International Journal of General Medicine

Dovepress

Open Access Full Text Article

ORIGINAL RESEARCH

Predicting Early Treatment Effectiveness in Bell's Palsy Using Machine Learning: A Focus on Corticosteroids and Antivirals

Jheng-Ting Luo[1], Yung-Chun Hung[1,2], Gina Jinna Chen[3], Yu-Shiang Lin[1]

[1]In-Service Master Program in Artificial Intelligence in Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan; [2]School of Health Care Administration, College of Management, Taipei Medical University, Taipei, Taiwan; [3]Department of Electronic and Electrical Engineering, Southern University of Science and Technology, Shenzhen, People's Republic of China

Correspondence: Yu-Shiang Lin, In-Service Master Program in Artificial Intelligence in Medicine, College of Medicine, Taipei Medical University, No. 250, Wuxing St., Xinyi District, Taipei City, 11031, Taiwan, Email eriklin@tmu.edu.tw

Purpose: Facial nerve paralysis, particularly Bell's palsy, manifests as a rapid onset of unilateral facial weakness or paralysis. Despite most patients recovering within three to six months, a significant proportion experience poor recovery. This study utilized six machine learning models to investigate the effectiveness of early treatment in Bell's palsy.

Patients and Methods: We applied data from 17 hospitals in Scotland to predict treatment outcomes. Patients were randomized into four groups: Prednisolone (corticosteroids), Acyclovir (antivirals), both, and placebo. Outcomes, defined as full resolution of symptoms, were assessed using the House-Brackmann scale at 3 and 9 months post-treatment. We employed six different machine learning models to predict recovery outcomes and evaluated model performance using AUC, precision, recall, and F1-score.

Results: Among 493 patients, 72.6% recovered after three months and 89.5% after nine months. Logistic regression demonstrated the highest predictive performance for both 3-month ($AUC = 0.751$) and 9-month recovery ($AUC = 0.720$). Additionally, several models achieved Precision levels exceeding 0.9. We further employed the best-performing logistic regression for feature ranking, indicating that the patient's age and prednisolone administration are the most significant predictors of recovery.

Conclusion: The results highlight the potential of machine learning models in predicting the effectiveness of early treatment. This study conducted a comprehensive comparison of six different machine learning models, with the logistic regression showing the highest predictive performance for both 3-month and 9-month recovery. Additionally, feature ranking using logistic regression supported the importance of Prednisolone in treatment. Notably, our findings revealed the significance of age in prognosis evaluation for the first time. This suggests that future research should further develop age-specific prognostic models, enabling clinicians to tailor individualized treatment strategies more effectively. This previously unrecognized discovery provides a foundation for prognostic analysis in Bell's palsy patients.

Keywords: Bell's palsy, machine learning, prognostic prediction, feature ranking

Introduction

Facial nerve paralysis is the loss of facial movement due to pathological changes in the facial nerve, leading to impaired function of the voluntary facial muscles innervated by the nerve, resulting in facial asymmetry.[1] Bell's palsy is the most common peripheral paralysis of the seventh cranial nerve. This idiopathic condition has a rapid onset.[2] Bell's palsy should be suspected in patients with acute onset of unilateral facial weakness or paralysis involving the forehead in the absence of other neurologic abnormalities.[3] The annual incidence is 15 to 20 per 100,000, with 40,000 new cases yearly. The lifetime risk is 1 in 60, and the recurrence rate is 8% to 12%. Most individuals start to recover within two weeks and are fully recovered within three to six months.[4] Although 70% of patients experience complete resolution without treatment, up to 30% have a poor recovery. The major cause of Bell's palsy is believed to be a viral infection of the facial nerve. The viral illnesses include herpes simplex virus, varicella-zoster virus, and Epstein-Barr virus.[5–10] Due to this viral

Received: 1 September 2024
Accepted: 2 November 2024
Published: 9 November 2024

infection, the facial nerve swells as it passes through the temporal bone and becomes compressed within the canal. This condition leads to the patient being unable to control the muscles on the affected side of the face.[11,12]

Recovery of facial nerve function is an important indicator guiding treatment recommendations. The House-Brackmann facial nerve grading system can be used to describe the degree of facial nerve weakness and evaluate treatment outcomes.[2,12,13] The primary pharmacologic therapy for Bell's palsy or facial nerve palsy is early short-term oral corticosteroid treatment. According to a review published by the American Academy of Family Physicians (AAFP) in 2023, an oral corticosteroid regimen (Prednisone, 50 to 60 mg per day for five days, followed by a five-day taper) is the first-line treatment. Starting this treatment within 48 hours of symptom onset yields the most effective results. The potential benefit of adding antiviral therapy to the corticosteroid regimen for patients with new-onset Bell's palsy remains uncertain. Due to the absence of more comprehensive data, we recommend that severe facial paralysis patients categorized as House-Brackmann grade IV or higher should receive a combination of oral antiviral medication and corticosteroids. This combination therapy may also reduce the incidence of synkinesis (involuntary contraction of certain facial muscles). Recommended antivirals include Acyclovir (400 mg five times per day for ten days). Treatment with antivirals alone is ineffective and not recommended.[14–18]

