HHS Public Access

Author manuscript

Adv Exp Med Biol. Author manuscript; available in PMC 2025 February 20.

Published in final edited form as:

Adv Exp Med Biol. 2024 ; 1462: 499–512. doi:10.1007/978-3-031-64892-2_31.

Machine Learning in Pain Neuromodulation

Tessa Harland,
Department of Neurosurgery, Albany Medical College, Albany, NY, USA

Trish Elliott,
Department of Neurosurgery, University of Arizona College of Medicine, Tucson, AZ, USA

Ilknur Telkes,
Department of Biomedical Science, Charles E. Schmidt College of Medicine, Florida Atlantic University, Boca Raton, FL, USA

Julie G. Pilitsis
Department of Neurosurgery, University of Arizona College of Medicine, Tucson, AZ, USA

Abstract

This chapter highlights the intersection of pain neuromodulation and machine learning (ML), exploring current limitations in pain management and how ML techniques can address these challenges. Neuromodulation technologies, such as spinal cord stimulation (SCS), have emerged as promising interventions for chronic pain, but limitations such as patient selection have resulted in high rates of failure and costly removal of these devices. ML offers a powerful approach to augment pain management outcomes by leveraging predictive modeling for enhanced patient selection, adaptive algorithms for programming optimization, and identification of objective biomarkers for improved outcome assessment. This chapter discusses various ML applications in pain neuromodulation and how we can expect it to shape the future of the field. While ML holds great promise, challenges such as algorithm transparency, data quality, and generalizability must be addressed to fully realize its potential in revolutionizing pain management.

Keywords

Pain neuromodulation; Machine learning; Spinal cord stimulation; Candidate selection; Programming optimization; Biomarkers; Chronic pain

1 Background

1.1 Current Limitations in Pain Neuromodulation

Chronic pain is a widespread health issue affecting 20% of adults globally, which is more patients than cancer, diabetes, and heart disease combined [1]. The significant morbidity associated with it leads to some of the highest healthcare and societal costs of any disease. Addressing the significant burden of chronic pain and the ongoing opioid epidemic [2] has prompted ongoing advancements in pain management with the emergence of

jpilitsis@arizona.edu .

neuromodulation technologies as a possible solution. The International Neurostimulation Society defines therapeutic neuromodulation as “the alteration of nerve activity through targeted delivery of a stimulation, such as electrical stimulation or chemical agents to specific neurological sites in the body” [3]. While this has grown to include noninvasive techniques such as transcranial magnetic stimulation, surgically implanted devices such as spinal cord stimulation (SCS), peripheral nerve stimulation (PNS), and intrathecal drug delivery systems (IDDS) are a mainstay of pain neuromodulation and fall within the purview of neurosurgeons and pain physicians.

As one of the most common procedures for pain neuromodulation, SCS stands as a well-established, FDA-approved surgical intervention for individuals with treatment-resistant chronic pain. This procedure is primarily utilized for conditions such as post-laminectomy syndrome, chronic neuropathic pain, and complex regional pain syndrome (CRPS). These devices have been increasingly used over the last 5 years with newly improved conditions such as painful diabetic neuropathy [4]. Though successful in many patients, SCS treatment costs more than USD 40,000 per patient and fails at a rate of 10–20%, largely in part because the factors that best predict which individual patients are most likely to benefit remain unclear. Moreover, current methods of patient selection require an invasive trial of the device and fail in 25–30% of cases at a cost of thousands of dollars [5, 6]. The success of these pain management interventions is largely dependent on appropriate patient selection. Identification of the chronic pain phenotypes that may benefit from SCS could dramatically improve the outcomes of these devices for patients while reducing healthcare costs. Consequently, the optimization of patient selection is essential in improving the overall outcomes of these pain neuromodulation devices.

Another significant limitation of pain neuromodulation is the subjectivity of its clinical outcomes. The evaluation of pain relief often relies on self-reported scales such as the numerical rating scale (NRS) or visual analog scale (VAS). See Table 1 for a comprehensive list of commonly used pain scores. While these subjective methods are considered the gold standard for quantifying pain and the effectiveness of pain management therapies, immeasurable factors such as psychological and environmental contributions limit the reliability and accuracy of these measures. Further, different groups use different measures, and it is unclear whether these are equivalent. Most commonly, a holistic approach, including some of the measures in Table 1, is preferred. Often, these subjective measures determine responsiveness to therapies and are used to guide decision-making for permanent implantation of neuromodulation devices such as SCS. The identification of objectively measured features or phenotypes of pain that would allow for better assessment of pain outcomes is essential to improving and maximizing the positive clinical potential of these devices.

Machine learning (ML)-derived algorithms serve as a powerful tool that can be used to improve both clinical outcomes of current neuromodulation devices and the means that we measure their success. Briefly, ML constitutes a subset of artificial intelligence focused on constructing algorithms capable of learning patterns and making predictions from data. Refer to Table 2 for a brief description of commonly used ML techniques. In the context of neuromodulation for pain and the current limitations of the field, ML stands as a crucial

tool for advancing the therapeutic impact of these devices and their outcomes (Fig. 1). This chapter discusses the role of ML in pain neuromodulation with a specific focus on SCS, encompassing candidate selection, program optimization, and identification of biomarkers of pain.

2 The Role of Machine Learning in Candidate Selection

2.1 Use of Preoperative Features to Predict SCS Response

Traditionally, SCS candidate selection for pain-related conditions is determined based on diagnosis, prior treatment response, lack of candidacy for alternative surgical procedures, psychological contribution to pain, drug use, and reasoning for stimulator placement [7]. Prior to SCS placement, all patients undergo psychological testing and a short trial period of stimulation that involves percutaneous placement of leads connected to an external generator. If the patient receives a greater than 50% reduction in NRS during the trial, they are typically considered candidates for permanent placement. Despite these guidelines, patient selection remains difficult for even experienced providers and requires consideration of a large number of parameters [8, 9]. As previously discussed, the overall SCS failure rates are unacceptably high, with great financial and societal costs suggesting that improved methods of patient selection are imperative for improvement of pain neuromodulation outcomes.

