Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Screening for Parkinson's Disease: Clinical and Community Evaluation of the CheckPD App Questionnaire

Jirada Sringean
King Chulalongkorn Memorial Hospital, Chulalongkorn University
Teerawat Laosombat
King Chulalongkorn Memorial Hospital, Chulalongkorn University
Tittaya Prasertpan
King Chulalongkorn Memorial Hospital, Chulalongkorn University
Saisamon Phumphid
King Chulalongkorn Memorial Hospital, Chulalongkorn University
Pattamon Panyakeaw
King Chulalongkorn Memorial Hospital, Chulalongkorn University
Onanong Phokaewvarangkul
King Chulalongkorn Memorial Hospital, Chulalongkorn University
Suppata Maytharakcheep
King Chulalongkorn Memorial Hospital, Chulalongkorn University
Chanawat Anan
King Chulalongkorn Memorial Hospital, Chulalongkorn University
Nattaphat Leabthong
King Chulalongkorn Memorial Hospital, Chulalongkorn University
Nurhussneeda Jeh-voh
King Chulalongkorn Memorial Hospital, Chulalongkorn University
Kornake Horaruengdecha
King Chulalongkorn Memorial Hospital, Chulalongkorn University
Jeeranun Meesri
King Chulalongkorn Memorial Hospital, Chulalongkorn University
Sunaree Wekhinhiran
King Chulalongkorn Memorial Hospital, Chulalongkorn University
Ornanong Udomsirithamrong
King Chulalongkorn Memorial Hospital, Chulalongkorn University
Appasone Phoumindr
King Chulalongkorn Memorial Hospital, Chulalongkorn University
Warongporn Phuenpathom

King Chulalongkorn Memorial Hospital, Chulalongkorn University

Roongroj Bhidayasiri

rbh@chulapd.org

King Chulalongkorn Memorial Hospital, Chulalongkorn University

---

Research Article

Keywords: Parkinson's disease, screening questionnaire, community screening, digital biomarkers, machine learning

Posted Date: November 10th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11486713/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Screening for Parkinson's Disease: Clinical and Community Evaluation of the CheckPD App Questionnaire

Jirada Sringean[1], Teerawat Laosombut[1], Tittaya Prasertpan[1,2], Saisamon Phumphid[1], Pattamon Panyakeaw[1], Onanong Phokaewvarangkul[1], Suppata Maytharakcheep[1], Chanawat Anan[1], Nattaphat Leabthong[1], Nurhussneeda Jeh-voh[1], Kornake Horaruengdecha[1], Jeeranun Meesri[1], Sunaree Wekhinhiran[1], Ornanong Udomsirithamrong[1,3], Appasone Phoumindr[1,4], Warongporn Phuenpathom[1], Roongroj Bhidayasiri[1,5*]

[1]Chulalongkorn Centre of Excellence for Parkinson's Disease and Related Disorders, King Chulalongkorn Memorial Hospital, The Thai Red Cross Society, Chulalongkorn University, Medicine, Bangkok, Thailand.

[2]Sawanpracharak Hospital, Nakornsawan, Thailand.

[3]Bangkok Hospital Chiang Mai, Chiang Mai, Thailand.

[4]Department of Internal Medicine, Faculty of Medicine, University of Health Sciences, Vientiane Capital, Lao PDR.

[5]The Academy of Science, The Royal Society of Thailand, Bangkok, Thailand.

Corresponding author:

Roongroj Bhidayasiri, MD, FRCP, FRCPI.

Chulalongkorn Centre of Excellence for Parkinson's Disease & Related Disorders,

Chulalongkorn University Hospital,

1873 Rama 4 Road, Bangkok 10330, Thailand.

Tel: (+662) 256-4000 ext. 70701; Fax: (+662) 256-4630

Email address: rbh@chulapd.org

Keywords: Parkinson's disease; screening questionnaire; community screening; digital biomarkers; machine learning

Word count: 3,601; abstract: 240 words; title: 90 characters

Abstract

Background: The CheckPD app combines a 20-item symptom questionnaire with smartphone sensor and voice assessments. We evaluated the questionnaire as a first-stage triage component and explored the contribution of digital features.

Methods: A hospital case-control evaluation included 60 participants with Parkinson's disease (PD) and 107 volunteer controls. Community screening included 1,510 participants; those meeting questionnaire or app referral criteria underwent specialist assessment. In a subgroup of 78 clinically evaluated participants, three prespecified model families were compared using questionnaire-only (Q) and questionnaire-plus-sensor (Q+S) features with matched participant-level five-fold cross-validation.

Results: In the hospital cohort, the questionnaire showed strong apparent discrimination (area under the receiver operating characteristic curve [ROC-AUC] 0.971, 95% confidence interval [CI] 0.951–0.992). A score of at least 5 yielded 90.00% sensitivity and 93.46% specificity. Each additional positive response was associated with higher odds of PD status after adjustment for age and sex (adjusted odds ratio 1.853, 95% CI 1.464–2.344). In community screening, 43 of 98 questionnaire-positive participants had confirmed PD (positive predictive value 43.88%; detection yield 2.85%). The app-based pathway identified four additional cases below the questionnaire threshold, giving 47 cases overall (3.11%). Mean ROC-AUC was higher with Q+S than with Q across all three models.

Conclusions: The questionnaire showed strong apparent hospital discrimination but more modest community performance. It should be used as a first-stage triage component within, rather than as a replacement for, the multimodal CheckPD pathway. External validation with blinded assessment and verification of screen-negative participants is required.

Introduction

Parkinson's disease (PD) is diagnosed clinically, but recognition may be delayed when symptoms are subtle or access to neurological expertise is limited. Diagnostic accuracy varies with clinical expertise and the stage at which patients are evaluated, highlighting the importance of specialist assessment and appropriate longitudinal evaluation (Rizzo et al. 2016). Scalable case-finding approaches could therefore help identify individuals who warrant further assessment and prioritise referral, particularly in settings where access to movement disorder specialists is limited.

Several questionnaires have been developed to screen for parkinsonism or PD, with performance varying according to item content, study population and method of administration (Mutch et al. 1991; Duarte et al. 1995; Rocca et al. 1998; Chan et al. 2000; Setthawatcharawanich et al. 2011; Dahodwala et al. 2012; Fereshtehnejad et al. 2014). Most have predominantly emphasised motor manifestations or parkinsonism, whereas non-motor features such as hyposmia, rapid eye movement (REM) sleep behaviour disorder (RBD), excessive daytime sleepiness, constipation and depression may occur before or alongside recognised motor manifestations (Heinzel et al. 2019). However, many of these symptoms are non-specific and may also occur with ageing or other common disorders. A symptom questionnaire may therefore be most useful as an accessible first-stage triage tool rather than as a stand-alone diagnostic instrument.

Digital technologies offer the opportunity to complement self-reported symptoms with objective measures of motor and non-motor function, although their clinical value depends on how they are integrated into a defined care pathway rather than on technical performance alone (Falkenburger et al. 2026). Smartphone-derived speech and other digital features have also shown potential for scalable PD classification across different populations and languages (Lim et al. 2025). The CheckPD programme was developed as a multimodal screening pathway

combining a 20-item symptom questionnaire with smartphone-based assessments of voice, tremor, dexterity, gait and balance (Bhidayasiri et al. 2024). Its subsequent national implementation demonstrated the feasibility, reach, adoption and real-world performance of the integrated screening programme (Bhidayasiri et al. 2026).