Previous studies have demonstrated the effectiveness of machine learning models in predicting clinical outcomes for various medical conditions and identifying prognostic factors. For instance, an AI-enabled ECG algorithm using a convolutional neural network was developed to detect atrial fibrillation during normal sinus rhythm, showing promising results in outcome prediction.[19] In another study, gradient-boosting decision trees (GBDT) were compared with logistic regression (LR) for diabetes prediction. The results revealed that GBDT models demonstrated greater reliability and accuracy, especially when analyzing big data.[20] In the prediction of cardiovascular risk, the integration of artificial neural networks (ANN) has been demonstrated to be effective and exhibit high accuracy.[21,22] Dong et al illustrate this through their development of six ML-based prognostic models designed to predict the overall survival of patients with alpha-fetoprotein (AFP)-positive hepatocellular carcinoma (HCC).[23] By leveraging algorithms such as XGBoost, the study demonstrates the robust performance of ML in forecasting survival outcomes at 1-, 3-, and 5-year intervals. The models' ability to accurately process a range of demographic and clinicopathological features provides valuable insights into patient prognosis. The XGBoost model's superior performance across multiple time points highlights the potential for these tools to assist physicians in personalizing treatment strategies for patients with AFP-positive HCC. Peng et al demonstrated this potential through the development of ML-based prognostic models to predict overall survival for prostate cancer patients with lymph node involvement. The study applied Gradient Boosting Survival Analysis (GBSA), Random Survival Forest (RSF), and Extra Survival Trees (EST) algorithms, comparing their performance against the Cox regression model. Across all models, the ML-based approaches demonstrated higher accuracy in predicting patient survival, outperforming the Cox regression model. The study also developed a web-based tool, making these advanced ML models more accessible to clinicians, which improves early medical interventions and treatment customization.[24] Doja et al explored the utility of machine learning in predicting survival outcomes for metastatic prostate cancer, which typically has a poor prognosis compared to earlier stages. The study developed an ensemble approach that provided the best results, with an accuracy of 81.4% for the entire dataset and 83.74% for the age-specific model. These findings underscore the potential of machine learning to enhance survival predictions by accounting for variations across different patient subgroups. Patients predicted to have better survival outcomes might be considered for more aggressive therapies, while those with poorer prognoses could benefit from palliative care and closer monitoring.[25] Age has consistently emerged as a significant prognostic factor in machine learning-based models, highlighting its crucial role in determining patient outcomes. In the study by Dong et al, age was identified as one of the most important variables influencing survival outcomes.[23] Similarly, Peng et al examined the prognostic factors for prostate cancer patients with lymph node-positive disease and utilized stepwise selection methods to identify eight key variables, including age.[24] In the study by Doja et al, the age-specific model, which achieved higher accuracy (83.74%) than the model for the complete dataset, demonstrated the significance of age as a predictor. These results highlight the importance of incorporating age into prognostic models.[25] By personalizing care based on prognosis, physicians can optimize resource allocation and provide more effective and timely interventions, ultimately enhancing patient outcomes. In summary, machine learning models have been proven effective in predicting various diseases.

However, there remains a significant gap in research regarding the prognostic prediction of Bell's palsy using machine learning models. Recent studies on machine learning in Bell's palsy have primarily focused on classification and diagnosis. For instance, a study on the classification of facial paralysis developed an evaluation system using real-time facial animation units (FAUs) and ensemble learning classifiers, achieving high accuracy, sensitivity, and specificity in classification outcomes.[26] Another study introduced a deep-learning-based method for the early diagnosis of facial paralysis diseases, including Bell's palsy and stroke, using a multi-task network to enhance diagnostic accuracy.[27] Lastly, research on predicting synkinesis caused by Bell's palsy or Ramsay Hunt syndrome utilized various machine learning algorithms to improve prediction accuracy over conventional methods.[28]

Despite extensive research using machine learning methods on Bell's palsy in recent years, most studies predominantly focused on severity classification or diagnosis rather than treatment effectiveness or prognostic prediction. Prognostic prediction provides the following significant benefits for patients with Bell's palsy. First, machine learning offers an effective method for prognostic prediction in large and complex datasets. Second, rather than preselecting factors for model development, machine learning allows the data to reveal which features are important for specific predictions.[29] According to Tedeschi et al, while several validated tools are widely used for assessing peripheral facial paralysis, including Bell's palsy, these tools have limitations that impact their effectiveness in clinical practice. The House-Brackmann scale, although extensively studied and easy to use, is limited by its low inter-rater reproducibility and inability to capture subtle differences in facial movement due to its broad categorization into six severity levels. This creates challenges for clinicians in accurately monitoring progress, particularly in patients with mild or moderate paralysis.[30] Therefore, our research primarily focuses on utilizing machine learning to establish prognostic models aimed at enhancing the accuracy of assessments for peripheral facial nerve palsy. In this research, we utilize six different machine-learning models to predict the effectiveness of early treatment with corticosteroids or antiviral agents in Bell's palsy patients. Although clinical guidelines for Bell's palsy are available, no related research has been done in the field of machine learning. Our study makes a significant medical contribution by employing machine learning models to address this gap. Distinguishing between House-Brackmann grade III (obvious but not disfiguring difference between the two sides) and grade IV (obvious weakness and/or disfiguring asymmetry) may pose challenges as well. Additionally, we explored the significant factors influencing treatment effectiveness. This innovative approach not only enhances the predictive accuracy for early treatment outcomes but also provides valuable insights into the critical factors that influence recovery in Bell's palsy patients.

Materials and Methods

Dataset

We used a public dataset from Kaggle, a data science competition platform.[31] This dataset comprises data from 17 hospitals across Scotland, where potential patients with Bell's palsy were referred to specialized receiving centers. Eligible patients were randomly assigned to different study groups and monitored for a period of 9 months. The inclusion criteria targeted adults aged 16 years or older who presented with unilateral facial nerve weakness of no identifiable cause. These patients were initially seen in primary care or emergency departments and referred to a collaborating otorhinolaryngologist within 72 hours of symptom onset. Exclusion criteria included pregnancy, breastfeeding, uncontrolled diabetes (glycated hemoglobin level > 8%), peptic ulcer disease, suppurative otitis media, herpes zoster, multiple sclerosis, systemic infection, sarcoidosis, and other rare conditions.

