Optimization of Spinal Cord Stimulation Using Bayesian Preference Learning and Its Validation

Zixi Zhao, Aliya Ahmadi, Caleb Hoover, Logan Grado, Nicholas Peterson, Xinran Wang, David Freeman, Thomas Murray, Andrew Lamperski, *Member, IEEE*, David Darrow, and Theoden I. Netoff

***Abstract*— Epidural spinal cord stimulation has been reported to partially restore volitional movement and autonomic functions after motor and sensory-complete spinal cord injury (SCI). Modern spinal cord stimulation platforms offer significant flexibility in spatial and temporal parameters of stimulation delivered. Heterogeneity in SCI and injury-related symptoms necessitate stimulation personalization to maximally restore functions. However, the large multi-dimensional stimulation space makes exhaustive tests impossible. In this paper, we present a Bayesian optimization strategy for identifying personalized optimal stimulation patterns based on the participant's expressed preference for stimulation settings. We present companion validation protocols for investigating the credibility of learned preference models. The results obtained for five participants in the E-STAND spinal cord stimulation clinical trial are reported. Personalized preference models produced by the proposed learning and optimization algorithm show that there is more similarity in optimal frequency than in pulse width across participants. Across five participants, the average model prediction accuracy is 71.5% in internal cross-validation and 65.6% in prospective validation. Statistical tests of both validation studies show that the ability of the preference models to correctly predict unseen preference data is significantly greater than chance. The personalized preference models are also shown to be significantly correlated with motor task performance across participants. We show that several aspects in participants' quality of life has been improved over the course of the trial. Overall, the results indicate that the Bayesian preference optimization algorithm could assist clinicians in the systematic programming of individualized therapeutic stimulation settings and improve the therapeutic outcomes.**

***Index Terms*— Optimization methods, statistical learning, Bayes methods, neural engineering.**

Manuscript received March 23, 2021; revised July 16, 2021; accepted August 2, 2021. Date of publication September 20, 2021; date of current version October 1, 2021. This work was supported in part by Minnesota Office of Higher Education SCI/TBI Grant Program under Grant 159800, in part by the MnDrive Fellowship Program, and in part by St. Jude/Abbott for a generous device donation. *(Corresponding author: Theoden I. Netoff.)*

This work involved human subjects in its research. Approval of all ethical and experimental procedures and protocols was granted by the Hennepin Healthcare Institutional Review Board.

Zixi Zhao, Logan Grado, and Theoden I. Netoff are with the Department of Biomedical Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: tnetoff@umn.edu).

Aliya Ahmadi, Caleb Hoover, and David Freeman, are with the Department of Neurosurgery, University of Minnesota, Minneapolis, MN 55455 USA.

Nicholas Peterson is with the Department of Rehabilitation Medicine, University of Minnesota, Minneapolis, MN 55455 USA.

Xinran Wang is with the School of Statistics, University of Minnesota, Minneapolis, MN 55455 USA.

Thomas Murray is with the Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis, MN 55455 USA.

Andrew Lamperski is with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA.

David Darrow is with the Department of Neurosurgery, University of Minnesota, Minneapolis, MN 55455 USA, and also with Hennepin County Medical Center, Minneapolis, MN 55415 USA (e-mail: darro015@umn.edu).

This article has supplementary downloadable material available at https://doi.org/10.1109/TNSRE.2021.3113636, provided by the authors.

Digital Object Identifier 10.1109/TNSRE.2021.3113636

I. Introduction

ELECTRICAL neuromodulation was first introduced in 1967 for the treatment of chronic pain. Initially transcutaneous to stimulate large fibers according to the gate theory of pain, it was later modified to be an implantable epidural spinal cord stimulation (eSCS) unit [1]. Neuromodulation of the spinal cord (directly or indirectly) has since been successfully used to treat chronic pain and was serendipitously found to restore volitional movement in patients with multiple sclerosis as early as the 1970s [2]. In the past decade, there has been growing interest in the utilization of eSCS to restore volitional movement after chronic spinal cord injury (cSCI).

Currently, there is no treatment for cSCI, and until now, patients with motor-complete injuries have minimal hope of substantial recovery of volitional movement chronically (>1 year from injury) [3], [4]. Modern reports of successful partial restoration of function after chronic and stable SCI provide mounting evidence that neuromodulation may provide a significant and scalable treatment [5]–[7]. Despite these encouraging case reports and series, numerous challenges remain to be solved before a definitive clinical trial. Technological advancement of implantable neuromodulation platforms provides increasing flexibility for tailoring spatial and temporal neuromodulation to each patient, but large, high-dimensional flexibility requires efficient algorithmic optimization tailored to each patient with particular pathology and each underlying physiologic system.

Research on animal models has demonstrated a varied response in volitional movement based on stimulation

parameters [8]. This variation was further indicated in human models, thus leading to the need for spatiotemporal optimization of eSCS parameters that are customized for each patient [9], [10]. The observed variation in response to the stimulation parameters may result from the heterogeneity of SCIs, thus indicating the importance of discovering optimal parameters that are specific to the patient. There are trillions possible settings for a 16-lead paddle; thus, for electrical stimulation therapy to be used clinically, an efficient methodology for selecting the optimal settings needs to be developed. Current literature on eSCS optimization methodologies for eliciting volitional movement after cSCI is limited. The majority of the literature is focused on animal models, and the optimization strategies vary widely among groups. In 2013, Capogrosso *et al.* developed a computational model that combined a 3D finite element method with a model of the rat spinal cord to determine the spinal circuits and fibers recruited by eSCS. This model allowed for simulations of various electrode configurations to predict the optimal parameters for standing and walking in rats [11]. A Bayesian optimization-based active learning algorithm was developed to select bipolar stimuli to elicit specific muscle responses, measured by electromyography (EMG), in spinally transected rats [12]. Another animal model focused on the optimization of stepping by varying the stimulus intensity, frequency, and pulse-to-pulse timing. The optimal intensity and frequency were first determined subjectively by study personnel. These parameters were then evaluated again using kinematic data and EMG along with differing stimulation pulse intervals [13]. Similarly, one group applied quantitative gait variables from 3D kinematic data recordings to obtain optimal parameter combinations [8].

The literature on stimulation optimization in human models is even more limited than that in animals, and different optimization objectives and strategies have been tried in different studies. One study obtained a motor neuron activation map for each participant to determine which regions of the spinal cord were activated during specific muscle tasks. A computational model was then employed to perform simulations identifying optimal electrode configurations [10]. EMG mapping data have also been utilized to select optimal parameters [14]. The limit in the methodologies for performing optimization demonstrates a clear need to devise a novel, comprehensive approach to the selection of optimal eSCS parameters for volitional movement.

In recent years, machine learning-based optimization techniques have been utilized to optimize spinal cord stimulation in human clinical trials. Preference feedback was collected to represent the multifaceted effects of stimulation on a human participant. Utilizing human preference feedback data for treatment optimization holds potential for achieving personalized treatment while maintaining the therapeutic effect without the need to perform various expensive medical measurements. The comparison of options and expression of a preference for one over others also avoids problems associated with human ratings, such as variations in the scale between participants and at different times by the same participant [15]. A related algorithm, called the "correlational dueling bandit" algorithm, allows the effectiveness of untested stimulations to be inferred based on similar tested ones. This algorithm was used to facilitate the discovery of optimal spatial stimulation configurations for standing-frame training using physicians' preference feedback [16]. Another algorithm, called stagewise safe Bayesian optimization, has the benefit of enforcing safety in making decisions of stimulation recommendation, and was used in finding optimal stimulation for a tetraplegic patient in gripping experiments [17].