These previous CheckPD publications addressed development of the broader platform and implementation of the integrated programme rather than the specific operating characteristics of the questionnaire within the staged screening pathway. The present study therefore evaluated the questionnaire as a first-stage symptom-triage component by (1) determining item-level and total-score operating characteristics and deriving a screening threshold in a hospital case-control cohort; (2) examining the association between questionnaire score and PD status after adjustment for age and sex; (3) estimating questionnaire-positive predictive value and confirmed PD detection yield during community deployment; and (4) exploring whether addition of sensor and voice features provided complementary discriminatory information in a clinically evaluated community subgroup.

Methods

Study design and relationship to previous CheckPD studies

This study comprised three complementary analyses of the 20-item CheckPD questionnaire: a hospital-based case-control evaluation of item-level and total-score operating characteristics, a community field evaluation of questionnaire-positive predictive value and confirmed PD detection yield, and an exploratory comparison of questionnaire-only and questionnaire-plus-sensor classification approaches. Development and preliminary evaluation of the broader CheckPD platform, including the questionnaire component, and subsequent national implementation of the integrated screening pathway have been reported previously (Bhidayasiri et al. 2024). The present study was designed to address distinct questionnaire-specific questions

and should therefore be interpreted as complementary to, rather than a re-evaluation of, the previously reported CheckPD programme. The study was approved by the Ethics Committee of the Faculty of Medicine of Chulalongkorn University (IRB number 0804/65) and registered with the Thai Clinical Trials Registry with the TCTR identification number TCTR20231025005. Development of the 20-item questionnaire and the CheckPD workflow are shown in Figure 1. The Thai version of the 20-item questionnaire and instructions for its use are provided in Supplementary Data 1 and 2. The questionnaire diagnostic accuracy evaluation is reported with reference to the Standards for Reporting Diagnostic Accuracy Studies (STARD 2015) recommendations (Bossuyt et al. 2015). Reporting of the exploratory machine-learning analysis was informed by applicable TRIPOD+AI recommendations for clinical prediction model studies (Collins et al. 2024). The corresponding reporting checklists accompany this submission.

Hospital-based case-control validation cohort

Between January and July 2024, consecutive patients with clinically diagnosed PD were recruited from the Chulalongkorn Centre of Excellence for Parkinson's Disease and Related Disorders (ChulaPD, www.chulapd.org), King Chulalongkorn Memorial Hospital. PD had been diagnosed by movement disorder specialists before questionnaire administration according to the UK Parkinson's Disease Society Brain Bank criteria (Gibb and Lees 1988) or the Movement Disorder Society clinical diagnostic criteria for PD (Postuma et al. 2015). Participants with PD were therefore aware of their diagnosis when completing the questionnaire.

Volunteer comparison participants aged ≥18 years were recruited from King Chulalongkorn Memorial Hospital and the Centre's outpatient clinics. Eligible volunteers had no history or current clinical evidence of PD, atypical or secondary parkinsonian disorder, or another major neurological disorder; were able to read, comprehend and independently complete the questionnaire; and provided written informed consent. Exclusion criteria included essential tremor, dystonia, stroke, epilepsy, dementia, peripheral neuropathy, active psychiatric illness

affecting response reliability, and advanced musculoskeletal, orthopaedic, visual or auditory impairment.

Community field screening with the questionnaire and CheckPD application

Community screening was conducted on 28–29 May 2026 in Prachuap Khiri Khan Province, Thailand, through active community-based health screening campaigns organised in collaboration with the provincial Thai Red Cross, local health authorities and Village Health Volunteers (VHVs). Recruitment was invitation-based with voluntary participation following community outreach announcements. Although outreach primarily targeted adults aged ≥40 years, eligibility included community-dwelling adults aged ≥18 years who could complete the 20-item CheckPD questionnaire independently or with technical assistance from trained research staff or VHVs and who provided written informed consent. Participants with severe acute illness or bedridden status and those with severe cognitive or communication impairment precluding reliable symptom reporting were excluded.

Participants completed the questionnaire and smartphone-based assessments within the CheckPD application. Participants meeting either referral criterion (a questionnaire score ≥5 or a high-risk CheckPD application result) underwent neurological assessment by movement disorder specialists during the same visit. PD was diagnosed according to the UK Parkinson's Disease Society Brain Bank criteria (Gibb and Lees 1988) or the Movement Disorder Society clinical diagnostic criteria for PD (Postuma et al. 2015). The specialist assessment therefore followed screening, and the specialists were aware of the questionnaire and application results. Participants in whom PD was confirmed were classified as confirmed PD; those assessed who did not meet the clinical diagnostic criteria for PD were classified as having no confirmed PD.

The deployed CheckPD high-risk classifier used a prespecified Random Forest ensemble majority-voting architecture (CheckPD ML engine v1.0; backend server build

testmodel20241128x). The classifier incorporated the 20-item questionnaire and up to eight smartphone-based digital assessments: resting tremor, postural tremor, gait, balance, sustained vowel phonation, sentence repetition, alternating-finger tapping and pinch-to-size. Predictions from the component classifiers were combined using majority voting with a threshold of 0.5 for high-risk classification. Model parameters, feature-extraction procedures and decision thresholds were trained and frozen before community deployment and were not modified using outcomes from the present field cohort.

Because reference-standard clinical assessment was restricted to participants meeting the screening referral criteria, the community analysis was limited to positive predictive value (PPV) and confirmed PD detection yield. Population sensitivity, specificity, negative predictive value (NPV) and prevalence could not be estimated because screen-negative participants were not clinically verified.

Missing data

Responses to all 20 questionnaire items were mandatory within the CheckPD application; therefore, submitted app-based questionnaires contained no missing questionnaire items. Paper-based hospital questionnaires with missing responses were excluded from analysis. For the exploratory machine-learning analysis, participants lacking one or more required smartphone task recordings or with unusable recordings were excluded. Among participants retained in the machine-learning dataset, isolated missing extracted feature values were median-imputed using values calculated exclusively from the training fold and subsequently applied to the corresponding test fold. Feature standardisation was likewise fitted using training data only.

Statistical analysis

Continuous variables are reported as mean ± standard deviation (SD), and categorical variables as number and percentage. Between-group comparisons were performed using independent-

samples t tests and chi-square or Fisher's exact tests, as appropriate. Apparent discrimination of the total questionnaire score in the hospital cohort was evaluated using receiver operating characteristic (ROC) analysis. The operating threshold was selected in the same cohort by maximising the Youden index and should therefore be considered internally derived rather than externally validated. Sensitivity, specificity, PPV, NPV, false-positive rate and false-negative rate were reported with 95% confidence intervals (CIs).

Each questionnaire item was evaluated in a separate binary logistic regression model, with PD status as the dependent variable and the individual questionnaire item, age and sex as independent variables. The total questionnaire score was evaluated separately in a hierarchical logistic-regression model containing age and sex in the first block and total questionnaire score in the second block. Odds ratios (ORs) with 95% CIs were reported. Community analyses were restricted to PPV and confirmed PD detection yield because reference-standard assessment was not performed in screen-negative participants. A two-sided p value <0.05 was considered statistically significant. Statistical analyses were performed using SPSS version 29.0.1.0.

Exploratory machine-learning analysis

The exploratory analysis included 78 clinically evaluated community participants with usable questionnaire and smartphone data, comprising 20 participants with confirmed PD and 58 participants without confirmed PD. The questionnaire-only feature set (Q) consisted of the 20 binary questionnaire items. The questionnaire-plus-sensor feature set (Q+S) comprised these 20 items plus 84 sensor- and voice-derived features obtained from resting tremor, postural tremor, gait, balance, sustained phonation and sentence-speech tasks (Bhidayasiri et al. 2024).