The initial study was carried out with the authorization of the Multicenter Research Ethics Committee for Scotland. All patients provided written informed consent after the aims and methods of the study had been described to them and after they had received an information sheet. Consequently, ethical approval was deemed unnecessary for the current secondary analysis. Furthermore, the initial research was conducted in accordance with the provisions of the Declaration of Helsinki and Good Clinical Practice guidelines, ensuring all protocols complied with relevant norms and regulations.[32] This secondary analysis followed the same approach.

Study Design

From June 2004 to June 2006, patients were recruited through various channels, including family doctors, emergency departments, the national 24-hour medical telephone consultancy service, and dentists' offices. The study maintained blinding for patients, recruiters, study visitors, and outcome assessors regarding group assignments. Patients underwent randomization twice, resulting in four study groups: Prednisolone (at a dose of 25 mg twice daily) and placebo (lactose), Acyclovir (400 mg five times daily) and placebo, Prednisolone and Acyclovir, and two placebo capsules. Each patient received two bottles of odorless capsules with an identical appearance to ensure blinding.

Within 3 to 5 days after randomization, a researcher conducted a baseline assessment either at the patient's home or a doctor's office. Follow-up assessments were conducted at three months. If recovery was incomplete (grade 2 or higher on the House-Brackmann scale) at this visit, an additional assessment was conducted at nine months.[32]

After thorough consideration, we collected 493 patients and ten variables related to patient characteristics. These variables comprise patient ID, sex, age, baseline score on the House-Brackmann scale, time between symptom onset and treatment initiation, treatment group, receipt of Prednisolone, receipt of Acyclovir, full recovery at three months, and full recovery at nine months.

Study Outcomes

Recovery of facial nerve function is an important outcome, and the initial severity of facial weakness provides valuable prognostic information.[11] To objectively assess facial function among Bell's palsy patients, clinicians employ standardized scales, with the House-Brackmann facial nerve grading system being the predominant choice. This system categorizes facial function into six grades, ranging from grade I (no weakness) to grade VI (complete weakness) (Table 1).[12–14]
The study outcomes include two key indicators:

1. Full recovery at three months: This indicates the patient's facial nerve function has returned to grade I within three months.
2. Full recovery at nine months: This indicates the patient's facial nerve function has returned to grade I within nine months.

Model Development

We perform model development and statistical analysis in JupyterLab. Initially, the overall study was randomized into a training (80%) and testing dataset (20%) as per the Pareto principle.[33]

Table 1 House-Brackmann Facial Nerve Grading System

Grade	Definition
I	Normal facial function in all areas
II	Gross: slight weakness noticeable on close inspection; may have very slight synkinesis At rest: normal symmetry and tone Motion: forehead - moderate to good function; eye - complete closure with minimum effort; mouth - slight asymmetry.
III	Gross: obvious but not disfiguring difference between two sides; noticeable but not severe synkinesis, contracture, and/or hemi-facial spasm. At rest: normal symmetry and tone Motion: forehead - slight to moderate movement; eye - complete closure with effort; mouth - slightly weak with maximum effort.
IV	Gross: obvious weakness and/or disfiguring asymmetry At rest: normal symmetry and tone Motion: forehead - none; eye - incomplete closure; mouth - asymmetric with maximum effort.
V	Gross: only barely perceptible motion At rest: asymmetry Motion: forehead - none; eye - incomplete closure; mouth - slight movement
VI	No movement

We preprocessed the dataset. Among the ten patient characteristics, seven were identified as categorical variables: sex, the time between the onset of symptoms and the start of treatment, treatment group, received Prednisolone, received Acyclovir, full recovery in 3 months, and full recovery in 9 months. After identifying the categorical variables, we proceeded with data labeling.

The study outcomes in the dataset exhibit an imbalance. This includes "full recovery in 3 months" and "full recovery in 9 months". We resolved data imbalance in classification by employing the random oversampling approach[34] for the training dataset. Subsequently, we evaluated the model's performance through 5-fold cross-validation. Cross-validation is a key technique in machine learning because it helps evaluate a model's performance on unseen data and detect overfitting. By using cross-validation, the risk of overfitting is mitigated, leading to a more accurate assessment of the model's performance.[35]

To predict study outcomes in Bell's palsy, we employed six methods from different machine learning areas. In addition to focusing on improving predictive accuracy, our rationale for selecting machine learning models also prioritized ensuring better interpretability of the results.[36] This allows for a more effective analysis of the key features driving the predictions, which led us to select the following six machine learning models for the predictive analysis.

1. Logistic regression: A process of modeling the probability of a discrete outcome given an input variable.[37]
2. eXtreme Gradient Boosting (XGBoost): A decision-based tree ensemble algorithm that is classified as embedded feature selection.[38]
3. Support vector machine (SVM): A machine learning algorithm that categorizes data by identifying an optimal line or hyperplane that maximizes the margin between classes.[39–41]
4. Random forest: An ensemble of tree predictors, which builds several decision trees during training and outputs the mode of the classes for classification.[42]
5. Decision tree: A decision support tool that uses a tree-like model of decisions and their possible consequences.[37]
6. Multilayer perceptron (MLP): A supplement of a feed-forward neural network that consists of at least three layers of nodes.[43]

By employing these diverse machine learning models, we aimed to identify the most effective approach for predicting treatment outcomes in patients with Bell's palsy. The performance of each model was rigorously evaluated to ensure the reliability and accuracy of our predictions.