In the last decade, attempts have been made to create models that distinguish the optimal eSCS parameters, but the validation of these methods has not been extensively considered. Experimental studies on anesthetized rats utilized evoked EMG responses to validate their hybrid computational model [11]. While machine learning-based optimization algorithms have been applied in experiments and clinical studies to optimize stimulation parameters, rigorous validation studies have not been presented in the work to validate the accuracy of the models and the reliability of the optimal stimulation parameters. In studies that applied those techniques, the optimal configuration selected by the algorithm was validated only by comparing it with the configuration selected by human experts [12], [16], [17].

The Epidural Stimulation After Neurologic Damage (E-STAND) trial examines the effect of eSCS on volitional movement and autonomic function in participants with motor-complete thoracic SCI [5]. Acknowledging the heterogeneity in SCI, the trial emphasizes the discovery of optimal stimulation parameters based on the individual. This paper presents a Bayesian preference learning-based optimization method for eSCS, discusses the optimization results of five participants from the E-STAND trial, and highlights the validation studies and their results to illustrate the effectiveness of the proposed optimization method.

II. Method

A. E-STAND Study Criteria

This study was performed in accordance with Hennepin County Medical Center's Institutional Review Board approval and Food and Drug Administration Investigational Device Exemption approval. The participants identified in this study have provided informed consent and authorization to present publicly identifiable information for research purposes. The study protocol is registered with ClinicalTrials.gov (NCT03026816). Patients with chronic (more than one year since injury), traumatic SCI were recruited if they met the following criteria: greater than 22 years of age, motor-complete AIS classification A or B with a neurological level of injury between C6 and T10, full arm and hand strength, and intact segmental reflexes below the level of injury. Participants were excluded if they had medical or psychological comorbidities that would significantly increase the risk associated with the operation or if they had severe dysautonomia with systolic blood pressure fluctuations below 50 or above 200 mmHg on autonomic testing, contractures, pressure ulcers, recurrent urinary tract infection refractory to antibiotics, unhealed spinal fracture, recent botulinum toxin use, or pregnancy.

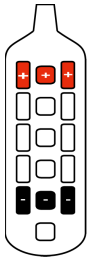

Fig. 1. Schematic of the Abbott stimulator's electrode paddle.

The demographics of the five participants can be found in the supplementary materials. These five participants are the early participants who have completed the 80% of trial visits at the time of writing this manuscript.

B. Setting Optimization

1) Spinal Cord Stimulation Optimization Protocol: Each participant in the E-STAND study is implanted with the stimulator a month prior to the stimulator being turned on and programmed with stimulation settings for optimization. The implanted stimulator is an Abbott stimulation unit composed of a Proclaim 7 Elite battery and a Lamitrode Tripole paddle with a 16-electrode array, as shown in Fig. 1.

The E-STAND study primarily focuses on the optimization of temporal parameters, i.e., frequency and pulse width. The optimal electrode spatial configurations are heuristically determined before the third visit through clinician tuning and EMG measurements. During the first two visits, several spatial electrode configurations are tested by the physician and evaluated based on the participant's acute movement response. A "mapping" session is also performed to collect EMG measurements of muscle twitch in response to a large set of spatial stimulation configurations at 2 Hz. The spatial configurations that induce the most movement or prominent muscle activation are individually combined with a series of temporal parameters for further at-home evaluation between visits. At the third visit, the spatial configuration that is most preferred by the participant is considered to be the optimal spatial configuration. After the third visit, the optimization method primarily focuses on temporal parameters revolving around the optimal spatial configuration. To be noted, the stimulation amplitude is controlled by the participant through a patient programmer each time when a stimulation setting is used, hence it is not optimized in conjunction with frequency and pulse width.

Each participant has 13 follow-up visits arranged 30-45 days apart. During the first 12 visits, the participant's stimulator is programmed with a set of 15 settings to be tested at home. Participants are blinded to the stimulation parameters associated with the settings, and any old settings that are reassigned to the participant are randomly given a new program number. The set includes both settings focused on restoring volitional movement and settings focused on restoring autonomic function. When at home, the participant is instructed to use one volitional setting each day according to a calendar that establishes the sequence of settings to be tested. The participant uses volitional programs for a triple flexion and extension leg exercise task (further described in the supplemental materials) and other movement-related activities. At the end of the day, the participant evaluates the current setting in an online survey by comparing it with the last one, and they may optionally compare it with any settings that the participant has used since the most recent visit. Here we enforce binary comparison evaluations instead of evaluation scores as a scoring metric tends to drift over time [15]. The optional comparisons data between current day's setting with any previous settings used in the current month are treated equally as the collected mandatory comparison data.

TABLE I
OPTIMIZATION PHASES FOR EACH ESTAND PARTICIPANT

Phase	Optimization objective
Follow ups 1 -3 (3~ 4.5 months)	Find the optimal spatial configuration. Promising spatial configurations are selected based on mapping results and in-clinic movement response. Those spatial configurations are paired with same pairs of temporal parameters for settings for at-home testing and evaluation.
Follow ups 4-11 (8~ 12 months)	Learn the preference map in the temporal parameter space. The spatial configuration of settings tested in this phased are the same, i.e., the optimal spatial found in phase 1.
Follow ups 12,13 (2 ~ 3.5 months)	Other reasonable spatial configurations are further tested with the optimal temporal parameters as a refinement step.

At the end of each month, a preference model is built based on the participant's daily preference evaluations. This model is able to give preference predictions on untested parameters, and new settings with high predicted preference scores will be included in the next set. In a case of 9 volitional settings, 6 are usually new, untested settings, and the other 3 are previously tested settings that gave desirable results to ensure an ethical distribution of settings. When the optimized temporal parameters converge and the preference model presents a decent prediction accuracy, other reasonable spatial configurations are further tested with the optimal temporal parameters as a refinement step in the last two or three months. The different phases of optimization and objectives are summarized in Table I.

2) Stimulation Optimization Algorithm: The algorithm for spinal cord stimulation optimization presented in this paper is under a Bayesian optimization framework, and the optimization goal is to maximize a patient's preference level by finding the optimal temporal stimulation parameters, i.e., frequency and pulse width. The optimization algorithm consists of three parts. The first part fits a preference model, in the form of a Gaussian process, to cumulatively collected comparison data; the second part, called sampling, determines a new set of settings based on the model; and the third part, called

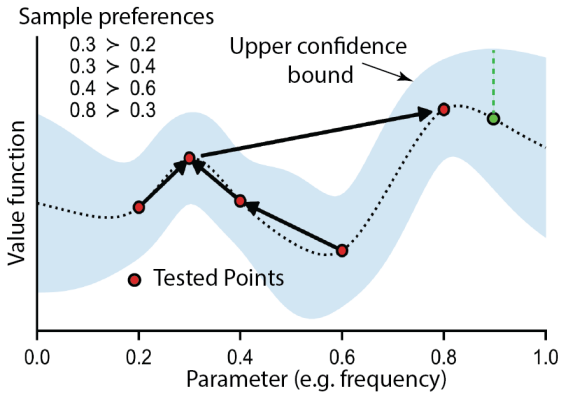

Fig. 2. Example of a one-dimensional Bayesian optimization plot with relational comparison data. The black dashed line is the fitted preference value function, and the blue region is the confidence bound.

sequencing, determines the sequence of settings to be used in the following month. The sequence essentially determines which two settings are going to be directly compared against each other.

a) Bayesian preference modeling: Preference modeling is performed to find a best-fit relationship between a latent variable representing a participant's unobserved preference level and options through Bayes' rule. The best-fit preference model is essentially the posterior distribution computed through Bayes' rule and the predictive distribution of preference values for unobserved options. Again, in our application, the data set used to extract the model is composed of binary comparisons.