Three model families were prespecified with fixed hyperparameters: L2-penalised logistic regression (C = 1.0), Gaussian naïve Bayes (var_smoothing=1×10^{-9}) (John and Langley 1995), and Random Forest (300 trees, maximum depth 3, minimum of 3 samples per leaf and square-

root feature sampling) (Breiman 2001). No class weighting, synthetic resampling, hyperparameter tuning or model-family selection was performed. The classification threshold was fixed at 0.5. Median imputation and feature standardisation, when required, were fitted exclusively using training-fold data.

All three models were evaluated separately using Q and Q+S with identical participant-level five-fold grouped cross-validation splits. No participant contributed data to both the training and test sets within the same fold. Performance was summarised descriptively as mean ± SD across the five matched test folds. No model was designated as the preferred model, and no formal inferential comparison between Q and Q+S was performed because of the small exploratory sample (Varoquaux 2018). Further details are provided in Supplementary Data 3.

Results

A hospital-based case-control cohort of clinically diagnosed patients with PD and volunteer controls was used to assess the questionnaire's discriminatory performance. We recruited 60 patients with PD (36 men and 24 women) and 107 volunteer controls (22 men and 85 women). The mean age was 68.08 $\pm$ 9.59 years for patients with PD versus 54.84 $\pm$ 11.84 years for controls ($p < 0.001$). The average educational level of both patients with PD and volunteer controls was at high-school level.

The mean total score was 11.70 ± 5.08 for patients with PD versus 1.12 ± 2.09 for controls ($p < 0.001$). There were significant differences in "yes" or "positive" responses between patients with clinically diagnosed PD and controls on all items, as shown in Table 1 and Figure 2. The most frequently reported symptom was item no. 3 (clumsy or slower in performing daily activities; 86.7%), while the least frequently reported was item no. 16 (hallucination; 31.7%).

In the hospital-based case-control cohort, ROC curve analysis of the total questionnaire score demonstrated strong apparent discrimination between patients with clinically diagnosed PD and volunteer controls, with an area under the curve (AUC) of 0.971 (95% CI: 0.951–0.992). A cut-off score of 5 or higher was selected using the highest Youden index (0.835) (Figure 3). Performance measures for all cut-off scores are shown in Supplementary Data 4. Sensitivity and specificity were 90.00% (95% CI: 79.49–96.24%) and 93.46% (95% CI: 87.01–97.30%), respectively. The PPV was 88.52% (95% CI: 77.82–95.26%) and the NPV was 94.34% (95% CI: 88.04–97.87%). The false-positive rate (FPR) was 6.54% (95% CI: 2.68–12.99%) and the false-negative rate (FNR) was 10.00% (95% CI: 3.76–20.51%). All 20 questionnaire items showed significantly different response frequencies between patients with PD and volunteer controls (all $p < 0.001$). The highest item-level sensitivity and specificity for discriminating between patients with PD and volunteer controls were 86.67% for item no. 3 (clumsy or slower in performing daily activities) and 99.07% for item no. 13 (hyposmia), respectively. The sensitivities and specificities of the other items are shown in Table 1.

Following adjustment for age and sex, several motor items showed strong associations with PD status. Item no. 2 (change in handwriting) yielded the highest odds ratio (adjusted OR = 125.89, 95% CI: 28.36–558.85, $p < 0.001$), followed by item no. 10 (unilateral onset of motor symptoms) (adjusted OR = 117.89, 95% CI: 25.59–543.04, $p < 0.001$). Among non-motor features, decreased sense of smell (item 13) also showed a strong age- and sex-adjusted association with PD status (adjusted OR 55.69, 95% CI 5.85–529.71). The other adjusted ORs are shown in Table 1.

In the hierarchical logistic-regression analysis, Nagelkerke pseudo-R^2 was 0.432 for the model containing age and sex and increased to 0.863 after addition of the total questionnaire score. Each additional positive questionnaire response was associated with an 85.3% increase in the odds of PD status after adjustment for age and sex (adjusted OR 1.853, 95% CI 1.464–2.344; $p <$

0.001). Age remained associated with PD status (adjusted OR 1.140, 95% CI 1.055–1.232; p < 0.001), whereas sex was not statistically significant after adjustment for total questionnaire score (p = 0.221).

Among 1,510 community participants, 107 were classified as high risk by the CheckPD application and underwent specialist assessment. Of 98 participants with questionnaire scores ≥5, 43 had confirmed PD and 55 did not, corresponding to a questionnaire-positive PPV of 43.88% (95% CI 33.86–54.30) and a questionnaire-detected confirmed PD yield of 2.85% (43/1,510). Nine app-high-risk participants had questionnaire scores <5; four had confirmed PD and five did not. The app-based pathway therefore identified 47 confirmed PD cases overall, including four cases below the questionnaire threshold. These four cases represented 8.5% of all confirmed cases and an absolute additional detection yield of 0.26% of the screened cohort. Overall confirmed PD detection yield was 3.11% (47/1,510), and the PPV among app-high-risk participants was 43.93% (47/107). The remaining 1,403 app-low-risk participants did not undergo reference-standard assessment; therefore, population sensitivity, specificity, NPV and prevalence could not be estimated.

Among the 107 clinically evaluated community participants, analysable multimodal data were available for 78, comprising 20 of 47 participants with confirmed PD and 58 of 60 participants without confirmed PD. Twenty-nine participants (27 with confirmed PD and two without confirmed PD) were excluded from the exploratory machine-learning analysis because the required sensor recordings were incomplete or unusable.

Using identical participant-level five-fold grouped cross-validation splits, mean ROC-AUC was higher with questionnaire-plus-sensor data (Q+S) than with questionnaire-only data (Q) for all three prespecified model families: L2-penalised logistic regression (0.683 ± 0.074 vs 0.431 ± 0.211), Gaussian naïve Bayes (0.810 ± 0.086 vs 0.577 ± 0.233), and Random Forest (0.675 ±

0.121 vs 0.513 ± 0.227). Mean sensitivity was also higher with Q+S for all three models, whereas the direction of change in specificity differed by model family (Table 3). These comparisons were descriptive; no preferred model was selected and no formal inferential comparison was performed. The corresponding test-fold ROC curves are shown in Figure 4.

Discussion

This study evaluated the 20-item CheckPD questionnaire as the symptom-based entry component of a broader multimodal PD screening pathway. The questionnaire showed strong apparent discrimination in an enriched hospital case-control cohort, but a more modest questionnaire-positive PPV during community deployment. The app-based pathway identified four additional confirmed PD cases below the questionnaire threshold. In a small clinically evaluated subgroup, addition of sensor and voice features was associated with higher mean ROC-AUC across three prespecified model families.

The item-level findings also highlight the importance of capturing bradykinetic manifestations in addition to tremor. Clumsiness or slowness during daily activities (item 3) showed the highest sensitivity (86.7%), followed by change in handwriting (item 2; 83.3%), whereas rest tremor (item 1) had a sensitivity of 73.3%. Earlier screening studies identified tremor as an important discriminating symptom between participants with PD and comparison participants (Mutch et al. 1991; Setthawatcharawanich et al. 2011). Taken together, these findings support inclusion of task-specific questions addressing slowness and dexterity alongside rest tremor rather than relying predominantly on tremor-based screening.

Several non-motor items demonstrated high item-level specificity. Hyposmia (item 13) had a specificity of 99.07%, with positive responses reported by 35.0% of participants with PD and 0.9% of volunteer controls, while the REM sleep behaviour disorder item (item 11) had a

specificity of 98.13%. These features are recognised as being associated with the prodromal phase of PD (Heinzel et al. 2019). Chronic constipation and excessive daytime sleepiness also showed relatively high item-level specificity. However, these symptoms are not specific to PD in clinical practice, and the present findings should not be interpreted as demonstrating prodromal PD detection.