Model Evaluation

Model performance was assessed with the area under the receiver operating characteristic (ROC) curve (AUC),[44] precision, recall, and F1 score. All evaluation metrics were obtained from the test dataset. A confusion matrix is important in assessing the classification efficiency of all classifiers. It offers details about the real and forecast classifications.[45] There are four components in the confusion matrix: true positive (TP), true negative (TN), false positive (FP), and false negative (FN). The evaluation indices are defined as:

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{F1 score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

TP indicates cases where the model accurately predicts that patients will achieve full recovery within the specified time frame (either three months or nine months, as indicated in the dataset). Conversely, *TN* signifies cases where the model correctly predicts that patients will not achieve full recovery within the same timeframe. These definitions of *TP* and *TN* are essential for evaluating the model's performance. By analyzing the confusion matrix, we can better understand the model's strengths and weaknesses, thereby enabling more informed decisions regarding the use of corticosteroids or antiviral agents in the early treatment of Bell's palsy.

Results

Baseline Characteristics

A total of 493 patients were eligible for analysis, and the baseline characteristics of clinical information are presented in Table 2. The mean age of the patients was 44.9 ± 14.6 years, and 48.3% were female. Out of the 250 people who received Prednisolone and the 248 people who received Acyclovir, a total of 127 people simultaneously took both medications. After treatment, 72.6% of patients recovered after three months, and nearly ninety percent (89.5%) recovered after nine months.

Model Performance

We employed six machine learning models, including logistic regression, eXtreme Gradient Boosting (XGBoost), support vector machine (SVM), random forest, decision tree, and multilayer perceptron (MLP), to predict the recovery outcomes after treatment. First, we investigate the full recovery in 3 months. Table 3 illustrates the performance of each machine-learning model. Compared with the other models, the models built using the logistic regression had the highest AUC value (*AUC* = 0.751). The Decision Tree model achieved the highest precision (0.850), while the Random Forest model demonstrated the best recall (0.765) and a competitive F1 score (0.769) (Figure 1A). Another study outcome is full recovery in 9 months, utilizing the aforementioned machine learning models. The performance of each machine learning model is depicted in Table 3. Similarly, we analyze the performance of each machine-learning model. Compared with the other models, the models built using the logistic regression had the highest AUC value (*AUC* = 0.720). Notably, the precision for predicting a full recovery in 9 months was exceptional across several models, with Logistic Regression, SVM, Decision Tree, and MLP all achieving precision scores above 90%. The Logistic Regression model, which achieved the highest AUC value, also showed a remarkable precision of 0.941. The Decision Tree demonstrated the highest precision (0.945), while the Random Forest model excelled with the best recall (0.909) and a high F1 score

Table 2 Baseline Characteristics of Bell's Palsy Patient

Characteristics	All (*n* = 493)
Age (Years)	44.9 ± 14.6
Baseline score on House-Brackmann scale	3.7 ± 1.1
Gender, *n* (%)	
Male	255 (51.7)
Female	238 (48.3)
Time between onset of symptoms and start of treatment, *n* (%)	
Within 24 hr	249 (50.5)
>24 to ≤48 hr	205 (41.6)
>48 to ≤72 hr	39 (7.9)
Treatment Group, *n* (%)	
Prednisolone–Placebo	123 (24.9)
Acyclovir–Prednisolone	127 (25.8)
Acyclovir–Placebo	121 (24.5)
Placebo–Placebo	122 (24.7)
Received Prednisolone, *n* (%)	250 (50.7)
Received Acyclovir, *n* (%)	248 (50.3)
Full Recovery in 3 Months, *n* (%)	358 (72.6)
Full Recovery in 9 Months, *n* (%)	441 (89.5)

Table 3 Comparison of the Performance of Different Machine Learning Models

Algorithm	AUC	Precision	Recall	F1 score
Prediction of full recovery in 3 months				
LR	0.751	0.837	0.687	0.755
XGBoost	0.663	0.766	0.749	0.757
SVM	0.711	0.835	0.662	0.738
Random Forest	0.676	0.772	0.765	0.769
Decision Tree	0.710	0.850	0.556	0.672
MLP	0.742	0.842	0.626	0.718
Prediction of full recovery in 9 months				
LR	0.720	0.941	0.685	0.793
XGBoost	0.597	0.899	0.891	0.895
SVM	0.691	0.929	0.537	0.681
Random Forest	0.609	0.897	0.909	0.903
Decision Tree	0.669	0.945	0.580	0.719
MLP	0.709	0.937	0.669	0.780

Abbreviations: AUC, area under the receiver operating characteristic (ROC) curve; LR, logistic regression; XGBoost, eXtreme Gradient Boosting; SVM, support vector machine; MLP, multilayer perceptron.

(0.903) (Figure 1B). Notably, the number of recovered patients at 3 months and 9 months were 358 and 441, respectively. We found that the precision of the Decision Tree model at 3 months and 9 months was 0.850 and 0.945, respectively, with a statistically significant difference (p<0.01). Additionally, the recall for the Random Forest model at 3 months and 9 months was 0.765 and 0.909, also showing a statistically significant difference (p<0.01). Furthermore, the F1 score for the 3-month and 9-month models was 0.769 and 0.903, respectively, with a statistically significant difference (p<0.01). These results further support the conclusion that the 9-month timeframe is sufficient to observe meaningful outcomes in Bell's palsy recovery.