We utilize probit modeling with Gaussian process-modeled latent variable to describe preferential, binary observations, i.e., the likelihood in Bayes' rule. Probit modeling enables us to derive the probability expressions of binary outcomes. The Gaussian process-modeled latent variable also dictates the prior distribution expression in Bayes' rule. The Gaussian process property is inherited in the posterior distribution with some approximation, which subsequently enables us to derive the predictive distribution. The details of this part of the algorithm can be found [15].

Fig. 2 shows an example of a preference model built based on four binary comparisons. The red dots connected by arrows represent comparisons that have been evaluated, and the arrows point from the losing setting to the winning setting. The dashed line represents the mean of the fitted Gaussian process preference variable, and the blue shaded area represents the confidence bound.

b) Batch sampling: After the preference model is developed, the sampling algorithm chooses new queries in the parameters to perform the next test. The new queries should strike a balance between the exploration of untested parameter regions and the exploitation of tested parameter regions that are known to give desirable results. The balance is dictated by the so-called acquisition function. There are different forms of acquisition functions with different properties, and we use the Upper Confidence Bound (UCB) for our application. The UCB acquisition function is a weighted sum of the mean and standard deviation predicted by the fitted preference model.

Pseudo-code of the batch sampling algorithm

Inputs: batch size: q; UCBPower acquisition function parameters: α, β; preferential data set $\mathbf{D}_t$; tested stimulation parameters: $\mathbf{X}_t = \{\boldsymbol{x}_{1:n}\}$; candidates of test parameters to be included in the batch: $\mathbf{X}_{\text{test}}$

Initialization: $\mathbf{D} = \mathbf{D}_t; \mathbf{X} = \mathbf{X}_t$; GP $= p(\mathbf{f}_t|\mathbf{D}_t, \mathbf{X}_t); \mathbf{f}_t = [f(\mathbf{x}_1), f(\mathbf{x}_2), \ldots, f(\mathbf{x}_n)]^{\mathrm{T}}$

Sampling:

For $k = 1{:}q$,

Based on GP, find $x^+ = \arg\max_{x \in \mathbf{X}} \text{UCBPower}(\mathbf{x}, \alpha, \beta)$

For $\hat{x} \in \mathbf{X}_{\text{test}}$,

Predict $\hat{d} = \text{Pref}\left(\mathbf{E}(f(\mathbf{x}^+)), \mathbf{E}(f(\hat{\mathbf{x}}))\right)$

Fit $\widehat{\text{GP}} = p(\hat{\mathbf{f}}|\mathbf{D} \cup \{\hat{d}\}, \mathbf{X} \cup \{\hat{\mathbf{x}}\})$

Calculate S-UCBPower$(\hat{\mathbf{x}}) = \sum_{\mathbf{x} \in \mathbf{X}_t \cup \{\hat{\mathbf{x}}\}} \text{UCBPower}(\mathbf{x}, \alpha, \beta)$

Obtain the kth point, $\mathbf{x}_k$, for the batch:

$\mathbf{x}_k = \arg\max_{\hat{\mathbf{x}} \in \mathbf{X}_{\text{test}}} \text{S-UCBPower}(\mathbf{x}, \alpha, \beta)$

Augment $\mathbf{D} = \mathbf{D} \cup \left\{\text{Pref}\left(\mathbf{E}(f(\mathbf{x}^+)), \mathbf{E}(f(\mathbf{x}_k))\right)\right\}$

Augment $\mathbf{X} = \mathbf{X} \cup \{\mathbf{x}_k\}$

Update GP $= p(\mathbf{f}|\mathbf{D}, \mathbf{X}), \mathbf{f} = [f(\mathbf{x}_1), f(\mathbf{x}_2), \ldots, f(\mathbf{x}_n)]^{\mathrm{T}}$

Update $\mathbf{X}_{\text{test}} = \mathbf{X}_{\text{test}} \backslash \{\boldsymbol{x}_k\}$

Fig. 3. Pseudo-code of the batch sampling algorithm.

For the example in Fig. 3, the weights in the UCB acquisition function are both one, and the green dot is the next sample point, as the UCB is the highest at that location.

The Abbott stimulator used in the study holds up to 15 different settings, and we are able to program the device with new settings only during once a month clinical visit. Therefore, the sampling algorithm needs to determine a batch of new parameters to sample instead of one. We designed a batch sampling algorithm that sequentially selects top q parameter locations (q is the batch size) from the candidate pool. To select each sample, the algorithm first simulates the outcomes of comparing of each setting in the pool against the incumbent, which is the tested setting that has the highest UCB; the algorithm then refits a model considering both the observed data and the simulated data. The parameter that produces a model with the highest sum UCB (S-UCB) is included in the batch and deleted from the candidate pool. Fig. 3 shows the pseudo-code of the batch sampling algorithm.

c) Sequencing: The last part of the algorithm arranges the settings to be tested in a sequence to maximize the information obtained from the pairwise comparisons; the resulting sequence becomes the "calendar" that guides program usage. The sequence algorithm is called the "hub-and-spoke" algorithm, producing sequence that only compares the top m settings (hub settings with the highest predicted preference values) to the rest and to each other. The produced sequence maximizes the information that confirms if the high preference value still holds for the hubs and quickly spots new settings that win over the hubs. The hub-and-spoke algorithm also eliminates repeated comparison between hub and non-hub settings but allows for repeated comparison between hubs. Another example of the "hub-and-spoke" algorithm sequence can be found in the supplemental materials.

C. Validation Protocol

We designed the validation studies' protocols with two goals. One goal is to examine the prediction accuracy of the learned preference models, which is achieved through internal

validation and prospective validation studies. In both studies, models trained with a sub-dataset are used to predict the outcomes of the rest, and the model prediction accuracy is then taken as the percentage of accurate predictions. The second goal is to examine the relevance of the preference model's prediction to volitional movement in response to stimulation.

Internal validation is conducted as 5-fold cross-validations across 5 participants. For each participant, the cross-validation randomly partitions the participant's preference dataset from the temporal optimization phase into 5 equal folds; each fold is then sequentially held as the testing set, and the preference model is trained with the remaining 4 folds. The internal validation accuracy is then the average model prediction accuracy averaged across all 5 folds. The results of this internal validation study from the 5 participants are discussed in section III.B.1.