Supervised ML-derived algorithms are being increasingly applied to pain management therapies to define the patient phenotypes that will respond and to create predictive models [10–14]. In one such study, a predictive model was developed utilizing classification and regression trees (CART) in patients who had previously failed standard SCS but responded positively to high-dose SCS, a salvage waveform [12]. They found that pain intensity scores, medication usage, paresthesia coverage for back pain, and EQ5D scores for leg pain served as predictive factors for high-dose SCS response after 12 months [12]. Despite identifying these predictive features, the model lacked internal and external validation, a critical step to avoid overfitting of the data and low performance when applied to an external cohort. In another study, a series of ML-based algorithms were used to predict high-frequency stimulation (HFS) SCS responders with a specificity and sensitivity of 90% but with limited predictive performance of 58.3% [13]. The study was further limited by utilization of a single split, with 80% of data used to train the model and 20% used to test the model only once [13]. The single split may not adequately capture the variability present in the dataset, potentially compromising the model’s performance.

Unsupervised ML-derived algorithms have also been used to assess the validity of evaluating patients naïve to SCS and those previously treated with SCS [15]. Goudman et al. used this technique on a retrospective cohort of 263 persistent spinal pain syndrome type 2 patients (185 individuals naïve to neurostimulation and 78 who had previously undergone SCS). Through analysis of pain intensity scores, health-related quality of life metrics, medication usage, and functional indicators, the study demonstrated that model-based clustering could not distinctly separate the two groups [15]. This suggests that the clinical profiles of SCS-naïve and previously treated individuals may be substantially similar, thus challenging the

conventional approach of evaluating them as distinct populations when assessing the success of SCS.

Our group recently developed a predictive model for long-term SCS response using a nuanced method that combines unsupervised and supervised approaches in an effort to avoid challenges inherent to ML applications on limited datasets [14]. We first used unsupervised ML techniques to identify specific subgroups or clusters within our dataset using K means. Here, we identified two distinct groups of patients that differed significantly in age, duration of chronic pain, initial NRS, and preoperative Pain Catastrophizing Scale (PCS) scores [14]. Supervised ML techniques were then applied to each cluster to develop a predictive algorithm for responders. Using the most predictive features, the performance of ML models was evaluated using a nested cross-validation scheme. Nested cross-validation is a technique used to evaluate machine learning models, particularly in scenarios where model performance estimation is crucial and data is limited [16]. It involves a nested loop structure: an outer loop segments the dataset into training and validation sets multiple times, while an inner loop fine-tunes model parameters on the training set and validates them on the inner validation set. This method provides a more robust estimate of model performance by reducing overfitting and offering a reliable evaluation of how a model would perform on unseen data. In future iterations of this model, we plan to incorporate more objective- and imaging-based features to further improve its predictive ability and to effectively identify the most meaningful features in SCS response.

Other research groups have already worked to integrate potential imaging biomarkers into predictive ML algorithms for assessing SCS response. De Andres et al. investigated the use of imaging biomarkers, specifically functional connectivity and brain volumetry, to predict SCS efficacy in post-laminectomy syndrome patients [17]. This prospective observational study used resting-state functional MRI and region of interest analysis to assess for volume and connectivity changes at baseline and at 6 months after SCS implantation. Among the 24 patients, seven were classified as such, defined by a minimum of ten-unit improvement on the pain detection questionnaire (PD-Q) at 6 months. Leveraging ML approaches, particularly linear discriminant analysis (LDA), they identified that responders had statistically significant volumetric differences in the left putamen and connectivity differences between the left planum temporale and right insular cortex when compared to nonresponders [17]. They ultimately demonstrated that LDA of clinical variables with the addition of the significant imaging biomarkers increased long-term prediction from 29% in current practice to 96% [17].

Another research group has built predictive SVM models using a high number of electroencephalography (EEG) features and principal component analysis (PCA) as a dimension reduction approach to distinguish pain phenotypes: healthy controls versus chronic pain patients with high pain (defined as VAS scores greater than 7/10 in average) [18]. Particularly in their three-way classification SVM, Levitt et al. first developed a binary SVM model to predict no-pain (healthy controls) versus high-pain group without SCS (chronic lumbar radiculopathy) and used the same PCA coefficients in the latter model following the same tenfold cross-validation approach. Performance analysis demonstrated a cross-validation accuracy of 71.9%, where chronic pain patients who were candidates

for thoracic SCS were distinguished accurately with a receiver operating characteristic-area under curve (ROC-AUC) of 0.96. This is particularly important because decisions regarding eligibility for SCS implants are based on subjective criteria.

To predict surgical success in patients who underwent SCS implantation surgery for their chronic pain, Adil et al. investigated how well a logistic regression (LR) model with recursive feature elimination (RFE) and deep neural networks (DNNs) could predict reduction in opioid use 1 year after SCS implantable pulse generator (IPG) surgery [19]. Researchers used a feature matrix composed of 30 predictors, including sex, opioid doses in milligram morphine equivalents (MME) with eight categories, diagnosis with six categories, and more than 17 comorbidities. Surgery was accepted as "successful" if the opioid doses were reduced or stayed stable over the course of the study period. The linear model, LR with parameter optimization (RFE) process, gave the highest prediction accuracy (ROC-AUC, 0.74; confidence interval (CI), 95%, 0.72–0.75) using the 5 most important variables out of 30. These key features were significantly associated with opioid use (e.g., pre-SCS opioid dose and long-term opioid usage) rather than comorbidities (e.g., hypertension, psychoses, and depression), suggesting the crucial role of pre-surgical measures and pharmacological patterns in outcome prediction. Reducing the total number of variables from 30 to 5 has a practical use where it is easier to interpret for clinicians, reduces complexity, and facilitates straightforward implementation in clinical decision-making. From an ML perspective, a model with fewer variables may require less data for training and validation, making it more practical in situations where comprehensive datasets are challenging to obtain.