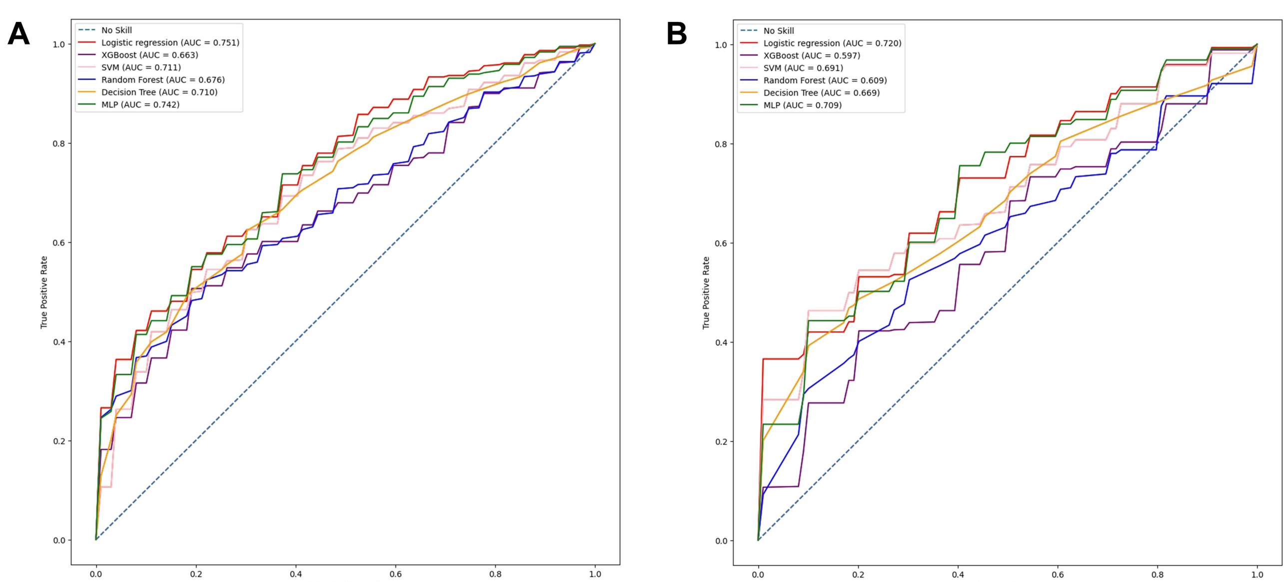

Figure 1 Plots of the model performance. (**A**), Receiver operating curve (ROC) in predicting full recovery in 3 months. (**B**), Receiver operating curve (ROC) in predicting a full recovery in 9 months.

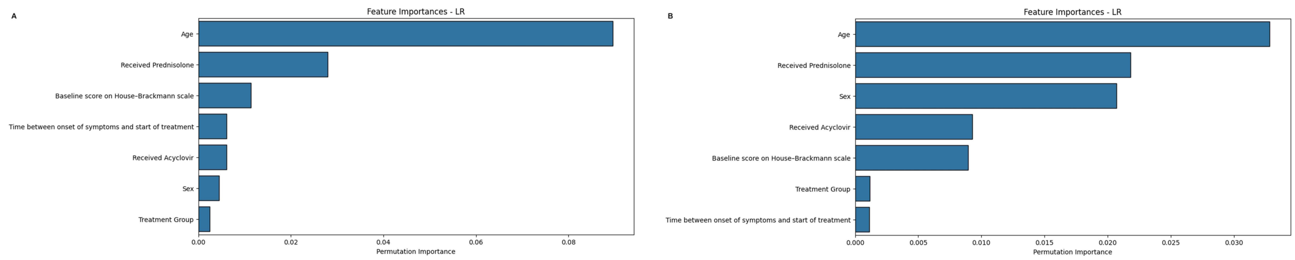

Figure 2 Feature importance analysis. (**A**), Full recovery in 3 months. (**B**), Full recovery in 9 months.

Feature Ranking

To investigate the factors influencing recovery in Bell's palsy, we perform feature selection in this dataset. Given that the logistic regression model achieved the highest AUC among the six machine learning models for predicting full recovery at both 3- months and 9-month, we selected logistic regression for the analysis of feature importance. Subsequently, logistic regression was used to generate the ranking results. The ranking results, as depicted in Figure 2, provide a detailed analysis of the importance of features in predicting full recovery at both three months and nine months.

For full recovery in 3 months (Figure 2A), age emerged as the most significant predictor with the highest permutation importance. The administration of Prednisolone was identified as the second most important factor, highlighting its therapeutic efficacy. Similarly, for full recovery in 9 months (Figure 2B), age remained the most significant predictor. This underscores its long-term influence on recovery. The administration of Prednisolone continued to be a contributing factor.

Despite the differences in permutation importance values for each feature between the 3-month and 9-month recovery periods, the results are consistent. Age and the administration of Prednisolone consistently ranked as the top two most important factors in both timeframes. These findings collectively highlight the dominant role of age and the therapeutic value of Prednisolone in both short-term and long-term recovery from Bell's palsy. Understanding the importance of these features can guide clinicians in making informed decisions. This insight enables them to adjust treatments based on individual patient characteristics, thereby optimizing clinical outcomes.

Discussion

Machine learning models are adept at managing large and complex datasets, improving the accuracy of prognostic predictions compared to traditional methods. Despite the growing interest in applying machine learning to prognostic prediction, several areas related to Bell's palsy remain under-researched.