Prospective validation is also performed to show the prediction accuracy of the preference model. This validation study protocol requires additional preference data to be collected at the end of each participant's temporal parameter preference learning phase. In the prospective validation study, the participant tests 9 or 10 different stimulation settings in a row and then tests the same settings again in a different order. The participants are asked to compare two sequential settings based on a 30 s triple flexion and extension task. The tested stimulation settings have been previously used by the participant, and the preference scores associated with those selected settings span the score range produced by the model. The prospective validation accuracy is determined by how well the model predicted outcomes match up with the observed outcomes. The results of the prospective validation study are described in section III.B.2.

Clinical outcome validation is performed to examine the correspondence between an objective measure of the participants' volitional movement response and the preference model prediction for the same stimulation parameters. The volitional movement measurement is extracted from the daily triple flexion and extension task performed by the participants. The analysis and results are further described in section III.C.

III. Results

A. Preference Model

The results of stimulation parameter optimization for five participants are presented here. Preference models developed at the end of the temporal parameter optimization phase are shown as heatmaps in Fig. 4. In the heat maps, regions that have higher preference values are indicated by a brighter hue. The tested stimuli's temporal parameters are indicated by plus-sign markers on the preference surfaces to illustrate the progress of stimuli testing. The marker's color indicates how recent the stimulus was tested; the color scale spans the rainbow, with blue indicating settings that were last tested earlier in the study and red representing settings that were last tested recently in the study. The size of the marker indicates how many times that stimulus was tested. Generally, stimuli in preferred regions are tested more than those in less preferred regions since a participant's favorite settings can be used many times within each month and for several months.

According to the preference maps, multiple preferred parameter regions can exist for each participant. At the end of the study, the participants' devices are programmed with a set of settings that covers parameters from those preferred areas.

B. Preference Model Validation

1) Internal Cross-Validation: Ten repeats of five-fold internal cross-validation was conducted on each participant's dataset to evaluate the prediction accuracy of the Bayesian preference model. For each time, the comparison data set collected for each participant are randomly divided into five even folds. Iteratively holding one-fold of data as the testing data set, and other 4 folds as the training data set to assess the model prediction accuracy. The average prediction accuracy across all folds and all repeats and its standard deviation are used here to present the average model performance for each participant. The individual cross-validation accuracies with the corresponding standard deviations are displayed as scatter plots in Fig. 5. We conducted a one-sample t-test with a null hypothesis that the mean cross-validation accuracy is 50%. The null hypothesis represents the situation in which the participant's response and stimulation setting preference provide no useful information and so there is a 50% chance of correctly guessing the outcome. The test shows that the mean cross-validation accuracy is significantly higher than 50% (p-value=0.002). The observed cross-validation accuracy, averaged across 5 participants, is 71.4% with a standard deviation of 7.1%.

We find that the cross-validation accuracy of the preference model is positively correlated with the participant's compliance and with the signal to noise ratio (SNR) estimated from each data set. The compliance is represented by the number of comparison pairs that have been collected for each participant and subsequently used to develop each participant's preference model shown in Fig. 5. Participants who assess stimulation settings persistently and provide answers to optional comparison question, i.e., comparing current-day setting with settings previously used from the same month, would produce a larger data set and hence higher compliance. Fig. 5 (a) plots the model's cross-validation accuracy against the participants' compliance and displays the regression line, which has a correlation coefficient of 0.94. The slope of the regression line is 0.025% (p-value=0.02), indicating that every additional 100 comparisons is predicted to improve the cross-validation accuracy by 2.5%. The positive correlation relationship is also found between the model's cross-validation accuracy and the estimated signal to noise ratio as plotted in Fig. 5 (b). The noise is incorporated in the probit preference model to handle the inconsistencies in the observations and can be estimated through a hyperparameter optimization routine when fitting a preference Gaussian process model [21]. Since the fitted preference value spectrum varies individually, the SNR is reported here instead of the noise alone as a quantification of the inconsistencies found in each participant's data. The SNR for participant i is defined as:

$$SNR_i = \frac{\text{Fitted preference score range for participant } i}{\text{Noise estimated from participant } i\text{'s data set}}$$

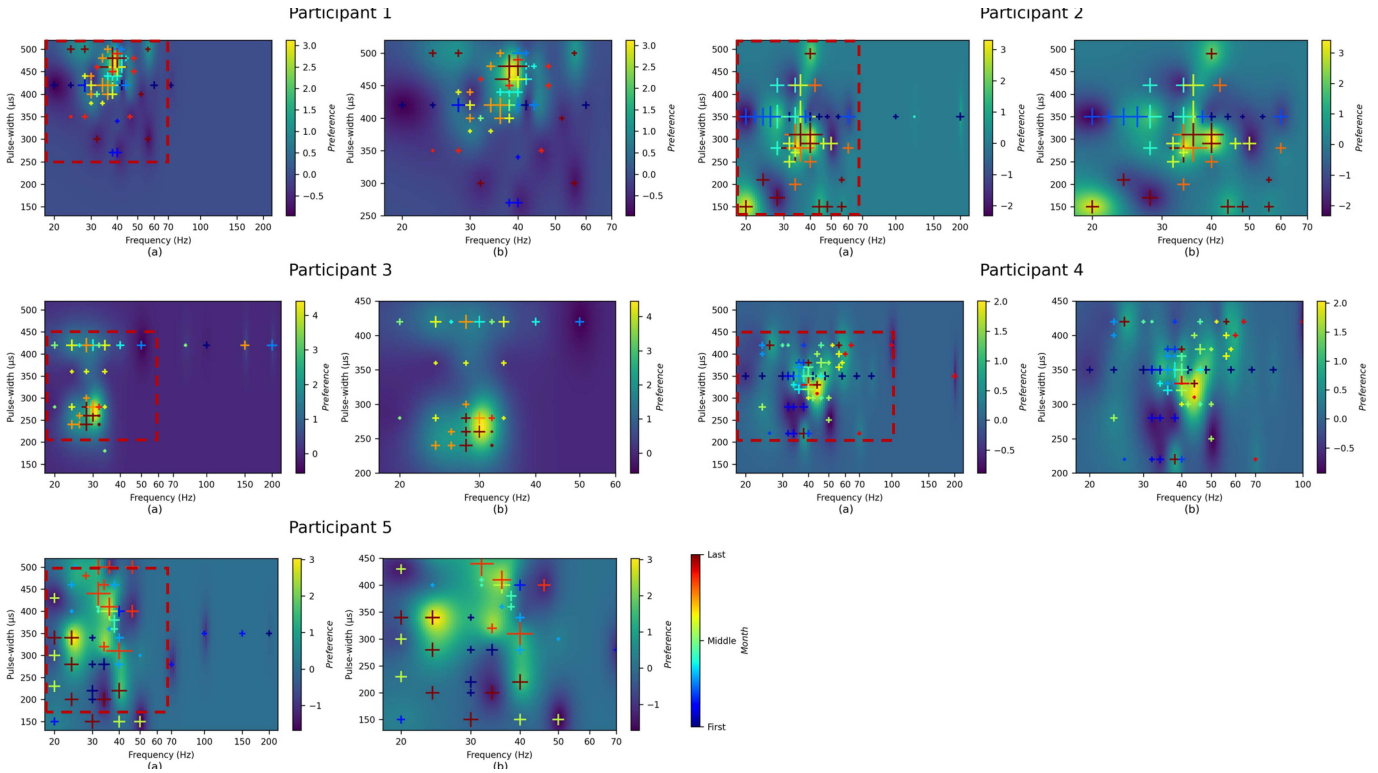

Fig. 4. Individual response surfaces representing the overall volitional preferences from 5 participants. For each participant, subplot (a) shows the preference map on a full parameter range enclosing all the setting tested, and subplot (b) shows a close-up map for the parameter range where includes most tested settings for each participant. The specific range plotted in (b) are highlighted by the dash squares in (a).