Unilateral onset of motor symptoms (item 10) also showed strong discrimination, with a specificity of 96.26% and an age- and sex-adjusted OR of 117.89. In contrast, dizziness on standing (item 17) was the most frequently reported positive item among volunteer controls (15.0%) and had the lowest specificity (85.05%) and adjusted OR (4.72), consistent with the non-specific nature of orthostatic symptoms. In the hierarchical analysis, Nagelkerke pseudo-R^2 increased from 0.432 for the model containing age and sex to 0.863 after addition of the total questionnaire score. Each additional positive response was associated with an 85.3% increase in the odds of PD status after adjustment for age and sex. These findings support further evaluation of the cumulative questionnaire score as a first-stage triage measure, while recognising that the ≥5 threshold was derived and evaluated within the same hospital cohort.

Community deployment provided a more pragmatic assessment of questionnaire performance. Among questionnaire-positive participants, 43 of 98 had specialist-confirmed PD, corresponding to a PPV of 43.88% and a confirmed PD detection yield of 2.85% of the screened cohort. The lower PPV than that observed in the enriched hospital case-control sample is expected given differences in disease frequency, case mix and symptom spectrum between clinical and community settings (Dahodwala et al. 2012; Usher-Smith et al. 2016). Self-reported symptoms may also reflect musculoskeletal disease, autonomic symptoms unrelated to PD, mood and sleep disorders, ageing and differences in health literacy. The questionnaire should therefore be interpreted as a broad symptom-triage tool rather than as a stand-alone diagnostic instrument.

Performance estimates from the present questionnaire analysis and previous CheckPD studies should not be interpreted as direct comparisons. The apparent AUC of 0.971 was obtained from the complete hospital case-control cohort in which the questionnaire threshold was also derived, whereas the previously reported CheckPD analyses used different cohorts, validation procedures, thresholds and outcome measures (Bhidayasiri et al. 2024; Bhidayasiri et al. 2026). AUC, accuracy and PPV describe different aspects of performance and are influenced by study design and case mix. The high apparent AUC in the hospital cohort therefore does not indicate superiority of questionnaire-only screening over the multimodal CheckPD pathway; rather, it supports the questionnaire's role as an accessible first-stage symptom-triage component.

Within the clinically evaluated high-risk subgroup, questionnaire positivity was almost identical between participants with and without confirmed PD, indicating substantial restriction of questionnaire-score range after initial screening selection. In this more challenging setting, addition of sensor and voice features was associated with higher mean ROC-AUC and sensitivity across all three prespecified model families, although changes in specificity differed between models. These findings suggest that objective digital measures may provide complementary information beyond self-reported symptoms. However, the small sample, high predictor-to-participant ratio, differential availability of analysable sensor recordings and fold-to-fold variability preclude claims of definitive incremental validity.

A practical strength of the questionnaire is its simplicity: it can be completed in approximately five minutes and produces an easily interpretable cumulative score, supporting its use as a scalable first-stage triage tool. The study also evaluated the questionnaire across complementary hospital and community settings and provided a transparent internally derived operating threshold with corresponding sensitivity and specificity. The exploratory multimodal analysis further illustrates how self-reported symptoms may be combined with objective digital measures,

while appropriately requiring larger prospective studies before conclusions about incremental clinical utility can be drawn.

Several limitations should be considered. First, the hospital analysis used an enriched case-control design comparing established PD with volunteer controls who were younger and differed in sex distribution. Participants with PD completed the questionnaire after diagnosis, and comparison participants with several neurological, musculoskeletal and sensory disorders were excluded; these features may have increased apparent discrimination. The operating threshold was derived and evaluated in the same cohort. Second, community recruitment occurred during a two-day screening campaign and may not represent the broader target population. Only participants meeting questionnaire or app referral criteria underwent specialist assessment, precluding estimation of population sensitivity, specificity, NPV and prevalence. Specialists were aware of the screening results, introducing potential review bias. Third, the exploratory machine-learning analysis included only 20 participants with confirmed PD and 58 without confirmed PD and used 104 candidate features. Twenty-seven of 47 confirmed PD participants, compared with two of 60 participants without confirmed PD, lacked analysable multimodal data, creating potential selection bias. Finally, although the questionnaire includes symptoms associated with prodromal PD and atypical parkinsonism, longitudinal prodromal prediction and discrimination of atypical syndromes were not evaluated.

Conclusion

Building on previously reported questionnaire development and psychometric evaluation, the 20-item CheckPD questionnaire showed strong apparent discrimination between established PD and volunteer controls in an enriched hospital case-control cohort, but a more modest questionnaire-positive PPV during community deployment. The app-based pathway identified four additional confirmed PD cases below the questionnaire threshold, while exploratory participant-level cross-

validation showed higher mean ROC-AUC after addition of sensor and voice features across three prespecified model families. These findings support the questionnaire as a scalable first-stage symptom-triage component within, rather than a replacement for, the multimodal CheckPD pathway. Prospective external validation in representative populations, with blinded assessment and reference-standard evaluation of screen-negative participants, is required.

Table and figure legends

Table 1. Numbers of participants (patients with PD and volunteer controls) who answered "yes" or gave a "positive" response to each item of the PD screening questionnaire, with sensitivity, specificity and adjusted odds ratios (ORs).

Table 2. Characteristics of clinically evaluated participants with and without confirmed PD who had high-risk results from the CheckPD application.

Table 3. Descriptive five-fold grouped cross-validation performance using questionnaire data alone (Q) and questionnaire data combined with sensor and voice features (Q+S).

Figure 1. Development of the 20-item questionnaire and the CheckPD workflow.

Figure 2. Proportion of participants reporting each questionnaire item in the hospital PD and volunteer-control groups.

Figure 3. Receiver operating characteristic (ROC) curve for the total 20-item CheckPD questionnaire score in the hospital case-control cohort. AUC = 0.971 (95% CI 0.951–0.992). A score ≥5 was selected using the maximum Youden index.

Figure 4. Receiver operating characteristic (ROC) curves across three fixed machine-learning model families comparing questionnaire-only (Q) and multimodal questionnaire-plus-sensor (Q+S) feature sets.

Supplementary Data 1. Thai version of the 20-item CheckPD questionnaire.

Supplementary Data 2. Instructions for completing the self-reported 20-item CheckPD questionnaire.

Supplementary Data 3. Detailed machine-learning features and preprocessing.

Supplementary Data 4. Performance measures for all cut-off scores in the hospital-based case-control cohort.

Code and data availability

https://drive.google.com/drive/folders/1fUu5cmvo8XYJXiLZ8_XO1mvz2cVoIxFw?usp=share_link

Author contributions

JS conceived the study, led the initial conceptualisation, item design, and development of the 20-item screening questionnaire, conducted clinical and field evaluations, performed statistical analyses, and drafted the main manuscript. TL conducted the machine-learning pipeline, digital sensor feature extraction, and model cross-validation. TP co-led the initial conceptualisation, item design, and development of the 20-item screening questionnaire. RB supervised the overall project, obtained funding, guided the clinical methodology, critically revised the manuscript for important intellectual content, and served as corresponding author. SP, PP, OP, SM, CA, NL, NJ, KH, JM, SW, OU, AP, and WP contributed to community field coordination, clinical assessments

and database management, and critically revised the manuscript. All authors reviewed and approved the final manuscript and agree to be accountable for the work.

