The Predictive Power of Computerized Posturography With and Without Clinical Features for Accurate Vestibular Migraine Diagnosis

Anita Singh[1], Peter James, MD[2], Matthew Crowson, MD[3], Biratal Wagle, MS[4], Michaela Geffert, MS[2], James Saunders, MD[1,2]

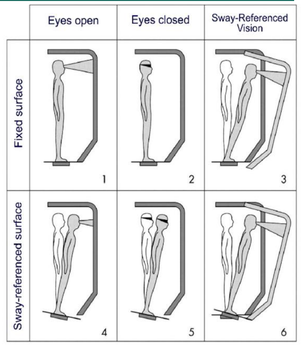
[1]Geisel School of Medicine at Dartmouth, [2]Dartmouth Hitchcock Medical Center, [3]Harvard Medical School Department of Otolaryngology, [4]Thayer School of Engineering

INTRODUCTION

Vestibular migraine (VM) affects 1–3% of the general population and up to 21% of migraine patients, yet remains frequently underdiagnosed due the absence of reliable biomarkers.[1-4]

Computerized dynamic posturography (CDP) objectively assesses postural control by quantifying vestibular, visual, and somatosensory integration under dynamic conditions (see figure for typical CDP test protocol), revealing subtle balance deficits even during symptom-free periods. [5-6]

We explored the predictive ability of CDP testing to enhance VM diagnostic accuracy with and without clinical features.



METHODS

Retrospective Cohort Analysis

- 334 patients (46 VM, 288 non-VM) | VM prevalence: 13.8%
- 6,121 CDP trials from DHMC
- Inclusion: <50% missing clinical data

CDP Biomechanical Features (15)

- Aggregated per trial using:
  - Center of pressure, sway angle, forces, torques, velocities

Clinical Features (9)

- Episode characteristics
- Migraine history
- Speech discrimination scores

Model Development

XGBoost
- scale_pos_weight = 10
- max_depth = 6
- n_estimators = 500

Random Forest
- Class weighting = 100:1
- n_estimators = 100
- Addresses class imbalance

5-Fold Stratified Cross-Validation

- Patient-level grouping: all trials from same patient in same fold (prevents data leakage)
- Stratification maintains VM/non-VM ratio across folds
- Standardization applied per fold using training data statistics
- Patient-level predictions: mean pooling of trial-level probabilities

ROC-Based Threshold Optimization

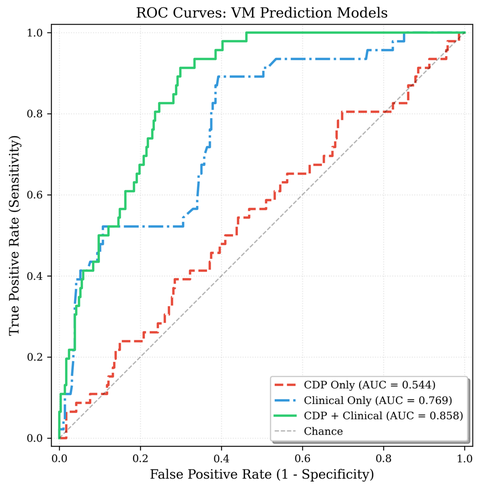
- Target: ≥ 80% sensitivity (per clinical framework)
- Evaluated trade-offs: sensitivity, specificity, PPV, NPV across multiple operating points
- Decision threshold adjusted per fold for optimal clinical utility

CDP Only (15 features) | **Clinical Only (9 features)** | **CDP + Clinical**

Performance Metrics

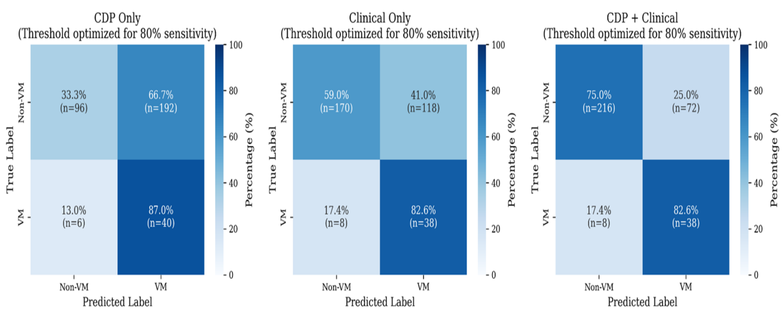
- Metrics: Mean ± 95% CI via bootstrap resampling (1,000 iterations)
- Feature importance: SHAP (SHapley Additive exPlanations) values

RESULTS



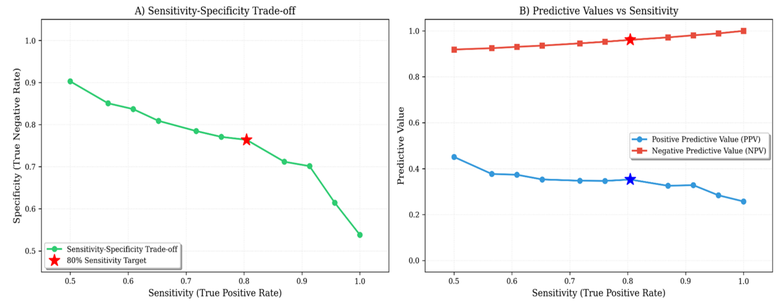
Model Comparison: CDP vs Clinical Features

ROC curve analysis demonstrated that clinical features alone substantially outperformed CDP-only models (AUC 0.80 vs 0.53). **The combined CDP + Clinical model achieved the highest discrimination** (AUC 0.86-0.94), with clinical features dominating predictive power. SHAP analysis confirmed this hierarchy: the top 3 features were clinical (photophobia, phonophobia, episode duration), with the highest-ranked CDP feature (vertical force, FZ_mean) appearing at position #2 with SHAP importance 1.79.