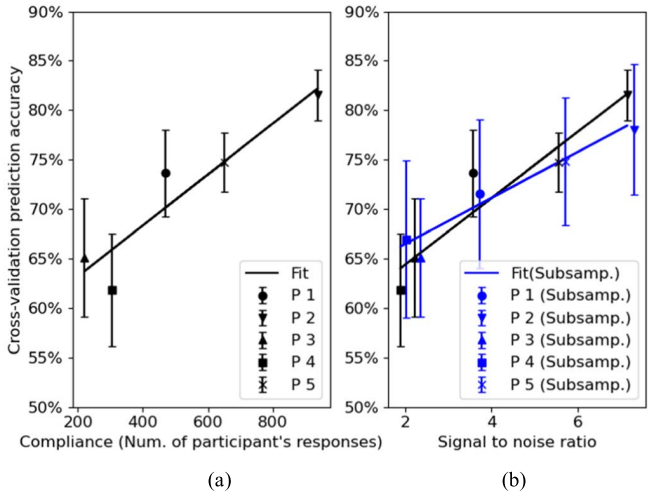

Fig. 5. Cross-validation accuracy. (a) Cross-validation accuracy vs. compliance. (b) Cross-validation accuracy vs. SNR. In both subplots, scatter plot with error bars to show the means and standard deviations for the 5-fold cross-validation accuracy for each participant.

The correlation coefficient is 0.95. The slope of the regression line is 3.37% (p-value = 0.01), indicating that an improvement SNR value of 1 is predicted to improve the cross-validation accuracy by 3.37%.

To further investigate the impact of SNR and compliance to the cross-validation accuracy, we conducted 5-fold cross-validation on 10 randomly subsampled datasets for P1, 2, 4, 5, respectively. Each randomly subsampled dataset has the same size as P3's original data set (220 pairs of comparisons), as P3 provided the least amount of comparison data. It shows that reducing number of data widens the range of cross-validation accuracy within each participant. The average cross-validation accuracies from the subsampled datasets still follow the positive linear correlation trend as shown using the blue regression line (slope = 2.33, p-value = 0.01, correlation coefficient = 0.97) in Fig. 5 (b).

2) Prospective Validation: Prospective validation studies were performed on the 5 participants as the protocol described in II.C. The prediction accuracy of each participant's preference model in matching the observed outcomes is reported in Table II. Overall, the average prediction accuracy is 65.63%, and the result of a one-sample t-test shows the average prediction accuracy is higher than 50% with a p-value of 0.0275, indicating that the preference model significantly helps to predict the correct preferred setting. Though participant 2's model prediction accuracy is lower than 50%, the accuracy increases to 72.73% if comparisons involving only the top 4 settings (according to the preference model prediction) are considered. Similar accuracy increases also appear for the other 3 participants when considering the subset involving only the top 4 settings. This is because the top settings are usually tested more times than the inferior settings when developing preference models and finding optimal stimulations. Consequently, the average prediction accuracy for this sample subset increases to 72.15%, and the result of the one-sample t-test shows that the mean accuracy from the prospective validation study is higher than 50% with a p-value of 0.001.

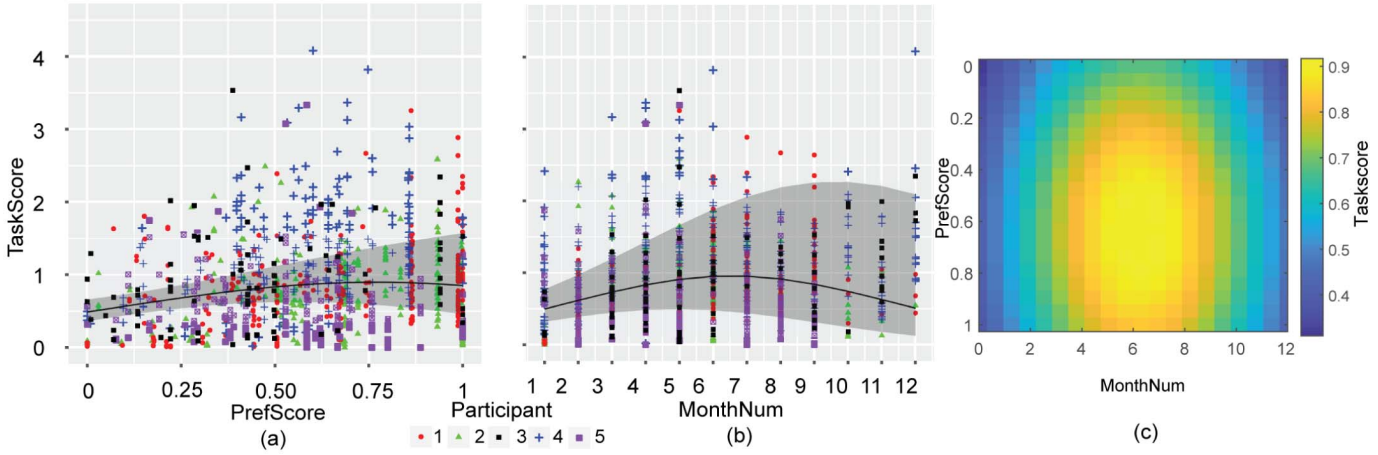

Fig. 6. Models of task score vs. preference score and/or month number and the corresponding data. (a) GLMM population-level predictions of task scores based on preference scores are shown by the black line, and the shaded area represents the 95% confidence interval of the population-level predictions. The GLMM is fitted to all five participants' data. The task score for each triple flexion and extension task and the corresponding preference score based on the stimulation parameter settings are shown in the scatter plot. (b) GLMM population-level predictions of task scores based on month number (c) The response surface shows the task score from the fixed-effects prediction in the fitted GLMM. The GLMM is fitted to all 5 participants' data.

We performed another statistical test showing that the success rate of the preference model predictions is higher than 50% based on the total number of comparisons and accurate predictions from all 5 participants. To construct the test, we considered the prediction of the outcome of a pairwise comparison a binomial process. For all 5 participants, the outcomes of 62 pairwise comparisons out of 96 match the model prediction; the probability of this observation happening is 0.0014 if the binomial process success rate is 50%. Therefore, the prospective validation study statistically shows that the preference model learned through Bayesian preference optimization is representative of the participant's true preference.

C. Clinical Validation of Optimization

1) Movement Metrics and Task Score: We constructed power-based metrics based on Inertial Movement Unit (IMU) signals to measure leg movement during each performed task. In each task, an average active trial is obtained by averaging all repeated active trials in the task; the peak and baseline values of the average active are used to compute the task performance metric values, i.e. task scores. The further description on the task scores construction can be found in the supplementary materials.