Funding: This research was funded by the following grants: the Thailand Science Research and Innovation Fund (Program Management Unit for Competitiveness, C01F670185), the National Economic and Social Development Council, Technology and Innovation in Life Sciences National Agency (TC (ERP) 31/2568), National Research Council of Thailand (N42A680591, N35E680087), Ratchadapiseksompotch Fund (RA-MF-22/69) and the Center of Excellence grant (CE68_028_3000_004), both awarded by Chulalongkorn University, and the Thai Red Cross Education and Research Committee Fund.

Declaration of competing interests

Jirada Sringean has received a grant from The Michael J. Fox Foundation for Parkinson's Research (MJFF-026889). Roongroj Bhidayasiri has received grants from the following agencies: the Thailand Science Research and Innovation Fund (Program Management Unit for Competitiveness, C01F670185), the National Economic and Social Development Council, Technology and Innovation in Life Sciences National Agency (TN 3/2569, TC (ERP) 31/2568), National Research Council of Thailand (N42A680591, N35E680087), Ratchadapiseksompotch Fund (RA-MF-22/69) and the Center of Excellence grant (CE68_028_3000_004), both awarded by Chulalongkorn University, and the Thai Red Cross Education and Research Committee Fund. He is also supported by the Royal Society of Thailand, has provided consultancy services, and has received honoraria and lecture fees from industries such as Britannia, Ipsen, Teva-Lundbeck, and Mitsubishi-Tanabe. Bhidayasiri holds patents for various innovations, including a laser-guided walking stick, a portable tremor device, technology for monitoring during the night, an electronic diary for Parkinson's disease symptoms, as well as the Parkinson insole and shoe, and anti-choking mugs. Furthermore, his copyrighted works comprise the application for screening Parkinson's disease (CheckPD^{TM}), a Parkinson's mascot, a daily activity diary for Parkinson's,

484 creative dopamine-themed lyrics, and educational video content focusing on common night-time
485 and gastrointestinal symptoms associated with Parkinson's disease.
486

References

Bhidayasiri R, Phumphid S, Sringean J, Anan C, Laosombut T, Meesri J, Wekhinhiran S, Leabthong N, Phokaewvarangkul O, Surangsrirat D, Vateekul P, Win ZK, Jeh-Voh N, Panyakaew P, Phuenpathom W, Maytharakcheep S, Jagota P, Ratanasirisawad V, Khongprasert S, Rattanachaisit W, Prasertpan T, Udomsirithamrong O, Phoumindr A, Likhitwitayawuid N, Saengmanee T, Rattanajun P, Horaruengdecha K, Sukmueng N, Dornbunlon P, Sittitrai W, Vorachetbancha S, Chanchalong W, Phonsrithong A, Kullavanijaya P, Sricholwattana J, Chirapravati V, Chariyavilaskul P, Sukpisan K, Thammatacharee J, Limpanyalers A, Juthasong J, Rattanamornsakul W, Bunnag T, Boonrach G (2026) National implementation of a digital Parkinson's disease screening programme in Thailand: reach, adoption, and real-world performance of the CheckPD app. BMC Public Health 26 (1). doi:10.1186/s12889-026-27654-0

Bhidayasiri R, Sringean J, Phumphid S, Anan C, Thanawattano C, Deoisres S, Panyakaew P, Phokaewvarangkul O, Maytharakcheep S, Buranasrikul V, Prasertpan T, Khontong R, Jagota P, Chaisongkram A, Jankate W, Meesri J, Chantadunga A, Rattanajun P, Sutaphan P, Jitpugdee W, Chokpatcharavate M, Avihingsanon Y, Sittipunt C, Sittitrai W, Boonrach G, Phonsrithong A, Suvanprakorn P, Vichitcholchai J, Bunnag T (2024) The rise of Parkinson's disease is a global challenge, but efforts to tackle this must begin at a national level: a protocol for national digital screening and "eat, move, sleep" lifestyle interventions to prevent or slow the rise of non-communicable diseases in Thailand. Front Neurol 15:1386608. doi:10.3389/fneur.2024.1386608

Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, Lijmer JG, Moher D, Rennie D, de Vet HC, Kressel HY, Rifai N, Golub RM, Altman DG, Hooft L, Korevaar

DA, Cohen JF, Group S (2015) STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ 351:h5527. doi:10.1136/bmj.h5527

Breiman L (2001) Random Forests. Machine Learning 45 (1):5–32. doi:10.1023/A:1010933404324

Chan DK, Hung WT, Wong A, Hu E, Beran RG (2000) Validating a screening questionnaire for parkinsonism in Australia. J Neurol Neurosurg Psychiatry 69 (1):117–120. doi:10.1136/jnnp.69.1.117

Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, Ghassemi M, Liu X, Reitsma JB, van Smeden M, Boulesteix AL, Camaradou JC, Celi LA, Denaxas S, Denniston AK, Glocker B, Golub RM, Harvey H, Heinze G, Hoffman MM, Kengne AP, Lam E, Lee N, Loder EW, Maier-Hein L, Mateen BA, McCradden MD, Oakden-Rayner L, Ordish J, Parnell R, Rose S, Singh K, Wynants L, Logullo P (2024) TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 385:e078378. doi:10.1136/bmj-2023-078378

Dahodwala N, Siderowf A, Baumgarten M, Abrams A, Karlawish J (2012) Screening questionnaires for parkinsonism: a systematic review. Parkinsonism Relat Disord 18 (3):216–224. doi:10.1016/j.parkreldis.2011.09.003

Duarte J, Claveria LE, de Pedro-Cuesta J, Sempere AP, Coria F, Calne DB (1995) Screening Parkinson's disease: a validated questionnaire of high specificity and sensitivity. Mov Disord 10 (5):643–649

Falkenburger B, Fanciulli A, Potter-Nerger M, van Eimeren T, Classen J, Hoglinger G, Redecker C, Winkler J, Klucken J (2026) Clinical utility of digital technologies in Parkinson's disease. J Neural Transm (Vienna) 133 (2):297–307. doi:10.1007/s00702-025-03062-3

Fereshtehnejad SM, Shafieesabet M, Rahmani A, Farhadi F, Hadizadeh H, Shahidi GA, Delbari A, Lökk J (2014) A novel 6-item screening questionnaire for Parkinsonism: validation and comparison between different instruments. Neuroepidemiology 43 (3-4):178–193. doi:10.1159/000368167

Gibb WR, Lees AJ (1988) The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease. J Neurol Neurosurg Psychiatry 51 (6):745–752

Heinzel S, Berg D, Gasser T, Chen H, Yao C, Postuma RB (2019) Update of the MDS research criteria for prodromal Parkinson's disease. Mov Disord 34 (10):1464–1470. doi:10.1002/mds.27802

John GH, Langley P Estimating Continuous Distributions in Bayesian Classifiers. In: Conference on Uncertainty in Artificial Intelligence, 1995.