In comparing our study on predicting treatment effectiveness in Bell's palsy with the work of Gaber et al,[26] it is evident that while both studies leverage machine learning techniques to address issues related to facial paralysis, they diverge in their primary objectives and methodologies. Gaber et al focus on classifying the severity of facial paralysis using real-time facial animation units (FAUs), emphasizing a quantitative assessment of facial paralysis severity rather than treatment prediction. In contrast, we aim to identify predictive factors that influence treatment success, thereby offering a clinical tool to optimize therapeutic strategies. In comparing our study on early treatment effectiveness prediction in Bell's palsy with the work of Umirzakova et al,[27] notable distinctions emerge regarding focus and methodology. While both studies emphasize the significance of early intervention, our research is pioneering in its application of machine learning to predict clinical treatment outcomes. Conversely, Umirzakova et al focus on developing a deep-learning-based diagnostic framework aimed at early detection of facial paralysis conditions, such as Bell's palsy and stroke. This image-based diagnostic emphasis, while innovative, does not directly address treatment outcomes, thus limiting its applicability in clinical prognosis. Our study focuses on clinical treatment outcomes, while the other emphasizes diagnosis using geometric and semantic cues in images. In comparing our study on predicting early treatment effectiveness in Bell's palsy with the research by Kishimoto-Urata et al,[28] we observe both similarities and differences in focus and methodology within the realm of machine learning applications in facial paralysis. Our study is groundbreaking in its exploration of logistic regression models to predict treatment outcomes. Conversely, Kishimoto-Urata et al focus on predicting synkinesis. Their study demonstrates a higher predictive capability for synkinesis, with an AUC of 0.910 based on electrophysiological tests, nerve excitability, and scaling systems, which is higher than our study's AUC. However, our study's focus on early treatment effectiveness is a different clinical outcome, emphasizing recovery instead of synkinesis probability. The common themes reveal an effort to leverage machine learning in clinical applications pertaining to Bell's palsy, encompassing classification, diagnosis, treatment prediction, and complication forecasting. Our research notably concentrates on treatment outcomes, while the other studies focus primarily on diagnostic accuracy and severity assessment. Moreover, feature selection plays a crucial role across all studies. Our study highlights age and Prednisolone as key predictors for Bell's palsy recovery, which had not been previously investigated. Future research might benefit from integrating the diagnostic methods in deep learning studies with our treatment outcome predictions to create a holistic predictive model.

To our knowledge, this study represents the first investigation into machine learning models aimed at predicting the effectiveness of early corticosteroid and antiviral treatment in patients with Bell's palsy. We employed six machine learning models to predict recovery outcomes and evaluated model performance using AUC, precision, recall, and F1 score. Our research findings demonstrate that the logistic regression model exhibited the highest predictive performance for both 3-month and 9-month recovery outcomes. Specifically, the logistic regression model achieved an AUC of 0.751 for a 3-month recovery and 0.720 for a 9-month recovery. Notably, the logistic regression also showed a remarkable precision of 0.941 for a 9-month recovery. The results indicate the robustness and reliability of the logistic regression model in predicting Bell's palsy treatment outcomes.

Additionally, we performed a feature ranking analysis based on logistic regression. The feature ranking analysis revealed that the patient's age and Prednisolone administration were the most significant predictors of recovery. Current clinical guidelines primarily prioritize the use of corticosteroids in initial treatment decisions. Our results are consistent with previous research that has highlighted the efficacy of corticosteroids in improving recovery rates among Bell's palsy patients. This further strengthens the robustness of the existing literature. Notably, we also discovered for the first time that the patient's age outweighs the importance of Prednisolone administration (Figure 2). This not only confirms the importance of corticosteroids emphasized in current guidelines but also identifies age as a key factor influencing recovery. Objective quantification of important factors can help refine existing treatment guidelines and ensure that clinicians incorporate age into treatment plans.

The identification of age and Prednisolone administration as critical predictors of recovery holds substantial implications for tailoring treatment plans. The broader impact of these findings lies in the potential for more targeted and effective treatment regimens, especially for patients who may be at higher risk for poor recovery. The main contribution of this study lies in emphasizing the role of Prednisolone, reinforcing the importance of early corticosteroid use, which aligns with current guidelines, while also providing quantitative evidence to support its continued use. Additionally, our findings that patient age significantly impacts prognosis introduce a new dimension to patient assessment, which may lead to more age-specific treatment considerations. This underscores the importance of early intervention, especially in older patients who may face a higher likelihood of incomplete recovery. From a clinical practice standpoint, the integration of machine learning models can enable clinicians to better stratify patients based on their risk of poor recovery, allowing for more aggressive or tailored treatments when necessary. Furthermore, these findings call for additional research into how machine learning methods can be extended and applied to other conditions with similar prognostic uncertainty. Finally, this study underscores the broader potential of machine learning as a predictive tool, highlighting its valuable role in advancing precision medicine.

This study's primary contribution lies in shaping future research directions and treatment strategies. For the first time, we have identified that the patient's age outweighs the importance of Prednisolone administration in prognostic prediction. Future research can comprehensively explore the impact of age as a prognostic factor across diverse populations. Our future work involves further research to validate our model in different populations and settings to ensure its broader applicability. We will replicate the proposed method with larger, more diverse datasets, which would strengthen its plausibility and generalizability. In the meantime, more varied data-balancing strategies will be considered. Expanding the dataset to include diverse populations and additional variables could further enhance the predictive accuracy of our model. Moreover, our study found that age is a significant factor influencing classification outcomes. Therefore, incorporating age into further prognostic analysis is crucial. Future studies could encompass a wider age range, including adolescent populations under 16 years old. Personalized treatment plans could include targeted interventions, such as optimizing medication regimens based on age-specific considerations. For instance, younger patients may respond differently to medications or therapies compared to elderly patients. Therefore, tailoring treatment plans according to the patient's age can ultimately improve recovery rates and overall patient care.