2) Validation of Preference Learning With Task Score: Motor response is an important metric for evaluating the effectiveness of the optimization-guided eSCS therapy. Motor responses are characterized by task scores computed based on IMU signals collected during daily triple flexion and extension tasks. In this section, we present the results from a generalized linear (mixed) model (GL(M)M) analysis correlating the movement task score to the preference score and/or the therapy time. The GLMM analysis was performed in R (version 4.0.0). In each following analysis, the Akaike information criterion (AIC) is used to select the optimal predictor formulation; the optimal probability distribution and link function are found through residual diagnosis performed by the DARHMa package [18]. For all the GLMMs presented in the following, the optimal probability distribution and link function are found to be the Gaussian distribution and the log link function.

TABLE II
SUMMARY OF PROSPECTIVE VALIDATION ACCURACY FROM 5 PARTICIPANTS

	Pred. accuracy for all comparisons		Prediction accuracy for comparisons involving only the top 4 settings	
P 1	15/19 correct	78.95%	10/11 correct	90.91%
P 2	9/19 correct	47.37%	8/11 correct	72.73%
P 3	12/17 correct	76.47%	7/13 correct	53.85%
P 4	10/19 correct	52.63%	8/14 correct	57.14%
P 5	16/22 correct	72.72%	14/16 correct	86.50%
Avg. accuracy (p-value of mean accuracy > 50%)	65.63% (p-value=0.0275)		72.15% (p-value=0.0013)	
Prediction results for all 5 participants (p-value of the success rate >50%)	62/96 correct (p-value= 0.0014)		47/65 (p-value = 7.61×10^{-5})	

Fig. 6 shows the GLMM population-level regressions of the task score with respect to the preference score in (a), the month number in (b), and both as in (c). The preference scores are normalized into the same [0, 1] range for each participant.

The population-level regression formula is as follows:

$$\log TaskScore = -0.71^{***} + 1.57 PrefScore^{**} - 1.02 PrefScore^{2(*)},$$

where both the preference score (PrefScore) and the squared preference score (PrefScore^2) are found to be significant covariates with the outcomes, indicated by the asterisks. As shown by Fig. 6 (a), the population-level regressed task score increases linearly up to a normalized preference score of 0.7, after which it plateaus, and the confidence interval widens. Similarly, we performed GL(M)M analysis of the task score with respect to the number of treatment months. The population-level regression formula is: $log(\text{TaskScore}) = -0.95(***) + 0.28\text{MonthNum}(***) - 0.02\text{MonthNum}^2(***)$, where both the linear and quadratic terms of the month number are found to be significant covariates. As shown in Fig. 6 (b), the task score improves for the first 7 months of the study and then appears to deteriorate up until the end of the study. GLM fitted to each participant's data are presented in the supplemental materials, presenting differences among individual's trends. To determine if task performance is dependent upon both the preference score and the length of treatment, a final GLMM is regressed using linear and quadratic terms of both the preference score and the treatment month number. The interaction term between preference score and the treatment month number was found to be insignificant and therefore was not included in the final GLMM based on the AIC. The population-level regression formula in the best-fit GLMM is as follows:

$$\log TaskScore = -1.17^{***} + 0.25 MonthNum^{***} + 0.94 PrefScore^{*} - 0.02 MonthNum^{2(***)} - 0.73 PrefScore^{2(*)}$$

Fig. 6 (c) shows the response surface of the task score given by the month number and preference score based on fixed-effects predictions. The surface shows that an average participant's task performance improves monotonically in the first 6 months with respect to the normalized preference score up to 0.6-0.7. Then, the task scores plateau and slightly decrease as the month number increases while staying around the peak with respect to the preference score covariate. The fitness of the regression in an individual participant's case is also shown in the supplementary materials.

3) Quality of Life Improvement: Patients with cSCI often experience adverse effects in multiple body systems, suffer from many secondary complications, and must cope with altered social roles and psychiatric comorbidities [19]. Though regaining volitional control is an obvious priority, other improvements such as reducing injury-related pain, elevating energy levels, and regulating blood pressure are also greatly appreciated by the patients and may be reflected in the improvement of the patient's quality of life. ESTAND participants fill out the Patient Health Questionnaire-9 (PHQ-9) and the WHO Quality of Life (QOL)-Bref questionnaire at each visit. Further explanation on the questionnaires can be found in the supplemental materials.

GLMM analysis was performed to detect significant trends in the quality of life scores over time in the 5 participants, and the results are shown in Fig. 7. The baseline scores recorded prior to each participant's device implantation surgery are labeled as the "−1" visit. Subfigure (a) shows that for the PHQ-9 score, the visit number was not a significant covariant, indicating that statistically, the stimulation usage did not have a significant negative impact on the patients' mental state. Through GLMM analysis of the WHOQOL-Bref scores, shown in subfigures (b-f), we found that the physical health, psychological health, and environmental domain scores are significantly positively correlated with the visit number at a population level. Based on the fitted values of the slopes, an average patient's quality of life scores may improve by 4.16 in the physical health domain, 1.69 in the psychological domain, and 2.34 in the environmental domain after a period of one to one and half years of stimulation therapy usage and optimization. The WHOQOL-Bref summary score and the social relation domain score were not found to be significantly correlated with the visit number at a population level, as shown in subfigures (b) and (e), respectively.

IV. Discussion

Epidural spinal cord stimulation has attracted growing interest as a promising treatment for chronic spinal cord injury patients. Due to the heterogeneity in injuries, the optimal therapy needs to be personalized. In this paper, we developed a Bayesian optimization-based preference learning algorithm to optimize the frequency and pulse width of stimulation based on an individual participant's at-home evaluation, which was conducted through online preference surveys. The algorithm constructs participant-specific preference models that are then used to suggest new therapeutic settings that efficiently explore the broad parameter space and also focus in optimal parameter regions. To validate the efficacy of the algorithm, we examined resulting preference models quality through internal and prospective validation studies. The preference models were also validated against motor outcomes, which is the primary clinical outcome, to confirm the correlation between the preference models' predictions and leg movement responses to stimulation. The clinical outcome validation study is important for proving that human perception data and learning are instructive to the development of personalized eSCS therapy for patients with cSCI.

The first objective of this study was to show that the Bayesian optimization-based preference learning algorithm produces credible participant-specific models. We validated the model internally, where the model was used to predict preferences in a 5-fold cross-validation study, and prospectively, where model predicted for new, unseen comparisons obtained at the end of the study. The average accuracy across 5 participants' preference model was 71.5 in internal validation, and 65.6% in prospective validation; both of these values are significantly higher than chance levels, with p-values of 0.001 and 0.03, respectively. We also found that the preference models better predict the outcomes of comparisons involving top-rated stimulation settings, which have been tested often, than those of comparisons between settings, which were not