Lim WS, Chiu SI, Peng PL, Jang JR, Lee SH, Lin CH, Kim HJ (2025) A cross-language speech model for detection of Parkinson's disease. J Neural Transm (Vienna) 132 (4):579–590. doi:10.1007/s00702-024-02874-z

Mutch WJ, Smith WC, Scott RF (1991) A screening and alerting questionnaire for parkinsonism. Neuroepidemiology 10 (3):150–156

Postuma RB, Berg D, Stern M, Poewe W, Olanow CW, Oertel W, Obeso J, Marek K, Litvan I, Lang AE, Halliday G, Goetz CG, Gasser T, Dubois B, Chan P, Bloem BR, Adler CH, Deuschl G (2015) MDS clinical diagnostic criteria for Parkinson's disease. Mov Disord 30 (12):1591–1601. doi:10.1002/mds.26424

Rizzo G, Copetti M, Arcuti S, Martino D, Fontana A, Logroscino G (2016) Accuracy of clinical diagnosis of Parkinson disease: A systematic review and meta-analysis. Neurology 86 (6):566–576. doi:10.1212/WNL.0000000000002350

Rocca WA, Maraganore DM, McDonnell SK, Schaid DJ (1998) Validation of a telephone questionnaire for Parkinson's disease. J Clin Epidemiol 51 (6):517–523. doi:10.1016/s0895-4356(98)00017-1

Setthawatcharawanich S, Sathirapanya P, Phabphal K, Limapichat K (2011) Short questionnaire for Parkinson's disease as a screening instrument. Clinical neurology and neurosurgery 113 (10):885–888. doi:10.1016/j.clineuro.2011.07.001

Usher-Smith JA, Sharp SJ, Griffin SJ (2016) The spectrum effect in tests for risk prediction, screening, and diagnosis. Bmj 353:i3139. doi:10.1136/bmj.i3139

Varoquaux G (2018) Cross-validation failure: Small sample sizes lead to large error bars. Neuroimage 180 (Pt A):68–77. doi:10.1016/j.neuroimage.2017.06.061

20-item questionnaire development

Content validation and reliability

Hospital-based case-control validation cohort (N=167)
60 Clinically diagnosed PD
107 Volunteer controls

Community questionnaire field screening

Community screening cohort (N=1,510)

CHECK PD



App Low-risk (N=1,403)
Not clinically verified

App High-risk & Neurologist evaluated (N=107)

Questionnaire Score ≥ 5 (N = 98)
43 Confirmed PD 55
No PD confirmed

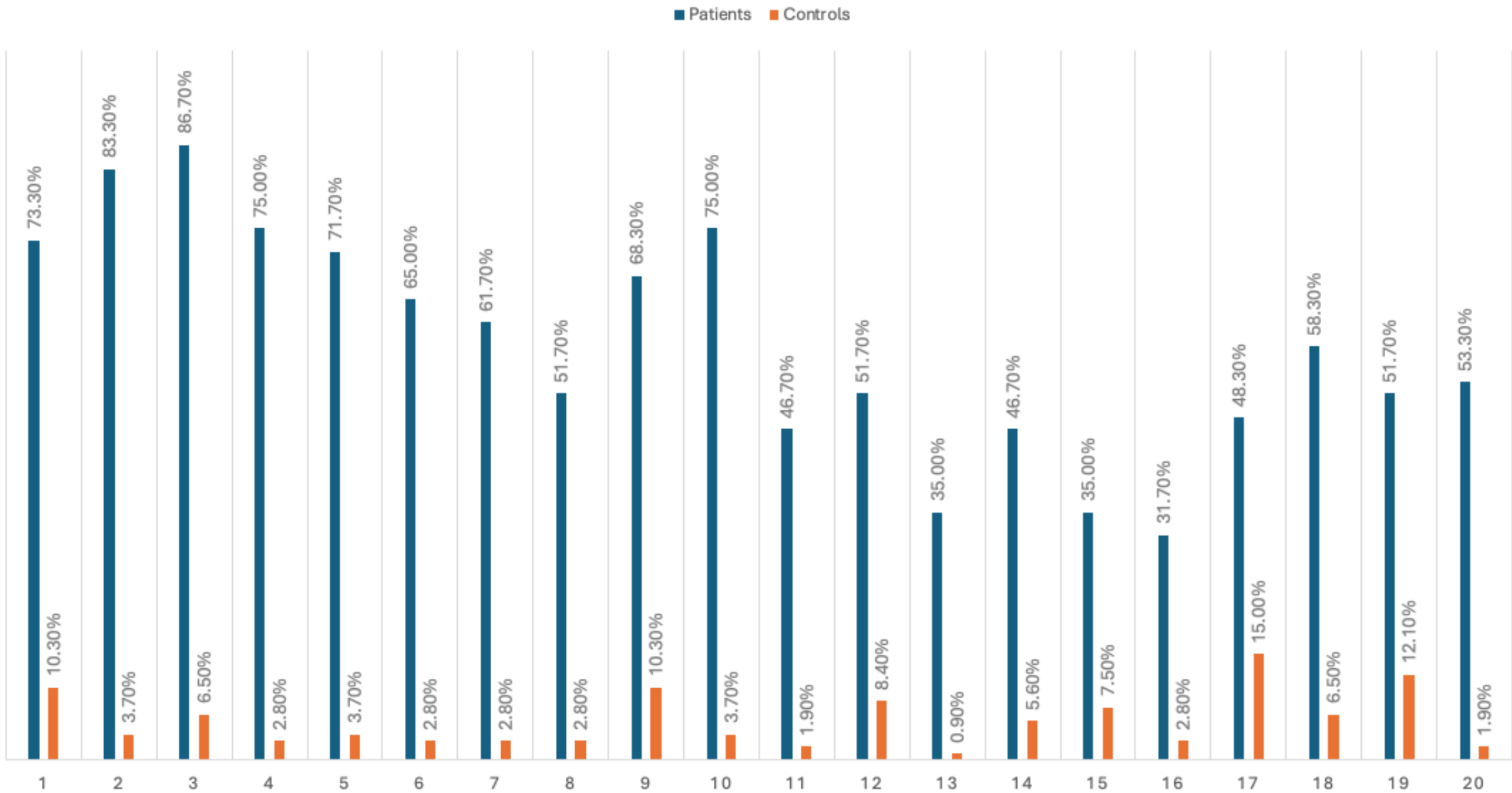
Questionnaire Score < 5 (N = 9)
4 Confirmed PD
5 No PD confirmed

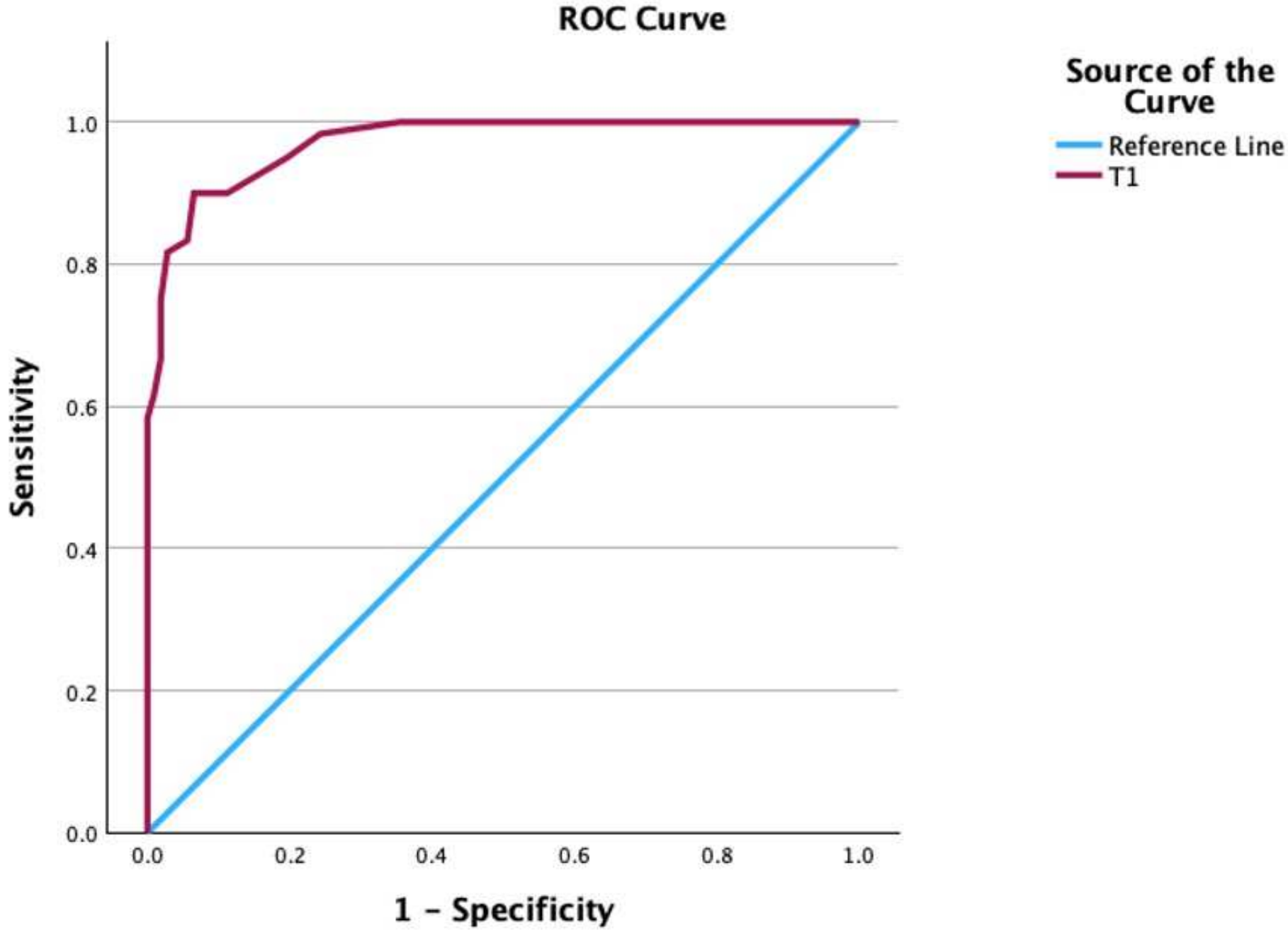
Specialist multimodal clinical assessment