Conclusions

This study demonstrates the potential of machine learning models in predicting recovery outcomes of early treatment in Bell's palsy patients. Particularly noteworthy is the logistic regression model, demonstrating the highest predictive performance in assessing recovery at both three months ($AUC = 0.751$) and nine months ($AUC = 0.720$). Notably, the logistic regression also showed a remarkable precision of 0.941 for a 9-month recovery. The quantitative rankings reveal that the patient's age and prednisolone administration are significant predictive factors for recovery. Our findings represent the initial discovery that the patient's age outweighs the importance of Prednisolone administration in recovery predictions. Age emerged as the top predictor, emphasizing its critical role in both short-term and long-term recovery outcomes. Prednisolone administration also demonstrated substantial importance, confirming its therapeutic value as outlined in current clinical guidelines. The study underscores the potential of machine learning in enhancing prognostic predictions for Bell's palsy treatment outcomes. By identifying age and Prednisolone administration as crucial predictive factors, researchers can further explore the impact of age on prognosis prediction. Age-specific therapeutic strategies could enhance personalized treatment. Further research across diverse healthcare settings may validate these models, advancing personalized medicine in facial nerve paralysis management.

Data Sharing Statement

The publicly available dataset presented in this study are available at the following website: https://www.kaggle.com/datasets/dillonmyrick/bells-palsy-clinical-trial.

Funding

This research was funded by the National Science and Technology Council, Taiwan (grant number NSTC 113-2222-E-038-001-MY3) and the Taipei Medical University (grant number TMU111-AE1-B30).

Disclosure

The authors report no conflicts of interest in this work.

References

1. Song A, Wu Z, Ding X, Hu Q, Di X. Neurologist standard classification of facial nerve paralysis with deep neural networks. *Future Internet*. 2018;10(11):111. doi:10.3390/fi10110111
2. Warner MJ, Hutchison J, Varacallo M. *Bell Palsy*. StatPearls; 2023. Available from: https://www.ncbi.nlm.nih.gov/books/NBK482290. Accessed November 06, 2024.
3. Dalrymple SN, Row JH, Gazewood J. Bell palsy: rapid evidence review. *Am Family Phys*. 2023;107(4):415–420.
4. Singh A, Deshmukh P. Bell's palsy: a review. *Cureus*. 2022;14(10):e30186. doi:10.7759/cureus.30186
5. James DG. All that palsies is not Bell's. *J R Soc Med*. 1996;89(4):184. doi:10.1177/014107689608900403
6. Jackson CG, von Doersten PG. The facial nerve: current trends in diagnosis, treatment, and rehabilitation. *Med Clin North Am*. 1999;83(1):179. doi:10.1016/S0025-7125(05)70096-1
7. Stjernquist-Desatnik A, Skoog E, Aurelius E. Detection of herpes simplex and varicella-zoster viruses in patients with Bell's palsy by the polymerase chain reaction technique. *Ann Otol Rhinol Laryngol*. 2006;115(4):306–311. doi:10.1177/000348940611500410
8. Theil D, Arbusow V, Derfuss T, et al. Prevalence of HSV-1 LAT in human trigeminal, geniculate, and vestibular ganglia and its implication for cranial nerve syndromes. *Brain Pathol*. 2001;11(4):408–413. doi:10.1111/j.1750-3639.2001.tb00408.x
9. Looker KJ, Magaret AS, May MT, et al. Global and regional estimates of prevalent and incident herpes simplex virus Type 1 infections in 2012. *PLoS One*. 2015;10:e0.
10. Kennedy PG, Rovnak J, Badani H, Cohrs RJ. A comparison of herpes simplex virus type 1 and varicella-zoster virus latency and reactivation. *J Gen Virol*. 2015;96(Pt 7):1581–1602. doi:10.1099/vir.0.000128
11. Grewal DS. Bell's palsy-tertiary ischemia: an etiological factor in residual facial palsy. *Indian J Otolaryngol Head & Neck Surg*. 2018;70(3):374–379. doi:10.1007/s12070-018-1381-9