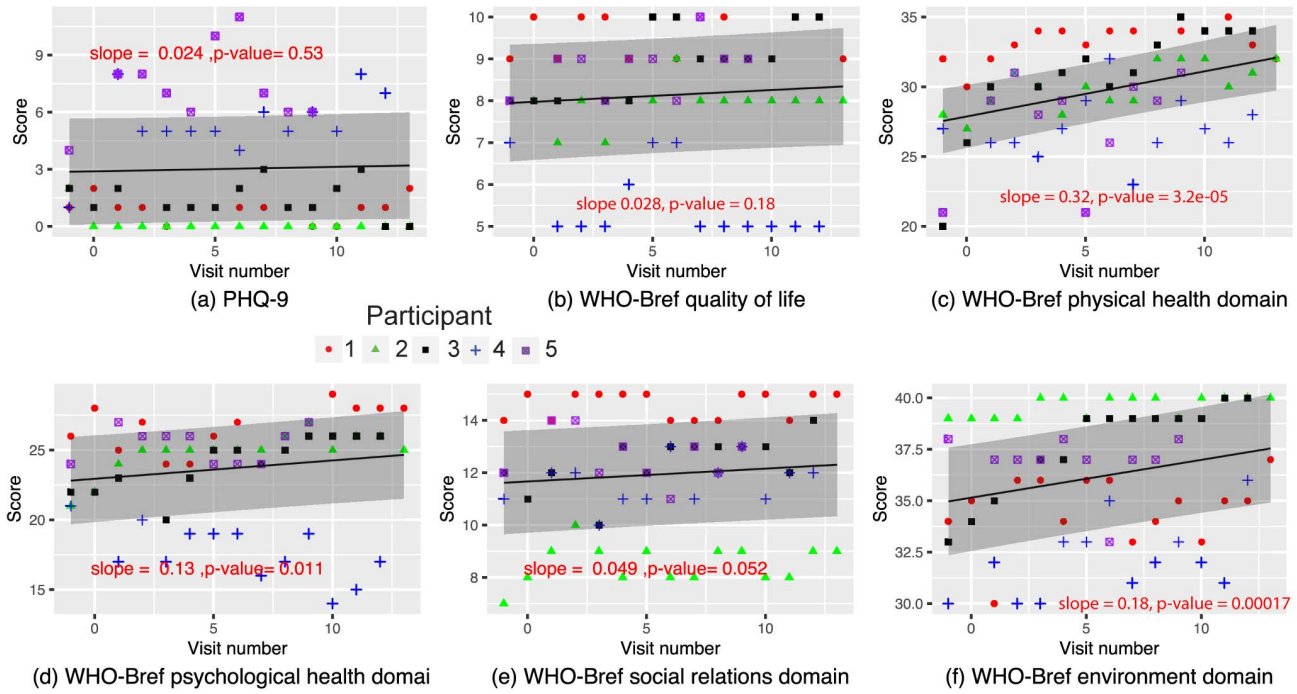

Fig. 7. Mixed models of quality of life scores vs. visit number. Participants' raw scores are presented as scatter plots. GLMM population-level predictions are shown as black lines, and the shaded areas represent the 95% confidence intervals. The GLMMs are fitted to all five participants' data in each case. The slopes and their p-values are provided for each GLMM. There are significant improvements in the physical health, psychological health, and environment domains from the WHOQOL-Bref analysis. The social relations domain is trending towards significance. The PHQ-9 score can range from 0 to 27, where a lower score indicates a healthier mental state. Subfigures (b)-(f) present the WHOQOL-Bref scores for which higher score values indicate better health states, and individually, each score in subfigures (b)-(f) can range from 2-10, 7-35, 6-30, 3-15, and 8-40, respectively.

favored and thus were not sampled as often and not well characterized. These results support the central hypothesis that Bayesian preference learning produces participant-specific models that can predict participant's preferences for settings better than chance.

The second objective of this study was to show that the preference models are correlated with the motor responses. In the study, we utilized daily leg movement tasks to track the motor response to a series of settings. Through GLMM analysis, we found that for an average participant, preference scores, i.e., the preference models' judgement of the effectiveness of a setting, are significantly correlated with motor response, as shown in Fig. 6 (a). This result confirms that the patient preference model reflects motor outcome. However, the motor task score plateaus at the high end of the preference score range. This result suggests that the preference score does not perfectly reflect the motor task score. A patient's preference depends on many aspects imparted by the therapy that are not well characterized by the motor task score, such as precision in volitional control, comfort, and the suppression of spasticity.

The third objective of the study was to determine if the length of treatment has a significant effect on the motor score. To test this relation, we looked at the time effects of the GLMM analysis and found a significant positive correlation between the motor score and time, as shown in Fig. 6 (b). The motor response improved for the first two-thirds of the study and then dropped slightly at the end. One possible explanation for the decrease in motor task score at the end of the study is that the scores were constructed based on the IMU power; this metric favors large magnitude gross movements and is not as sensitive to lower magnitude and finer controlled movements. As participants gain better control of their limb movements over the course of the study, their objective switches from gross movement to finer controlled movement and precision. This change may result from improvements in motor control with use of the device as a result of learning and plasticity and from the optimization of the stimulation parameters based on preference. This study was not designed to distinguish between these effects. Finally, it is possible that the stimulation efficacy decreases over the one or one-and-half years study period after implantation. Further studies will be needed to identify the precise cause of this decrease in motor score over time. However, from this model of motor score vs. time, it is not possible to determine if the changes in motor score arise from improvements imparted by the use of the device or from Bayesian optimization.

The fourth objective of the study was to determine if Bayesian optimization improved patient outcomes. To test this premise, we fit a GLMM that regressed motor scores to both treatment time and preference score. The resulting model, whose response surface is shown in Fig. 6 (c), shows that the motor scores are dependent on both the preference score and treatment time, allowing us to conclude that preference optimization influences the motor score that is independent of time. Furthermore, the population-level dependency of the motor responses on treatment time and preference score

was found to be nonlinear and non-monotonic, as seen in the models that fit the preference score and treatment time independently. To further evaluate the influence of Bayesian optimization on patient outcomes, we also tested the premise that stimulation therapy has a positive impact on a patient's quality of life in general. Fig. 7 shows that three out of six quality of life scores improved significantly at a population level over the course of the trial; these scores were the WHOQOL-Bref physical health, psychological health, and environmental domains. The WHOQOL-Bref summary score and social relations domain score are positively correlated with the visit number, but the correlations are not statistically significant. The PHQ-9 score shows that there is an insignificant negative correlation between the mental health state and the visit number. Three participants' PHQ-9 scores constantly remained below 5, suggesting that these patients were not suffering from depression, and participant 4 and 5's PHQ-9 scores increased after the start of the clinical trial but stayed below 10, which is categorized as mild depression. Though there was no evidence that the use of stimulation was the major cause of their increased depression scores, neither of the participants experienced complications caused by the stimulation, as indicated in their daily evaluation of stimulation notes. While the current optimization did not include quantifiable biometrics, they could be added into the objective function along with the preference in the optimization.

The fifth objective was to determine if there is any commonality in the preference models of temporal stimulation parameters and optimal parameter regions across participants. We found that the preferred stimulation frequencies all fall in the range 28 Hz to 44 Hz, whereas the preferred pulse widths exhibit more individual variation, as shown in Fig. 4. The response surfaces for the motor function preferences are shown in Supplemental Fig. 2, and the corresponding optimal parameters are similar to the overall preference maps and their optimal parameters. This similarity indicates that the motor responses are the main factor considered in a participant's assessment of overall preference. However, the optimal parameters for motor function are not necessarily the best for other outcomes. Supplemental Fig. 3 shows preference maps based on Participant 1's preference answers on other outcomes, i.e., transfer, posture, comfort, spasticity, and rehabilitation. The optimal parameters for posture and spasticity shifted to [36 Hz, 400 μs] from [40 Hz, 450 μs], which are the optimal parameters for overall preference. These results suggest that the optimal stimulation settings may have some common ranges across patients but that significant differences still exist across patients and outcomes, and these differences merit patient-specific optimization.