Confirmed PD (N = 47) (Standard PD Treatment)

Non-PD Controls (N = 60) (Annual follow-up)

Exploratory Multimodal ML Subgroup Analysis
(N= 78: 20 Confirmed PD vs. 58 No PD)
[Matched 5-Fold Grouped Cross-Validation: Questionnaire-only (Q) vs. Questionnaire-plus-Sensor (Q+S)]


PERCENTAGE OF SUBJECTS WHO ANSWERED "YES" OR "POSITIVE" TO EACH ITEM OF PD SCREENING QUESTIONNAIRE BETWEEN HOSPITAL PD AND VOLUNTEER-CONTROL GROUPS
Patients
Controls
73.30%
10.30%
83.30%
3.70%
86.70%
6.50%
75.00%
2.80%
71.70%
3.70%
65.00%
2.80%
61.70%
2.80%
51.70%
2.80%
68.30%
10.30%
75.00%
3.70%
46.70%
1.90%
51.70%
8.40%
35.00%
0.90%
46.70%
5.60%
35.00%
7.50%
31.70%
2.80%
48.30%
15.00%
58.30%
6.50%
51.70%
12.10%
53.30%
1.90%
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20


ROC Curve
Sensitivity
1 – Specificity
0.0
0.2
0.4
0.6
0.8
1.0
Source of the Curve
Reference Line
T1

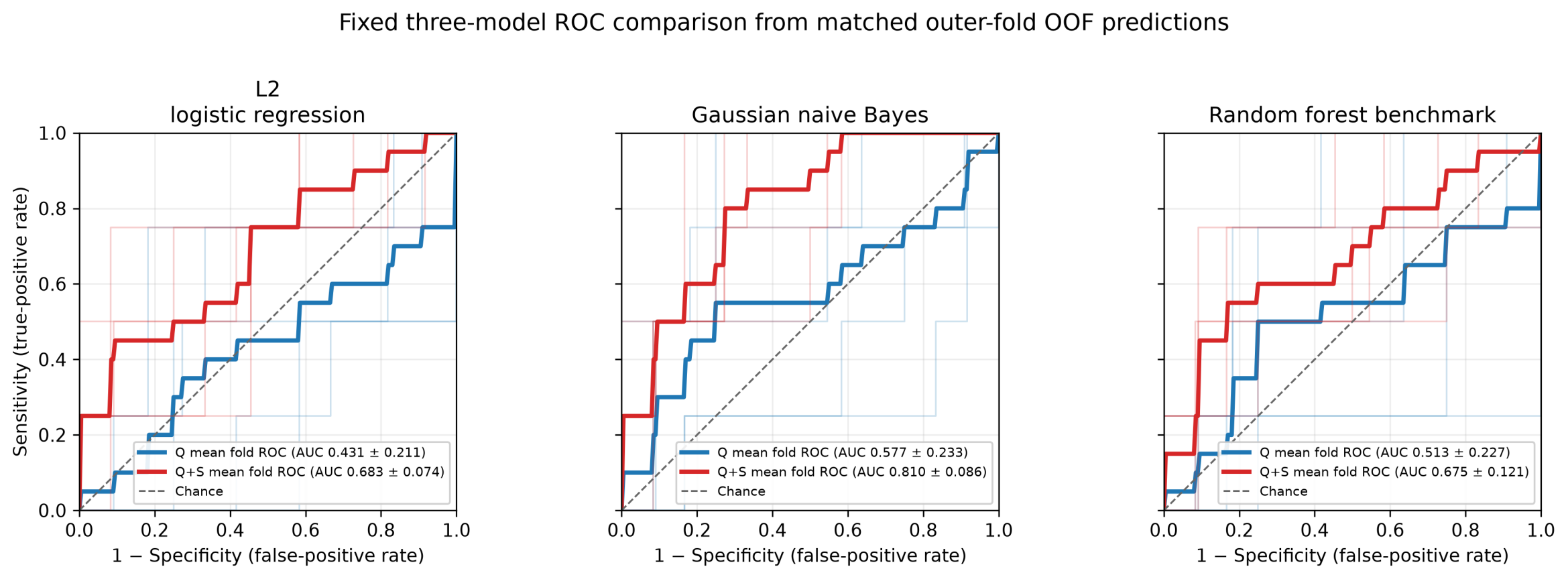

Fixed three-model ROC comparison from matched outer-fold OOF predictions
L2
logistic regression
Gaussian naive Bayes
Random forest benchmark
Sensitivity (true-positive rate)
1 − Specificity (false-positive rate)
Q mean fold ROC (AUC 0.431 ± 0.211)
Q+S mean fold ROC (AUC 0.683 ± 0.074)
Chance
Q mean fold ROC (AUC 0.577 ± 0.233)
Q+S mean fold ROC (AUC 0.810 ± 0.086)
Chance
Q mean fold ROC (AUC 0.513 ± 0.227)
Q+S mean fold ROC (AUC 0.675 ± 0.121)
Chance

Table 1. Number of subjects (PD and healthy controls) who answered "yes" or "positive" response to each item of PD screening questionnaire with sensitivity and specificity and adjusted odd ratio (OR).