2.2 Use of Machine Learning to Replace Lead Trialing

Prior to permanent SCS placement, patients undergo trial implantation to assess candidacy and determine parameters for pain relief, relying on subjective processes that are time-consuming and costly, yet essential in current practice and required for insurance coverage. However, recent studies have shown the potential for ML to enhance or even replace these trial processes [20]. Ounajim et al. employed an ML-assisted analysis to compare the predictive capabilities of traditional trialing versus ML algorithms in assessing the long-term efficacy of SCS in patients [21]. They utilized primary outcomes such as pain intensity measured by VAS, Oswestry Disability Index (ODI) scores, EuroQol with five dimensions and three levels (EQ5D-3 L), and the Montgomery-Asberg Depression Rating Scale (MADRS). Through PCA on the percent changes in these scores, they established a Global Health Improvement Score (GHIS) to classify patients as responders (GHIS >0) or nonresponders (GHIS <0). ML models were trained using various patient features, including demographics, BMI, pain characteristics, and duration, such as linear regression, regularized logistic regression (RLR), naïve Bayes classifier, artificial neural networks, support vector machines, classification and regression trees, random forests, and gradient-boosted trees [21]. External validation and model assessment revealed that RLR achieved an area under the curve (AUC) of 81% with a sensitivity of 83.3%. The ML-based models consistently exhibited superior predictive power for SCS outcomes compared to traditional trial-based approaches [21]. The incorporation of ML into candidate selection processes, especially regarding demographic considerations, presents a promising avenue to augment neurosurgeons' decisions and improve SCS patient outcomes.

3 The Role of Machine Learning in Programming Optimization

Traditional SCS applies stimulation to the dorsal column with fixed parameters based on amplitude, frequency, and pulse width [22]. This method, termed open-loop, lacks adaptation to the ongoing physiological changes within the patient and spinal cord. In contrast, closed-loop systems dynamically adjust simulation parameters in real time, leveraging control systems and, in more advanced systems, ML algorithms. While various studies highlight the potential of closed-loop SCS for greater pain reduction and improved quality of life in chronic back and leg pain patients [23], the formal application of ML in adaptive SCS remains unexplored. This is despite the demonstration of superior efficacy and precision compared to non-ML closed-loop systems in other neurostimulation therapies, notably DBS [24, 25]. Within pain neuromodulation and SCS, positional changes in the body [26–28] and evoked compound action potentials (ECAPs) [22, 23] have emerged as effective inputs for enhancing pain relief in SCS patients. Recent research underscores the potential of these inputs to be effectively processed using ML-based classifiers [29, 30]. Here, we discuss the role of these variables in closed-loop SCS and how this could be augmented with use of ML.

3.1 Posture-Controlled SCS

Automatic accelerometer-controlled stimulation adjustments in SCS involve a system that adapts stimulation parameters based on real-time feedback from patient movement. Traditional SCS operates in fixed settings, but patients often experience variations in paresthesia perception due to posture changes, leading to overstimulation or understimulation. Previous studies have explored the use of accelerometers to monitor patient movement and adjust SCS stimulation parameters accordingly [26, 27]. Consistently, patients preferred automatically adjusted stimulation over manual adjustment rates [26] while demonstrating that position-adaptive stimulation significantly reduced pain and improved convenience compared to manual programing [27]. Patients reported increased comfort during position changes, better activity, and improved sleep [26].

The application of ML-derived algorithms to accelerometer-based closed-loop SCS could further enhance adaptability. A systematic review by Narayanan et al. highlighted the accuracy of supervised ML techniques in interpreting accelerometer data to estimate physical activity components or sleep patterns [29]. Such ML-driven interpretation could enable tailoring SCS settings based on a patient’s lifestyle, potentially improving outcomes by providing better adaptation to changing needs. Analysis of 53 studies and 31 different ML models (the most common of which were vector machines, random forests, and artificial neural networks) revealed that 80% of studies demonstrated an accuracy of at least 85%, and 45% of studies achieved an accuracy greater than 95% [29]. The incorporation of ML to improve the accuracy of variable interpretation could lead to improved adaptability of posture-controlled SCS and further improve patient outcomes.

3.2 ECAP-Controlled SCS

ECAPs represent the summation of neural activity produced by a group of axons in response to an electrical pulse. It has been used in SCS to assess efficacy by measuring

the degree of neural response, providing real-time feedback that helps adjust and optimize stimulation parameters. The morphology of the ECAP, including amplitude, latency, and configuration, can be used to classify the physiologic effect of SCS and inform the necessary adjustments of stimulation parameters. The EVOKE trial compared the effectiveness of ECAP-controlled, closed-loop SCS against fixed-output, open-loop SCS. In this randomized trial, participants with chronic back and leg pain were assessed for pain reduction and holistic treatment response. Results indicated that after 36 months, a higher percentage of participants experienced a reduction of at least 50% in overall back and leg pain with closed-loop SCS (77.6%) compared to open-loop SCS (49.3%). Similarly, a greater proportion reported at least an 80% reduction in pain intensity with closed-loop SCS (49.3%) compared to open-loop SCS (31.3%). The closed-loop SCS group showed better outcomes in holistic treatment response (44.8% vs. 28.4%) and achieved greater neural activation and therapy delivery accuracy without significant differences in adverse events.

ML offers a means to further enhance the performance and efficacy of closed-loop SCS. In a study by Koh et al., researchers investigated the potential of using multi-contact nerve cuff electrodes to extract spatiotemporal signatures from peripheral neural signals in rats [30]. These signals were evoked in the sciatic nerve's different fascicles using mechanosensory stimuli. Natural ECAPs were then classified by employing ML-derived algorithms and evaluating their performance based on classification accuracy, F1 score, and the ability to reconstruct the original firing rates of neural pathways. The best-performing ML-based classifier achieved a mean accuracy of 0.686, an F1 score of 0.605, and a correlation coefficient of 0.728 between original and estimated firing rates [30]. These results underscore the potential of ML techniques in accurately classifying naturally evoked CAPs from peripheral neural signals, suggesting their promising application in enhancing closed-loop SCS programming by effectively extracting and classifying evoked neural responses.

3.3 Smart Neuromodulation

Smart neuromodulation represents the next frontier in advancing SCS and pain neuromodulation. While closed-loop systems dynamically modify stimulation based on real-time feedback from input such as posture changes or ECAPs, adaptive systems involve the ability of a device to learn from a patient's response to stimulation and modify its behavior over time based on gathered data or experience. By integrating diverse data sources, including patient demographics, physiological signals, imaging data, and lifestyle information, smart neuromodulation systems enable a holistic analysis and aim to provide more personalized pain management for individuals with an increased likelihood of improved outcomes. Closed-loop systems have already demonstrated potential to improve patient outcomes. Expanding on this, the application of ML to closed-loop SCS is a promising direction that offers a means to finely tailor stimulation parameters to individual patient needs. By leveraging ML-driven interpretation of data, adaptive SCS could provide personalized and responsive pain management, marking a transformative step toward improved patient outcomes with pain neuromodulation.

