Additional outcome(s) Secondary outcomes will include study, participant, dataset, and modelling characteristics. These will include gestational, post-menstrual, and chronological age; prematurity status; pain context and procedure; number of infants and number of samples, windows, or pain events; data source and dataset provenance; reuse of datasets across studies; pain ground truth and label-generation procedures; data modality and acquisition method; temporal unit of analysis; preprocessing; feature extraction and feature selection; data augmentation; algorithm and model architecture; training procedures; hyperparameter tuning and model-selection procedures; handling of class imbalance and missing data; and ablation analyses, where reported.

Data management Records will be imported into Covidence for duplicate removal, title and abstract screening, and full-text screening. Screening will involve the lead author (LH) and three research assistants (BK, MI, SZ), with a proportion of records independently double-screened to assess agreement. Full-text exclusions will be reviewed by

the lead author (LH). Data will be extracted using a custom extraction form. Disagreements or uncertainties will be resolved through discussion and review by the senior author (RPR).

Quality assessment / Risk of bias analysis Risk of bias will be assessed using the Prediction model Risk of Bias Assessment Tool (PROBAST). Inter-rater reliability will be assessed, and disagreements will be resolved via discussion and consultation with the senior author (RPR) as needed.

Strategy of data synthesis A descriptive and narrative synthesis will be conducted. Study characteristics will be summarized using frequencies and percentages for categorical variables and appropriate descriptive statistics for continuous variables where possible. Findings will be synthesized according to participant characteristics, pain context, modality, computational approach, and reported model performance.

Subgroup analysis Subgroup narrative synthesis may be performed based on infants' age, pain context, assessment modality, and the type of automated or computational method.

Sensitivity analysis Sensitivity analyses will be conducted to assess whether the overall patterns and conclusions of the narrative and descriptive synthesis change by age and pain stimulus type.

Language restriction No language restrictions are imposed in the search strategy.

Country(ies) involved Canada.

Other relevant information Crispin AR, Mahendra Arifionang LD, Alfrino M, Sitompul CP, Simanullang IF. A systematic literature review of methods and technologies for detecting pain in infants through facial expression analysis. *International Journal of Management Science and Information Technology*. 2026;6(2):1303-1310. doi:10.35870/ijmsit.v6i2.7431.

Moons KGM, Damen JAA, Kaul T, et al. PROBAST+AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. *BMJ*. 2025;388:e082505. Published 2025 Mar 24. doi:10.1136/bmj-2024-082505

Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R. The PRISMA

2020 statement: An updated guideline for reporting systematic reviews. *BMJ*. 2021 Mar 29;372.

Wang Z, Hu J, Zeng J, et al. Artificial intelligence in pediatric pain: a systematic review. *BMC Med Inform Decis Mak*. 2026;26:303. doi:10.1186/s12911-026-03571-5.

Zamzmi G, Kasturi R, Goldgof D, Zhi R, Ashmeade T, Sun Y. A review of automated pain assessment in infants: features, classification tasks, and databases. *IEEE Rev Biomed Eng*. 2018;11:77-96. doi:10.1109/RBME.2017.2777907.

Zhou Q, Huang X, Lv Y, et al. Application of Intelligent Neonatal Pain Assessment Tools Based on Multimodal Data Fusion: A Systematic Review. *J Pain Res*. 2026;19:578721. Published 2026 Mar 27. doi:10.2147/JPR.S578721

Keywords Neonatal pain; infant pain; pain assessment; artificial intelligence; machine learning; deep learning; facial expression; cry analysis; electroencephalography; physiological monitoring; multimodal.

Dissemination plans Review findings will be published in a peer-reviewed journal and presented at academic conferences.

Contributions of each author

Author 1 - Lojain Hamwi - Study conception and design; data management; data extraction; data analysis; data interpretation; protocol and manuscript writing; manuscript submission.

Email: lhamwi@yorku.ca

Author 2 - Bahar Khosrojerdi - Data screening and extraction; manuscript editing.

Email: bahar87@my.yorku.ca

Author 3 - Susan Zahir - Data screening and extraction; manuscript editing.

Email: szahir@yorku.ca

Author 4 - Mahbod Ebrahimi - Data screening and extraction; manuscript editing.

Email: mahbod99@yorku.ca

Author 5 - Hossein Kassiri - Study design; data interpretation; manuscript editing.

Email: kassiri@yorku.ca

Author 6 - Rebecca Pillai Riddell - Study conception and design; data analysis; data interpretation; manuscript writing and editing; providing funding.

Email: rpr@yorku.ca