12. de Almeida JR, Guyatt GH, Sud S, et al. Management of Bell palsy: clinical practice guideline. *Can Med Assoc J*. 2014;186(12):917–922. doi:10.1503/cmaj.131801
13. Portelinha J, Passarinho MP, Costa JM. Neuro-ophthalmological approach to facial nerve palsy. *Saudi J Ophthalmol*. 2015;29(1):39–47. doi:10.1016/j.sjopt.2014.09.009
14. Gronseth GS, Paduga R. Evidence-based guideline update: steroids and antivirals for bell palsy. report of the guideline development subcommittee of the American academy of neurology. *Neurology*. 2012;79(22):2209–2213. doi:10.1212/WNL.0b013e318275978c
15. Schwartz SR, Jones SL, Getchius TSD, Gronseth GS. Reconciling the clinical practice guidelines on Bell's palsy from the AAO-HNSF and the AAN. *Comment Otolaryngol*. 2014;150(5):709–711. doi:10.1177/0194599814529079
16. Grogan PM, Gronseth GS. Practice parameter: steroids, Acyclovir, and surgery for Bell's palsy (an evidence-based review): report of the quality standards subcommittee of the American academy of neurology. *Neurology*. 2001;56(7):830–836. doi:10.1212/wnl.56.7.830
17. Rowlands S, Hooper R, Hughes R, Burney P. The epidemiology and treatment of Bell's palsy in the UK. *Eur J Neurol*. 2002;9(1):63–67. doi:10.1046/j.1468-1331.2002.00343.x
18. Murthy JM, Saxena AB. Bell's palsy: treatment guidelines. *Ann Indian Acad Neurol*. 2011;14(Suppl 1):S70–S2. doi:10.4103/0972-2327.83092
19. Attia ZI, Noseworthy PA, Lopez-Jimenez F, et al. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. *Lancet*. 2019;394(10201):861–867. doi:10.1016/S0140-6736(19)31721-0
20. Seto H, Oyama A, Kitora S, et al. Gradient boosting decision tree becomes more reliable than logistic regression in predicting probability for diabetes with big data. *Sci Rep*. 2022;12(1):15889. doi:10.1038/s41598-022-20149-z
21. Lin C-M, Lin Y-S. Utilizing a two-stage Taguchi method and artificial neural network for the precise forecasting of cardiovascular disease risk. *Bioengineering*. 2023;10(11):1286. doi:10.3390/bioengineering10111286
22. Lin C-M, Lin Y-S. TPTM-HANN-GA: a novel hyperparameter optimization framework integrating the Taguchi method, an artificial neural network, and a genetic algorithm for the precise prediction of cardiovascular disease risk. *Mathematics*. 2024;12(9):1303. doi:10.3390/math12091303
23. Dong B, Zhang H, Duan Y, Yao S, Chen Y, Zhang C. Development of a machine learning-based model to predict prognosis of alpha-fetoprotein-positive hepatocellular carcinoma. *J Transl Med*. 2024;22(1). doi:10.1186/s12967-024-05203-w
24. Peng ZH, Tian JH, Chen BH, et al. Development of machine learning prognostic models for overall survival of prostate cancer patients with lymph node-positive. *Sci Rep*. 2023;13(18424). doi:10.1038/s41598-023-45804-x
25. Doja MN, Kaur I, Ahmad T. Age-specific survival in prostate cancer using machine learning. *Data Technol Appl*. 2020;54(2):215–234. doi:10.1108/DTA-10-2019-0189
26. Gaber A, Taher MF, Wahed MA, Shalaby NM, Gaber S. Classification of facial paralysis based on machine learning techniques. *Biomed Eng Online*. 2022;21(1). doi:10.1186/s12938-022-01139-5
27. Umirzakova S, Ahmad S, Mardieva S, Muksimova S, Whangbo TK. Deep learning-driven diagnosis: a multi-task approach for segmenting stroke and Bell's palsy. *Pattern Recogn*. 2023;144:109866. doi:10.1016/j.patcog.2023.109866
28. Kishimoto-Urata M, Urata S, Nishijima H, et al. Predicting synkinesis caused by Bell's palsy or Ramsay hunt syndrome using machine learning-based logistic regression. *Laryngoscope Investigative Otolaryngol*. 2023;8(5):1189–1195. doi:10.1002/lio2.1145
29. Riley RD. Machine learning in prognosis research. In: editor, Riley RD. *Prognosis Research in Healthcare: Concepts, Methods, and Impact*. Oxford; 2019. doi:10.1093/med/9780198796619.003.0017
30. Tedeschi R, Donati D, Giorgi F. Beyond the smile: a systematic review of diagnostic tools for peripheral facial paralysis. *Acta Neurologica Belgica*. 2024. doi:10.1007/s13760-024-02630-w
31. Myrick D. *R | Early Treatment in Bell's Palsy*. Kaggle; 2021. Available from: https://www.kaggle.com/datasets/dillonmyrick/bells-palsy-clinical-trial. Accessed November 06, 2024.
32. Sullivan FM, Swan IRC, Donnan PT, et al. Early treatment with prednisolone or acyclovir in Bell's palsy. *N Engl J Med*. 2007;357(16):1598–1607. doi:10.1056/NEJMoa072006
33. Dunford R, Su Q, Tamang E (2021). The Pareto Principle. The Race. Available from: https://api.semanticscholar.org/CorpusID:15925174. Accessed November 06, 2024.
34. Hayaty M, Muthmainah S, Ghufran SM. Random and synthetic over-sampling approach to resolve data imbalance in classification. *Int J Artificial Intelligence Res*. 2020;4(2). doi:10.29099/ijaair.v4i2.152
35. Kee E, Chong JJ, Choong ZJ, Lau M. A comparative analysis of cross-validation techniques for a smart and lean pick-and-place solution with deep learning. *Electronics*. 2023;12(11):2371. doi:10.3390/electronics12112371
36. Sadr H, Salari A, Ashoobi MT, et al. Cardiovascular disease diagnosis: a holistic approach using the integration of machine learning and deep learning models. *Eur J Med Res*. 2024;29(1):455. doi:10.1186/s40001-024-02044-7
37. Nettleton D. Chapter 9 - Data Modeling. In: Nettleton D editor. *Commercial Data Mining*. Morgan Kaufmann; 2014:137–157. doi:10.1016/B978-0-12-416602-8.00009-1
38. Moslemi A. A tutorial-based survey on feature selection: recent advancements on feature selection. *Eng Appl Artif Intell*. 2023;126(Part D):107136. doi:10.1016/j.engappai.2023.107136
39. Lau KW, Wu QH. Online training of support vector classifier. *Pattern Recogn*. 2003;36(8):1913–1920. doi:10.1016/S0031-3203(03)00038-4
40. Alam S, Sonbhadra SK, Agarwal S, Nagabhushan P. One-class support vector classifiers: a survey. *Knowledge-Based Syst*. 2020;196:105754. doi:10.1016/j.knosys.2020.105754
41. Cervantes J, Garcia-Lamont F, Rodríguez-Mazahua L, Lopez A. A comprehensive survey on support vector machine classification: applications, challenges and trends. *Neurocomputing*. 2020;408:189–215. doi:10.1016/j.neucom.2019.10.118
42. Breiman L. Random Forests. *Machine Learning*. 2001;45(1):5–32. doi:10.1023/A:1010933404324
43. Popescu M-C, Balas VE, Perescu-Popescu L, Mastorakis N. Multilayer perceptron and neural networks. *WSEAS Transactions Circuits Sys*. 2009;8(7):579–588. doi:10.5555/1639537.1639542
44. Bradley AP. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recogn*. 1997;30(7):1145–1159. doi:10.1016/S0031-3203(96)00142-2
45. Düntsch I, Gediga G. Confusion matrices and rough set data analysis. *J Phys Conf Ser*. 2019. doi:10.48550/arXiv.1902.01487