4 Machine Learning Techniques to Identify Biomarkers of Pain

In an effort to identify objective indicators of pain, several studies have attempted to define physiological and radiographic correlates of pain. For example, prior imaging-based investigations have demonstrated the activation of specific cortical and subcortical areas during pain episodes [31], while resting-state functional connectivity (rsFC) studies have also demonstrated variations in neural connectivity that correlated with pain severity [32]. Similarly, electroencephalography data have demonstrated peak alpha frequency recorded at the temporal scalp correlates with pain severity [33], and spontaneous oscillations are reduced with pain [34]. Although these studies have added to our understanding of pain's neural underpinnings, a definitive, reliable biomarker of pain remains elusive. To address this gap, ML-derived approaches have emerged as promising approaches to delineate pain features and objectively define the pain experience.

A recent study by Lamichhane et al. utilized ML to explore the brain's reorganization in chronic low back pain (LBP) patients [35]. Structural and resting-state functional MRI data, alongside clinical scores, were collected from 24 LBP patients and 27 matched controls. The study revealed distinct cortical thickness differences and altered functional connectivity in LBP patients, notably in visual networks and key cortical hubs within motor and visual processing regions [35]. Their model, using cortical thickness data, achieved 74.51% accuracy in distinguishing LBP subjects. In a related effort, brain structural data from MRI of 131 healthy controls, particularly cortical thickness, successfully predicted physical pain thresholds [36]. The model's robustness (Pearson $r = 0.36$, $P < 0.0002$, $R2 = 0.13$) remained specific to pain thresholds and was not influenced by potential confounders such as anxiety- or center-related effects. The analysis identified specific brain regions, including the rostral anterior cingulate gyrus and parahippocampal gyrus, which inversely correlated with pain sensitivity [36].

Similarly, a study by Fernandez Rojas et al. used ML-derived techniques to analyze functional near-infrared spectroscopy (fNIRS) to identify potential biomarkers of pain [37]. In this investigation, 18 participants underwent thermal pain testing involving varying temperature levels and pain intensities. After extensive feature extraction of 69 features and training across multiple models, SVM achieved the highest predictive power with an accuracy of 94.17% in categorizing four types of pain using only 25 features [37]. It is important to note that these studies treat pain as a dichotomous variable and did not consider pain severity.

While these imaging-based studies show promise in the identification of a biomarker for pain, some believe that structural-based markers limit temporal resolution and do not fully capture the dynamic nature of pain perception and experience. Hence, to address this limitation, some research groups have shifted their focus toward exploring more cost-effective and less invasive alternatives such as EEG. In one such study, Furman et al. examined alpha brainwave frequency and pain intensity during capsaicin-heat-induced pain in 44 subjects [38]. They found that the alpha frequency measured in a pain-free period before prolonged pain correlated with subsequent pain reports. A slower initial alpha frequency at this stage predicted higher pain during the prolonged pain condition.

Additionally, the degree of alpha frequency decreases between pain-free and prolonged pain states correlated with pain intensity [38]. Similar predictive abilities have been demonstrated across multiple studies using ML-based techniques applied to EEG data, demonstrating accuracies ranging from 62 to 100% in predicting pain intensity [39]. While these findings suggest the potential of ML to reliably predict pain-related outcomes, challenges related to inadequate reporting and potential bias observed in the reviewed studies underscore the need for improved reporting standards and external validation of models to enhance their reliability and facilitate practical translation into clinical settings [39].

These biomarkers are not limited to identification of pain but can also be used to define treatment response. In a study from our group, we compared recorded EEG from chronic pain patients during SCS surgery to understand neural patterns and their link to stimulation-induced pain relief [40]. We found that 10-KHz HFS showed stronger alpha power in somatosensory areas and an increased alpha/theta peak power ratio in the frontal cortex compared to 60-Hz tonic stimulation. HFS induced a shift from theta to alpha rhythms, unlike baseline and tonic stimulation, which maintained slower theta activity [40]. Moreover, a positive correlation was observed between changes in ODI scores and HFS-induced alpha/theta peak power ratio in frontal and somatosensory regions. These results suggest that changes in theta-alpha band interactions and their spatial distribution could be early neural indicators of pain relief induced by HFS in chronic pain. We expect that an increased understanding of objective markers of pain relief will enable improved assessment and subsequent optimization of pain neuromodulation therapies.

Wearable technology, such as smartwatches, allows for continuous and objective monitoring of physiological and behavioral data in real-world settings. This provides a more accurate representation of an individual's pain experience compared to sporadic assessments in a clinical setting. However, compared to the wide use of wearable technology in other fields, such as assessment of Parkinson's disease progression [41] or evaluation of DBS [42], little research has been done in SCS field in terms of investigating the potential chronic pain biomarkers. Especially within the framework of extensive datasets generated by these technologies, incorporation of ML methodologies becomes imperative. In a recent study, Patterson et al. demonstrated that physiological and behavioral variables measured via smartwatch in chronic pain patients from baseline to 6 months post-implantation of SCS were able to predict sleep disturbance outcomes (based on PROMIS-29) with the highest accuracy (95% with F1 score of 0.92 ± 0.001) [43]. Similarly, Heros et al. showed that RF classification model could predict the NRS outcomes (three levels: mild, moderate, and severe) with an accuracy of 78% (F1 score = 0.75 ± 0.07) when only objective measures captured via wearable devices were implemented [44]. The accuracy of the model increased to 81% (F1 score = 0.80 ± 0.07) when the objective features were combined with PROMs. In both studies, heart rate and step count were the prominent predictive measures.