	Patient (%)	**Control (%)**	**P-value**	**Sensitivity**	**Specificity**	**Adjusted OR by age and gender, (P value)**
1. Have you ever noticed a tremor of your hands or legs at rest?	44/60 (73.3%)	11/107 (10.3%)	<0.001*	73.33%	89.72%	48.70, [95% CI: 13.69, 173.30] ($p< 0.001^{\gamma}$)
2. Have you written slower, or your handwriting changed or become smaller?	50/60 (83.3%)	4/107 (3.7%)	<0.001*	83.33%	96.26%	125.89, [95% CI: 28.36, 558.85] ($p< 0.001^{\gamma}$)
3. Have you noticed that you become more clumsy or slower in any tasks: for example, combing,	52/60 (86.7%)	7/107 (6.5%)	<0.001*	86.67%	93.46%	57.34, [95% CI: 18.15, 181.14] ($p< 0.001^{\gamma}$)

dressing, or bathing or have more difficulty involving fine hand motor control task: for example, doing up your buttons or opening the bottle?						
4. Have you noticed that your voice has become softer or more monotonous or that listeners have to ask repeatedly because they cannot hear you speak?	45/60 (75.0%)	3/107 (2.8%)	<0.001*	75.00%	97.20%	101.52, [95% CI: 22.32, 461.80] ($p< 0.001^{\gamma}$)
5. Have you noticed that your arm swings less or	43/60 (71.7%)	4/107 (3.7%)	<0.001*	71.67%	96.26%	68.94, [95% CI: 17.24, 275.63] ($p< 0.001^{\gamma}$)

does not swing when you walk?						
6. Do you walk with short steps or shuffling gait?	39/60 (65.0%)	3/107 (2.8%)	<0.001*	65.00%	97.20%	63.08, [95% CI: 14.45, 275.32] ($p< 0.001^{\gamma}$)
7. Do you have difficulty to walk or being freeze while starting to walk, walking, or turning?	37/60 (61.7%)	3/107 (2.8%)	<0.001*	61.67%	97.20%	97.15, [95% CI: 19.90, 474.25] ($p< 0.001^{\gamma}$)
8. Do you have trouble with throwing yourself forward while walking causing you to keep up with your step or difficulty to stop walking immediately?	31/60 (51.7%)	3/107 (2.8%)	<0.001*	51.67%	97.20%	67.49, [95% CI: 13.90, 327.72] ($p< 0.001^{\gamma}$)

9. Do you have any of the following problems? Difficulty turning over when sleeping or getting out of bed or after sitting down.	41/60 (68.3%)	11/107 (10.3%)	<0.001*	68.33%	89.72%	23.26, [95% CI: 7.99, 67.73] ($p< 0.001^{\gamma}$)
10. Do your tremor, slowness, or stiffness start on one side of your body first?	45/60 (75.0%)	4/107 (3.7%)	<0.001*	75.00%	96.26%	117.89, [95% CI: 25.59, 543.04] ($p< 0.001^{\gamma}$)
11. Do you know that you have symptoms of speaking or shouting or moving your arms and legs that may be consistent with dreams or	28/60 (46.7%)	2/107 (1.9%)	<0.001*	46.67%	98.13%	78.50, [95% CI: 13.38, 460.64] ($p< 0.001^{\gamma}$)

fall off the bed while sleeping? Or have you ever been told by a bed partner or caregiver that you have these symptoms?						
12. Do you usually have excessive sleepiness during the day or fall asleep while doing activity?	31/60 (51.7%)	9/107 (8.4%)	<0.001*	51.67%	91.59%	13.51, [95% CI: 4.61, 39.58], ($p<0.001^{\gamma}$)
13. Do you feel your sense of smell has decreased?	21/60 (35.0%)	1/107 (0.9%)	<0.001*	35.00%	99.07%	55.69 [95% CI: 5.85, 529.71] ($p<0.001^{\gamma}$)
14. Do you have chronic constipation, defined as	28/60 (46.7%)	6/107 (5.6%)	<0.001*	46.67%	94.39%	9.75 [95% CI: 3.21, 29.66]

defecating less than 3 times a week in the last 3 months?						(p<0.001[γ])
15. Do you have symptoms of depression? Crying more easily than usual or lack of interest in the surrounding environment or things that used to be fun in the past.	21/60 (35.0%)	8/107 (7.5%)	<0.001*	35.00%	92.52%	12.74 [95% CI: 3.85, 42.17] (p<0.001[γ])
16. Have you ever seen a hallucination or heard a sound without a person?	19/60 (31.7%)	3/107 (2.8%)	<0.001*	31.67%	97.20%	18.36 [95% CI: 3.32, 101.69] (p<0.001[γ])
17. Do you often feel lightheaded or dizzy when changing positions	29/60 (48.3%)	16/107 (15.0%)	<0.001*	48.33%	85.05%	4.72 [95% CI: 1.92, 11.61] (p<0.001[γ])

from supine or sitting to standing, and symptoms usually improve or disappear after sitting or lying down?						
18. Do you usually have urinary control problems, such as being unable to control urine, urinary incontinence, or urinary retention?	35/60 (58.3%)	7/107 (6.5%)	<0.001*	58.33%	93.46%	17.25 [95% CI: 5.71, 52.14] ($p<0.001^{\gamma}$)
19. Do you have problems with analytical thinking, memory, or calculation that have worsened for more than 1 year?	31/60 (51.7%)	13/107 (12.1%)	<0.001*	51.67%	87.85%	5.72 [95% CI: 2.26, 14.45] ($p<0.001^{\gamma}$)

20. Do you have balance problems or frequent falls at the very beginning of tremor, slowness, or stiffness?	32/60 (53.3%)	2/107 (1.9%)	<0.001*	53.33%	98.13%	78.64 [95% CI: 14.64, 422.49] (p<0.001[γ])

Chi-square test

* Statistical significance when the p-value is less than 0.05.

Binary logistic regression test

[γ] Statistical significance when the p-value is less than 0.05.

Table 2. Characteristics of confirmed PD and evaluated non-PD among who had high risk results from CheckPD application

Parameters	**Confirmed PD (n=47)**	**Evaluated non-PD (n=60)**	**P value**
Age, mean ± SD (years)	65.9 ± 12.5	63.3 ± 14.0	0.313[a]
Gender, n % (male)	21/47 (44.7%)	9/60 (15.0%)	<0.001*[b]
Questionnaire score ≥ 5, n %	43/47 (91.5%)	55/60 (91.7%)	1.000[c]
Total questionnaire score, mean ± SD	8.5 ± 4.1	7.5 ± 3.4	0.144[a]

[a] Unpaired t-test

[b] Chi-square test

[c] Fisher's Exact test

* Statistical significance when the p-value is less than 0.05.

Table 3. Descriptive five-fold grouped cross-validation performance using questionnaire data alone (Q) and questionnaire data combined with sensor and voice features (Q+S)

Model	Metric	Q	Q+S	Difference (Q+S − Q)
L2 logistic regression	Sensitivity	0.200 ± 0.326	0.350 ± 0.224	+0.150 ± 0.137
	Specificity	0.779 ± 0.123	0.862 ± 0.118	+0.083 ± 0.105
	ROC-AUC	0.431 ± 0.211	0.683 ± 0.074	+0.252 ± 0.261
Gaussian naïve Bayes	Sensitivity	0.550 ± 0.326	0.850 ± 0.224	+0.300 ± 0.326
	Specificity	0.742 ± 0.054	0.655 ± 0.135	−0.088 ± 0.162
	ROC-AUC	0.577 ± 0.233	0.810 ± 0.086	+0.233 ± 0.251
Random forest	Sensitivity	0.000 ± 0.000	0.250 ± 0.177	+0.250 ± 0.177
	Specificity	0.967 ± 0.046	0.914 ± 0.059	−0.053 ± 0.076
	ROC-AUC	0.513 ± 0.227	0.675 ± 0.121	+0.161 ± 0.125

ROC-AUC, area under the receiver operating characteristic curve. Values are mean ± standard deviation across five matched grouped test folds. Differences are calculated as Q+S − Q within each fold. Positive values indicate an increase, and negative values indicate a decrease.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- STARD2015Checklistrbh24Oct2026.docx
- Suppl1CheckPDQuestionnairerbh24Oct2026.docx
- TripodChecklistrbh24Oct2026.doc