A patient's preference may not be static and may slowly migrate over time. Preferences might not be static for several reasons, one being that the patient's objective may change over time. For example, their initial preference may focus on the amplitude of movement and then later shift to the control of movement or other more subtle benefits that may align better with improvements in quality of life, i.e., the suppression of spasticity and comfort. We have also observed that function has been restored in patients even after stimulation is turned off [20], indicating that there may be plasticity effects that lead to changes in the response to stimulation and possibly the temporal parameter preference. If preference drifts over time, it might not be appropriate to weight more recent data heavier than earlier samples, but this study treated them equally.

V. Conclusion

Here we presented an eSCS optimization framework based on preference learning and Bayesian optimization for the treatment of patients with cSCI. The novelty of this stimulation optimization approach lies in the utilization of patients' self-reported preference data to build personalized preference models and Bayesian optimization to suggest new stimulation parameters to be tested. Using patients' self-reported at-home evaluation preference data balances several factors related to the patient's objectives, reduces the cost associated with in-clinic measurements, and increases the number of patient evaluations of settings.

The personalized preference models obtained for the five E-STAND participants show that there is more similarity in optimal frequency than pulse width across participants. Validation studies and analysis indicate that there are benefits in using preference models to determine optimal stimulation settings. Both internal cross-validation and prospective validation studies show that the prediction accuracy of the preference models produced through optimization are significantly higher than chance. Across five participants, the preference scores from individual preference model are significantly positively correlated with motor outcome and treatment month at a population level.

Overall, the results indicate that the Bayesian preference optimization algorithm may assist clinicians in the systematic programming of individualized therapeutic stimulation settings. Quality of life scores also indicate significant population-level improvement through the course of the therapy.

To our knowledge, our work is the first to systematically validate eSCS optimization results across multiple participants in terms of both the algorithmic accuracy and clinical outcome. The validation of clinical outcome can be further improved by constructing more interpretable and revealing movement metrics, such as leg travel distance, designing precision measured tasks, and including muscle activation measures, such as EMG data.

References

[1] P. L. Gildenberg, "History of electrical neuromodulation for chronic pain," *Pain Med.*, vol. 7, pp. S7–S13, May 2006.
[2] A. W. Cook, "Electrical stimulation in multiple sclerosis," *Hospital Pract.*, vol. 11, no. 4, pp. 51–58, Apr. 1976.
[3] A. Seifalian, M. Tsintou, and K. Dalamagkas, "Advances in regenerative therapies for spinal cord injury: A biomaterials approach," *Neural Regeneration Res.*, vol. 10, no. 5, p. 726, 2015.
[4] A. K. Varma *et al.*, "Spinal cord injury: A review of current therapy, future treatments, and basic science frontiers," *Neurochem. Res.*, vol. 38, no. 5, pp. 895–905, May 2013.
[5] D. Darrow *et al.*, "Epidural spinal cord stimulation facilitates immediate restoration of dormant motor and autonomic supraspinal pathways after chronic neurologically complete spinal cord injury," *J. Neurotrauma*, vol. 36, no. 15, pp. 2325–2336, Aug. 2019, doi: 10.1089/neu.2018.6006.

[6] V. R. Edgerton and S. Harkema, "Epidural stimulation of the spinal cord in spinal cord injury: Current status and future challenges," *Expert Rev. Neurotherapeutics*, vol. 11, no. 10, pp. 1351–1353, Oct. 2011.

[7] M. L. Gill *et al.*, "Neuromodulation of lumbosacral spinal networks enables independent stepping after complete paraplegia," *Nature Med.*, vol. 24, no. 11, pp. 1677–1682, Nov. 2018, doi: 10.1038/s41591-018-0175-7.

[8] M. Capogrosso *et al.*, "Configuration of electrical spinal cord stimulation through real-time processing of gait kinematics," *Nature Protocols*, vol. 13, no. 9, pp. 2031–2061, Sep. 2018.

[9] D. G. Sayenko, C. Angeli, S. J. Harkema, V. R. Edgerton, and Y. P. Gerasimenko, "Neuromodulation of evoked muscle potentials induced by epidural spinal-cord stimulation in paralyzed individuals," *J. Neurophysiol.*, vol. 111, no. 5, pp. 1088–1099, Mar. 2014.

[10] F. B. Wagner *et al.*, "Targeted neurotechnology restores walking in humans with spinal cord injury," *Nature*, vol. 563, no. 7729, pp. 65–71, Nov. 2018.

[11] M. Capogrosso *et al.*, "A computational model for epidural electrical stimulation of spinal sensorimotor circuits," *J. Neurosci.*, vol. 33, no. 49, pp. 19326–19340, Dec. 2013.

[12] T. A. Desautels *et al.*, "An active learning algorithm for control of epidural electrostimulation," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 10, pp. 2443–2455, Oct. 2015.

[13] P. K. Shah *et al.*, "Unique spatiotemporal neuromodulation of the lumbosacral circuitry shapes locomotor success after spinal cord injury," *J. Neurotrauma*, vol. 33, no. 18, pp. 1709–1723, Sep. 2016.

[14] S. Harkema *et al.*, "Effect of epidural stimulation of the lumbosacral spinal cord on voluntary movement, standing, and assisted stepping after motor complete paraplegia: A case study," *Lancet*, vol. 377, no. 9781, pp. 1938–1947, Jun. 2011.

[15] E. Brochu, V. M. Cora, and N. de Freitas, "A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning," 2010, *arXiv:1012.2599*. [Online]. Available: http://arxiv.org/abs/1012.2599

[16] Y. Sui, Y. Yue, and J. W. Burdick, "Correlational dueling bandits with application to clinical treatment in large decision spaces," 2017, *arXiv:1707.02375*. [Online]. Available: http://arxiv.org/abs/1707.02375

[17] Y. Sui, V. Zhuang, J. W. Burdick, and Y. Yue, "Stagewise safe Bayesian optimization with Gaussian processes," 2018, *arXiv:1806.07555*. [Online]. Available: http://arxiv.org/abs/1806.07555

[18] F. Hartig. (2020). *DHARMa: Residual Diagnostics for Hierarchical (Multi-Level/Mixed) Regression Models. R Package Version 0.1.6*. CRAN/GitHub. [Online]. Available: https://CRAN.R-project.org/package=DHARMa

[19] D. S. Tulsky *et al.*, "Overview of the spinal cord injury–quality of life (SCI-QOL) measurement system," *J. Spinal Cord Med.*, vol. 38, no. 3, pp. 257–269, 2015.

[20] I. Peña Pino *et al.*, "Long-term spinal cord stimulation after chronic complete spinal cord injury enables volitional movement in the absence of stimulation," *Frontiers Syst. Neurosci.*, vol. 14, p. 35, Jun. 2020.

[21] W. Chu and Z. Ghahramani, "Preference learning with Gaussian processes," in *Proc. 22nd Int. Conf. Mach. Learn. (ICML)*, 2005, pp. 137–144.