Given that wearable sensors facilitate longitudinal tracking of pain-related variables over extended periods, these can reveal trends, fluctuations, and patterns in pain experiences, contributing to a deeper understanding of chronic pain conditions and real-time feedback on neuromodulation-induced changes. ML provides a powerful tool that will continue to advance our understanding of pain and its management. Its capacity to detect patterns

and analyze large datasets enables the objective assessment of EEG, neuroimaging, rsFC, neural activity, and other forms of data in a meaningful manner that objectifies pain into a quantitative measurement free of subjectivity. Ultimately, the use of objectively measured biomarkers enables improved accuracy in assessing pain outcomes and classification of responders from nonresponders to neuromodulation therapy. In the future, pain management will evolve into a more individualized process, leveraging distinct bio-derived pain signatures to guide tailored therapeutic interventions.

5 Limitations of Machine Learning for Pain Neuromodulation

ML algorithms have inherent limitations, often labeled as “black box” models due to their opaque rationale and calculations behind generated outputs. Additionally, many studies are confined to single institutions and limited subject pools, potentially restricting the applicability of findings to broader pain populations. Small datasets can exacerbate the “accuracy paradox” in ML, where imbalanced outcome classes (e.g., more nonresponders than responders) artificially inflate predictive performance by predominantly classifying cases as nonresponders. To mitigate these issues, it is crucial to expand research beyond single institutions, integrating new external datasets to enhance algorithm generalizability and predictive accuracy.

6 Summary

The rise of ML-derived techniques represents a paradigm shift within pain medicine and neuro-modulation. With its capacity for processing large and complex datasets, ML provides innovative tools poised to reshape the current construct of pain management. These advancements encompass the potential to objectify pain, identify distinct pain phenotypes, and predict patient outcomes, thereby refining the selection process for tailored treatment modalities. However, it is crucial to underscore that effective utilization of ML hinges on the availability of high-quality datasets and a clear understanding of methodological nuances for maximizing its application potential.

Disclosures

Dr. Telkes has grant support from NIH R00NS119672 and FAU COECS/I-SENSE.

Dr. Pilitsis receives grant support from Medtronic, Boston Scientific, Abbott, NIH 2R01CA166379, NIH R01EB030324, and NIH U44NS115111. She is the medical advisor for Aim Medical Robotics and has stock equity. RTA receives fellowship support from Medtronic, Abbott, and Boston Scientific.

References

1. Zajacova A, Grol-Prokopczyk H, Zimmer Z. Pain trends among American adults, 2002–2018: patterns, disparities, and correlates. Demography. 2021;58(2):711–38. 10.1215/00703370-8977691. [PubMed: 33834222]
2. Leung N, Tsourmas NF, Yuspeh L, et al. Increased spinal cord stimulator use and continued opioid treatment among injured workers: a regional pilot study. J Occup Environ Med. 2020;62(8):e436–41. 10.1097/jom.0000000000001933. [PubMed: 32541622]
3. Mekhail NA, Cheng J, Narouze S, Kapural L, Mekhail MN, Deer T. Clinical applications of neurostimulation: forty years later. Pain Pract. 2010;10(2):103–12. 10.1111/j.1533-2500.2009.00341.x. [PubMed: 20070547]

4. Petersen EA, Stauss TG, Scowcroft JA, et al. Effect of high-frequency (10-kHz) spinal cord stimulation in patients with painful diabetic neuropathy: a randomized clinical trial. JAMA Neurol. 2021;78(6):687–98. 10.1001/jamaneurol.2021.0538. [PubMed: 33818600]

5. Brinzeu A, Cuny E, Fontaine D, et al. Spinal cord stimulation for chronic refractory pain: long-term effectiveness and safety data from a multicentre registry. Eur J Pain (London, England). 2019;23(5):1031–44. 10.1002/ejp.1355.

6. Nissen M, Ikäheimo TM, Huttunen J, Leinonen V, von Und Z, Fraunberg M. Long-term outcome of spinal cord stimulation in failed Back surgery syndrome: 20 years of experience with 224 consecutive patients. Neurosurgery. 2019;84(5):1011–8. 10.1093/neuros/nyy194. [PubMed: 29788145]

7. Lee AW, Pilitsis JG. Spinal cord stimulation: indications and outcomes. Neurosurg Focus. 2006;21(6):E3. 10.3171/foc.2006.21.6.6.

8. Olson KA, Bedder MD, Anderson VC, Burchiel KJ, Villanueva MR. Psychological variables associated with outcome of spinal cord stimulation trials. Neuromodulation. 1998;1(1):6–13. 10.1111/j.1525-1403.1998.tb00025.x. [PubMed: 22150881]

9. Hussaini SMQ, Murphy KR, Han JL, et al. Specialty-based variations in spinal cord stimulation success rates for treatment of chronic pain. Neuromodulation. 2017;20(4):340–7. 10.1111/ner.12582. [PubMed: 28370989]

10. Alexander J Jr, Edwards RA, Manca L, et al. Integrating machine learning with microsimulation to classify hypothetical, novel patients for predicting Pregabalin treatment response based on observational and randomized data in patients with painful diabetic peripheral neuropathy. Pragmat Observ Res. 2019;10:67–76. 10.2147/por.S214412.

11. Azimi P, Benzel EC, Shahzadi S, Azhari S, Mohammadi HR. Use of artificial neural networks to predict surgical satisfaction in patients with lumbar spinal canal stenosis: clinical article. J Neurosurg Spine. 2014;20(3):300–5. 10.3171/2013.12.SPINE13674. [PubMed: 24438428]

12. De Jaeger M, Goudman L, Brouns R, et al. The long-term response to high-dose spinal cord stimulation in patients with failed Back surgery syndrome after conversion from standard spinal cord stimulation: an effectiveness and prediction study. Neuromodulation. 2020; 10.1111/ner.13138.

13. Goudman L, Van Buyten JP, De Smedt A, et al. Predicting the response of high frequency spinal cord stimulation in patients with failed Back surgery syndrome: a retrospective study with machine learning techniques. J Clin Med. 2020;9(12) 10.3390/jcm9124131.