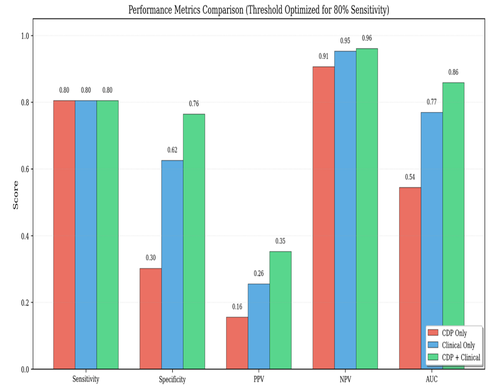
Confusion Matrix Analysis

At the 80% sensitivity threshold, the combined model correctly classified 7.6/9.2 VM patients per fold on average, with specificity of 75-76%. The confusion matrices reveal the trade-off inherent in high-sensitivity operation: achieving 82.8% VM detection required accepting 14.4 false positives per fold (24% false alarm rate among non-VM patients). Positive predictive value was 36.2% (95% CI: 26.3-46.9%), indicating that approximately 2 of every 3 positive predictions were false alarms. However, negative predictive value remained excellent at 96.4% (95% CI: 95.6-97.3%), providing confidence in ruling out VM.



Sensitivity-Specificity Trade-off Curves

Threshold optimization across multiple operating points revealed a steep sensitivity-specificity trade-off. The 80% sensitivity target (marked with red star) balanced clinical utility with acceptable false positive burden.



Performance Metrics Comparison

Direct comparison of all three model configurations shows that CDP features alone provided minimal discriminative power (sensitivity <10%, near-chance AUC). Clinical features achieved moderate performance (sensitivity 50%, AUC 0.80). The combined model's performance ceiling (sensitivity 82.8%, AUC 0.86) was driven primarily by clinical features, with CDP providing incremental value of +0.06 AUC points.

DISCUSSION

Key Features of Our Model

- High sensitivity achieved: 82.8% VM detection (95% CI: 80.7-85.2%) with stable performance across all folds
- Excellent negative prediction: 96% NPV enables confident rule-out of VM, reducing resource and time-intensive diagnostic workups
- Validated clinical criteria: Top predictors (photophobia, phonophobia, episode duration) confirm Bárány diagnostic framework for VM
- Quantified feature contributions: SHAP analysis precisely measures CDP's incremental value (+0.06 AUC) in multimodal assessment

Clinical Utility

- Triage tool potential: High NPV (96%) identifies low-probability cases who can avoid extensive vestibular testing
- Decision support: Quantitative risk stratification offers data-driven complement to clinical judgment
- Potential cost reduction: Efficiently rule out VM, reserving comprehensive workups for higher-risk patients
- Multimodal framework: Establishes baseline for integrating CDP with clinical, imaging, and laboratory data

CONCLUSION

This study demonstrates that machine learning models combining CDP biomechanical features with clinical criteria achieve 82.8% sensitivity for VM diagnosis (95% CI: 80.7-85.2%). This proof-of-concept creates a foundation for objective VM risk stratification and opens pathways for enhanced diagnostic accuracy through longitudinal CDP measurements, expanded clinical variables, and prospective multicenter validation. The models' ability to rule out VM represents an important advance toward more efficient, data-driven vestibular diagnostics.

REFERENCES

1. Yollu U, Uluduz DU, Yilmaz M, et al. Vestibular migraine screening in a migraine-diagnosed patient population, and assessment of vestibulocochlear function. Clin Otolaryngol. 2017;42(2):225-233. doi:10.1111/coa.1269.
2. Lempert T, Neuhauser H. Epidemiology of vertigo, migraine and vestibular migraine. J Neurol. 2009;256(3):333-338. doi:10.1007/s00415-009-0149-22.
3. Furman JM, Balaban CD. Vestibular migraine. Ann N Y Acad Sci. 2015;1343:90-96. doi:10.1111/nyas.12645
4. Huang TC, Wang SJ, Kheradmand A. Vestibular migraine: An update on current understanding and future directions. Cephalalgia. 2020;40(1):107-121. doi:10.1177/0333102419893176.
5. Keshner EA, Mallinson AI, Longridge NS, Sinno S, Petersen H, Perrin P. Evolution of postural control assessment: From dynamic posturography to virtual reality. Front Neurol. 2023;13. doi:10.3389/fneur.2022.10543463.
6. Black FO. What can posturography tell us about vestibular function? Ann N Y Acad Sci. 2001;942:446-464. doi:10.1111/j.1749-6632.2001.tb03765.x