14. Hadanny AHT, Khazen O, Dimarizo M, Marchese A, Telkes I, Sukul V, Pilitsis JG. Development of machine learning based models to predict treatment response to spinal cord stimulation. Neurosurgery. 2021;

15. Goudman L, Rigoard P, Billot M, et al. Spinal cord stimulation-Naïve patients vs patients with failed previous experiences with standard spinal cord stimulation: two distinct entities or one population? Neuromodulation. 2023;26(1):157–63. 10.1016/j.neurom.2022.04.037. [PubMed: 35551868]

16. Vabalas A, Gowen E, Poliakoff E, Casson AJ. Machine learning algorithm validation with a limited sample size. PLoS One. 2019;14(11):e0224365. 10.1371/journal.pone.0224365. [PubMed: 31697686]

17. De Andres J, Ten-Esteve A, Harutyunyan A, et al. Predictive clinical decision support system using machine learning and imaging biomarkers in patients with Neurostimulation therapy: a pilot study. Pain Physician. 2021;24(8):E1279–e1290. [PubMed: 34793655]

18. Levitt J, Edhi MM, Thorpe RV, et al. Pain phenotypes classified by machine learning using electroencephalography features. NeuroImage. 2020;223:117256. 10.1016/j.neuroimage.2020.117256. [PubMed: 32871260]

19. Adil SM, Charalambous LT, Rajkumar S, et al. Machine learning to predict successful opioid dose reduction or stabilization after spinal cord stimulation. Neurosurgery. 2022;91(2):272–9. 10.1227/neu.0000000000001969. [PubMed: 35384918]

20. Eldabe S, Duarte RV, Gulve A, et al. Does a screening trial for spinal cord stimulation in patients with chronic pain of neuropathic origin have clinical utility and

cost-effectiveness (TRIAL-STIM)? A randomised controlled trial. Pain. 2020;161(12):2820–9. 10.1097/j.pain.0000000000001977. [PubMed: 32618875]

21. Ounajim A, Billot M, Goudman L, et al. Machine learning algorithms provide greater prediction of response to SCS than Lead screening trial: a predictive AI-based Multicenter study. J Clin Med. 2021;10(20) 10.3390/jcm10204764.
22. Vallejo R, Chakravarthy K, Will A, Trutnau K, Dinsmoor D. A new direction for closed-loop spinal cord stimulation: combining contemporary therapy paradigms with evoked compound action potential sensing. J Pain Res. 2021;14:3909–18. 10.2147/jpr.S344568. [PubMed: 35002310]
23. Mekhail N, Levy RM, Deer TR, et al. Long-term safety and efficacy of closed-loop spinal cord stimulation to treat chronic back and leg pain (Evoke): a double-blind, randomised, controlled trial. Lancet Neurol. 2020;19(2):123–34. 10.1016/s1474-4422(19)30414-4. [PubMed: 31870766]
24. Merk T, Peterson V, Köhler R, Haufe S, Richardson RM, Neumann WJ. Machine learning based brain signal decoding for intelligent adaptive deep brain stimulation. Exp Neurol. 2022;351:113993. 10.1016/j.expneurol.2022.113993. [PubMed: 35104499]
25. Sand D, Rappel P, Marmor O, et al. Machine learning-based personalized subthalamic biomarkers predict ON-OFF levodopa states in Parkinson patients. J Neural Eng. 2021;18(4) 10.1088/1741-2552/abfc1d.
26. Schade CM, Schultz DM, Tamayo N, Iyer S, Panken E. Automatic adaptation of neurostimulation therapy in response to changes in patient position: results of the posture responsive spinal cord stimulation (PRS) research study. Pain Physician. 2011;14(5):407–17. [PubMed: 21927044]
27. Schultz DM, Webster L, Kosek P, Dar U, Tan Y, Sun M. Sensor-driven position-adaptive spinal cord stimulation for chronic pain. Pain Physician. 2012;15(1):1–12. [PubMed: 22270733]
28. Wolf EW 2nd. Dynamic detection of spinal cord position during postural changes using near-infrared reflectometry. Neuromodulation. 2015;18(6):448–59; discussion 459. 10.1111/ner.12319. [PubMed: 26095007]
29. Narayanan A, Desai F, Stewart T, Duncan S, Mackay L. Application of raw accelerometer data and machine-learning techniques to characterize human movement behavior: a systematic scoping review. J Phys Act Health. 2020;17(3):360–83. 10.1123/jpah.2019-0088. [PubMed: 32035416]
30. Koh RGL, Nachman AI, Zariffa J. Classification of naturally evoked compound action potentials in peripheral nerve spatiotemporal recordings. Sci Rep. 2019;9(1):11145. 10.1038/s41598-019-47450-8. [PubMed: 31366940]
31. Mouraux A, Iannetti GD. The search for pain biomarkers in the human brain. Brain. 2018;141(12):3290–307. 10.1093/brain/awy281. [PubMed: 30462175]
32. Hemington KS, Wu Q, Kucyi A, Inman RD, Davis KD. Abnormal cross-network functional connectivity in chronic pain and its association with clinical symptoms. Brain Struct Funct. 2016;221(8):4203–19. 10.1007/s00429-015-1161-1. [PubMed: 26669874]
33. Nir RR, Sinai A, Moont R, Harari E, Yarnitsky D. Tonic pain and continuous EEG: prediction of subjective pain perception by alpha-1 power during stimulation and at rest. Clin Neurophysiol. 2012;123(3):605–12. 10.1016/j.clinph.2011.08.006. [PubMed: 21889398]
34. Peng W, Hu L, Zhang Z, Hu Y. Changes of spontaneous oscillatory activity to tonic heat pain. PLoS One. 2014;9(3):e91052. 10.1371/journal.pone.0091052. [PubMed: 24603703]
35. Lamichhane B, Jayasekera D, Jakes R, et al. Multimodal biomarkers of low back pain: a machine learning approach. Neuroimage Clin. 2021;29:102530. 10.1016/j.nicl.2020.102530. [PubMed: 33338968]
36. Kotikalapudi R, Kincses B, Zunhammer M, et al. Brain morphology predicts individual sensitivity to pain: a multicenter machine learning approach. Pain. 2023;164(11):2516–27. 10.1097/j.pain.0000000000002958. [PubMed: 37318027]
37. Fernandez Rojas R, Huang X, Ou KL. A machine learning approach for the identification of a biomarker of human pain using fNIRS. Sci Rep. 2019;9(1):5645. 10.1038/s41598-019-42098-w. [PubMed: 30948760]
38. Furman AJ, Meeker TJ, Rietschel JC, et al. Cerebral peak alpha frequency predicts individual differences in pain sensitivity. NeuroImage. 2018;167:203–10. 10.1016/j.neuroimage.2017.11.042. [PubMed: 29175204]

39. Mari T, Henderson J, Maden M, Nevitt S, Duarte R, Fallon N. Systematic review of the effectiveness of machine learning algorithms for classifying pain intensity, phenotype or treatment outcomes using electroencephalogram data. J Pain Offi J Am Pain Soc. 2022;23(3):349–69. 10.1016/j.jpain.2021.07.011.

40. Telkes L, Hancu M, Paniccioli S, et al. Differences in EEG patterns between tonic and high frequency spinal cord stimulation in chronic pain patients. Clin Neurophysiol. 2020;131(8):1731–40. 10.1016/j.clinph.2020.03.040. [PubMed: 32504934]

41. Li P, van Wezel R, He F, Zhao Y, Wang Y. The role of wrist-worn technology in the management of Parkinson's disease in daily life: a narrative review. Front Neuroinform. 2023;17:1135300. 10.3389/fninf.2023.1135300. [PubMed: 37124068]

42. Habets JGV, Heijmans M, Kuijf ML, Janssen MLF, Temel Y, Kubben PL. An update on adaptive deep brain stimulation in Parkinson's disease. Mov Disord. 2018;33(12):1834–43. 10.1002/mds.115. [PubMed: 30357911]

43. Patterson DG, Wilson D, Fishman MA, et al. Objective wearable measures correlate with self-reported chronic pain levels in people with spinal cord stimulation systems. NPJ Digit Med. 2023;6(1):146. 10.1038/s41746-023-00892-x. [PubMed: 37582839]

44. Heros R, Patterson D, Huygen F, et al. Objective wearable measures and subjective questionnaires for predicting response to neurostimulation in people with chronic pain. Bioelectron Med. 2023;9(1):13. 10.1186/s42234-023-00115-4. [PubMed: 37340467]

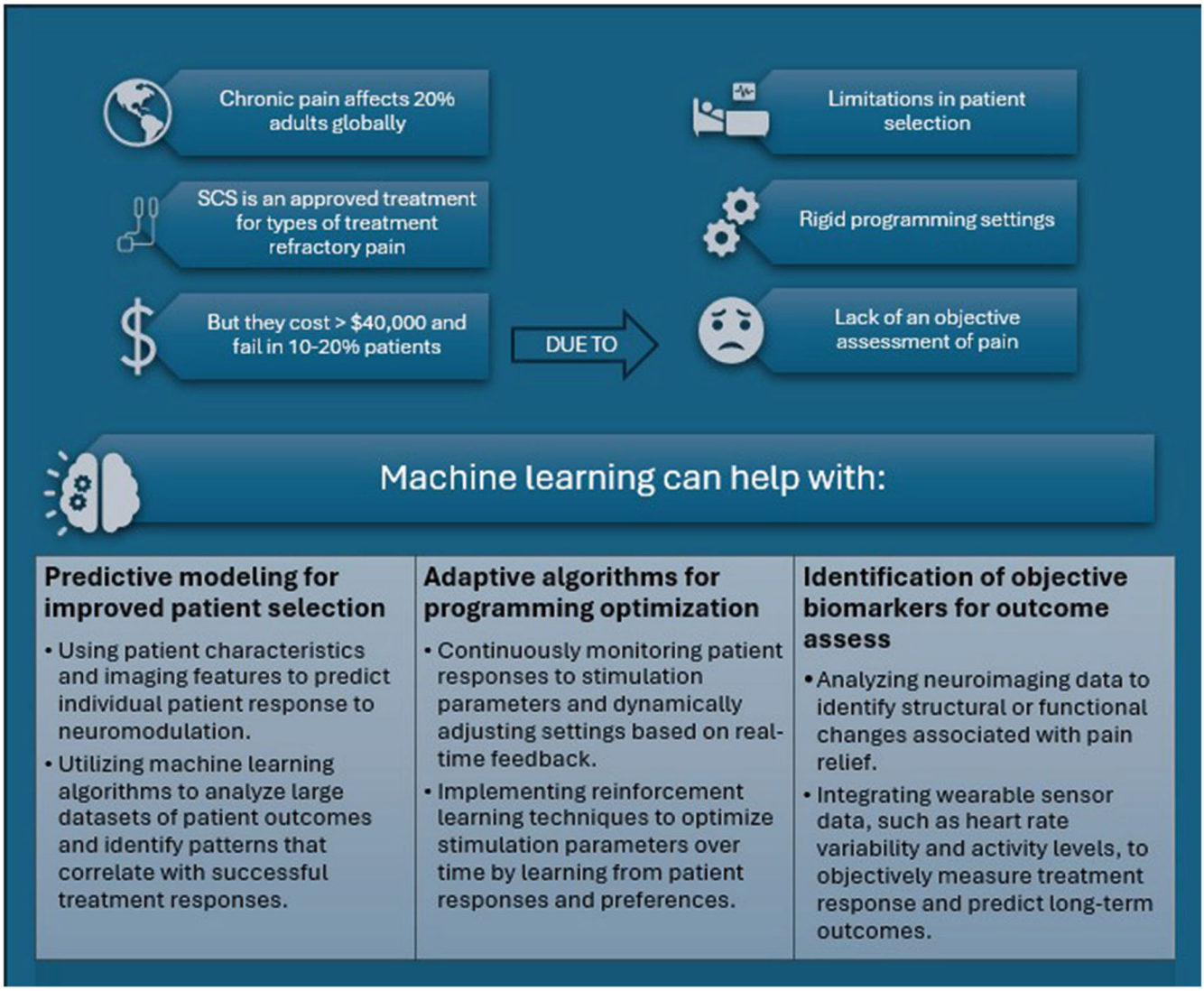

Fig. 1.
Summary figure of pain neuromodulation and machine learning

Table 1

Summary of commonly used pain assessment tools

Assessment tool	Primary focus	Description
Visual analog scale (VAS)	Sensory	A 10 cm line that patients mark anywhere from 0 (no pain) to 10 (pain as bad as it could possibly be) to measure their pain intensity
Numeric rating scale (NRS)	Sensory	Patients circle a number between 0 and 10, 0 and 20, or 0 and 100 to best describe their pain intensity. 0 representing “no pain at all” and the highest number representing “the worst pain ever possible”
McGill pain questionnaire	Sensory	Patients choose word descriptions tied to a numerical point system, resulting in the measurement of present and changing pain intensity, which may be used to determine intervention effectiveness
Percent pain relief (PPR)	Sensory	Measures patients’ pain relief over time as a percentage
Neuropathic pain symptom inventory (NPSI)	Sensory	Patients answer questions to assess five dimensions of neuropathic pain: Burning spontaneous, pressing spontaneous, paroxysmal, evoked, and paresthesia/dysesthesia, with a numerical score recorded for each subset and the total neuropathic pain
painDETECT questionnaire	Sensory	Patients answer seven sensory descriptor questions (never to very strongly) and two questions referring to spatial and temporal pain characteristics in order to determine whether the pain is nociceptive or neuropathic, with a score of 19 or above as likely neuropathic pain
Medication usage	Sensory	Measurement of the amount of medication necessary to relieve a patient’s pain
Global Health improvement score (GHIS)	Sensory/affective	Measures whether a patient is a responder or nonresponder to intervention
Patient-reported outcomes measurement information system (PROMIS)	Sensory/affective/ psychosocial	Measures a patient’s pain symptoms and function, and the extent at which pain hinders a patient’s physical, mental, cognitive, emotional, recreational, and social abilities or activities, with sections including sleep disturbance and life enjoyment. PROMIS records these measures as T-scores
West haven-Yale multidimensional pain inventory (MPI)	Sensory/affective/ psychosocial	Measures the patient’s comprehensive chronic pain experience within three parts: (1) pain intensity, interference, spousal support, self-control, and negative mood, (2) significant other’s responses to patient’s pain, and (3) patient’s activities
Pain catastrophizing scale	Affective	Patients indicate how often they have catastrophizing feelings or thoughts on a scale from 0 (not at all) to 4 (all the time) while they are in pain, with a total score of 0–52 and three subscores measuring rumination, magnification, and helplessness
Beck depression inventory	Affective	Patients circle a number from 0 to 4 as they answer a questionnaire of 21 depression-related statements. The total score may be anywhere from 0 to 63, with scores 10 and under considered normal and scores over 18 considered clinically relevant depression
Montgomery-Asberg depression rating (MADRS)	Affective	Patient is clinically interviewed and rated on a scale of 0–6 for ten depression indicators: Apparent sadness, reported sadness, inner tension, reduced sleep, reduced appetite, concentration difficulties, lassitude, inability to feel, pessimistic thoughts, and suicidal thoughts
Health-related quality of life metrics	Affective/ psychosocial	Obtain health utility values from patients to determine the effectiveness and patient satisfaction of an intervention
EuroQol with five dimensions and three levels (EQ5D-3 L)	Affective/ psychosocial	Patients indicate their current functioning as one of three different levels: No problems, some problems, and unable to/extreme problems, and health state on a VAS from 0 to 100, within five dimensions: Mobility, self-care, usual activities, pain/discomfort, and anxiety/ depression
Functional indicators	Psychosocial	Measures the effect of a patient’s pain on their ability to function
Oswestry disability index (ODI)	Psychosocial	Patients answer questions tied to a numerical system that measures their permanent functional disability, particularly in relation to back pain

Table 2

Types of machine learning algorithms classified by learning paradigm

Supervised learning	Requires labeled data in accordance with the correct response, class, label, or outcome per case, from which the ML learns, creates an algorithm, and predicts future outcomes on new datasets. May be used for classification and regression of a continuous outcome
Linear regression	A supervised learning algorithm that provides a linear relationship between an independent variable and a dependent variable, used to predict future outcomes
Regularized logistic regression (RLR)	A technique that adds a penalty term to the error function the model is trying to minimize, reducing the complexity of the prediction function and preventing overfitting in models
Naïve Bayesian classifier	A supervised learning algorithm that seeks to model the distribution of inputs of a certain class or category, used for classification tasks
Support vector machine (SVM)	A set of supervised learning algorithms that construct a hyperplane or set of hyperplanes in the data space for maximum separation between data points. May be used for classification, regression, and outlier detection, and known for its predictive accuracy, effectiveness in high-dimensional spaces, and ability to handle nonlinearly separable data
Decision tree model	A predictive model that splits the dataset based on inferred rules from the identification of which features contribute the most to the decision-making process, via information gain. This model correlates observations about an item to conclusions about its target value
Classification and regression tree (CART)	A classification predictive model with decision tree output, where each fork splits in a predictor variable and each end node reveals a prediction for the outcome variable, providing a visual of the rules set for an outcome prediction on other variables
Artificial neural networks	A set of classification or regression algorithms, which mimics the human brain's structure of interacting nerves, with a network of interconnected nodes. An algorithm is applied to the input of each node which feeds the processed output of the next layer of nodes until reaching the output layer, which has the final prediction or classification
Unsupervised learning	Uses raw, unlabeled data with no associated outcomes to recognize patterns of natural occurrence which may divide the data into subgroups; also known as clustering
K means	A technique for data clustering which can classify unlabeled data into a predetermined number of clusters according to their similarities, often used for unsupervised machine learning
Linear discriminant analysis (LDA)	A supervised learning algorithm that reduces the number of features to a more manageable amount prior to classification, often used for solving classification problems greater than two classes
Ensemble methods	Uses amalgamation or a combination of diverse model with unique learning patterns, strengths, and weaknesses to improve predictive accuracy and robustness by aggregating their individual predictions
Random forest model	Collects data from a multitude of decision trees and outputs the class that is the mode of the classes output by individual trees using a tree learning algorithm, which selects a random subset of features at each fork split. This reduces variance and overfitting by averaging the trees
Gradient-boosted trees	A decision tree technique that trains a new model to predict the error, or difference between prediction and regressive labels, of the current strong model. this creates an optimal predictive value of the model
XGBoost	An implementation of gradient-boosting decision trees, which has faster training and may be parallelized across clusters