A simple-to-use web-based calculator for survival prediction in Parkinson's disease

Yunliang Tang[1], Jiao Wang[2], Gengfa Chen[1], Wen Ye[1], Nao Yan[3], Zhen Feng[1]

[1]Department of Rehabilitation Medicine, First Affiliated Hospital of Nanchang University, Nanchang 330006, Jiangxi, China
[2]Department of Endocrinology and Metabolism, First Affiliated Hospital of Nanchang University, Nanchang 330006, Jiangxi, China
[3]Department of Neurology, Zhongnan Hospital of Wuhan University, Wuhan 430071, Hubei, China

Correspondence to: Zhen Feng; **email:** fengzhen@email.ncu.edu.cn

ABSTRACT

Background: To establish and validate a nomogram and corresponding web-based calculator to predict the survival of patients with Parkinson's disease (PD).
Methods: In this cohort study, we retrospectively evaluated patients (n=497) with PD using a two-stage design, from March 2004 to November 2007 and from July 2005 to July 2015. Predictive variables included in the model were identified by univariate and multiple Cox proportional hazard analyses in the training set.
Results: Independent prognostic factors including age, PD duration, and Hoehn and Yahr stage were determined and included in the model. The model showed good discrimination power with the area under the curve (AUC) values generated to predict 4-, 6-, and 8-year survival in the training set being 0.716, 0.783, and 0.814, respectively. In the validation set, the AUCs of 4- and 6-year survival predictions were 0.85 and 0.924, respectively. Calibration plots and decision curve analysis showed good model performance both in the training and validation sets. For convenient application, we established a web-based calculator (https://tangyl.shinyapps.io/PDprognosis/).
Conclusions: We developed a satisfactory, simple-to-use nomogram and corresponding web-based calculator based on three relevant factors to predict prognosis and survival of patients with PD. This model can aid personalized treatment and clinical decision-making.

INTRODUCTION

Parkinson's disease (PD) is a progressive neurodegenerative disorder mostly defined by motor symptoms such as resting tremor, muscular rigidity, bradykinesia, and postural instability [1]. Globally, an estimated 6 million people suffer from PD, and this number is expected to double by 2040, making PD the world's fastest-growing neurological disorder [2]. The majority of PD cases are idiopathic, although 10-15% are genetic, known as familial PD [3]. Although its

precise etiology remains unclear, the pathophysiological basis is degeneration of dopaminergic neurons in the substantia nigra [4]. Current treatment for PD involves pharmaceutical, surgical, and dietary interventions, and/or rehabilitation exercises [5]. Dopamine replacement strategies have been the most effective pharmacotherapy for motor symptoms in PD [6]. Subthalamic nucleus deep brain stimulation is a particularly promising new therapy, shown to effectively improve not only motor function but also executive function in PD [7]. However, there is no

well-established treatment to stop or slow the disease progression in PD patients. Rapid PD progression is associated with a poor prognosis [8]. Therefore, there is an urgent need for reliable and accurate prognostic models for PD.

Numerous studies have aimed to discover potential PD biomarkers, including physiological or bio-chemical measurements, metabolic and genetic assessments, imaging findings, and rating scales that could represent candidate indicators [9–11]. For instance, Majbour et al. reported that cerebrospinal fluid alpha-synuclein species correlated with the progression of motor impairment [12], Chung et al. constructed a model to predict later development of gait freezing in PD [13], and Hideyuki et al. reported that high C-reactive protein (CRP) levels were associated with shorter survival time and suggested that they may predict survival prognosis in patients with PD [14]. Although considerable effort has been made, few of these biomarkers have been adopted in routine clinical practice. More importantly, to the best of our knowledge, there is no effective prognosis prediction model for patients with PD.

In the present study, we aimed to establish a model by using Cox regression based on long-term follow-up of the cohort to predict the prognosis of PD patients. We first established a nomogram based on the primary predictive model and it was displayed as a web-based calculator for convenient clinical use. The predictive value of the model was evaluated based on discrimination, calibration, and clinical utility in the training and validation sets. This simple-to-use model might serve as an early warning and prediction system for patients with PD.

RESULTS

Baseline patient characteristics

As shown in Figure 1, after excluding patients with unknown HY stage, MMSE scores, albumin, or NSAIDs history, the rest were included in the training (n=235) and validation sets (n=184). The baseline characteristics of the training and validation sets are described in Table 1. The median follow-up times in the training and validation sets were 69.3 (2.9, 118.9) and 19.3 (1.8, 79.3) months, respectively. During the follow-up period, 43 patients died and 192 survived in the training set. Further, the mortality in the training set after 4, 6, and 8 years was 24.4%, 26.8%, and 20.7%, respectively. In the validation set, 17 patients died and 170 survived. Additionally, the mortality in the validation set after 4 and 6 years was 7.1%, and 7.7%, respectively.

Prognostic factors in PD

Four variables including age, PD duration, Hoehn-Yahr (HY), and albumin, were significantly associated with overall survival as per the univariate regression analysis. Thereafter, three variables (age, PD duration, and HY)
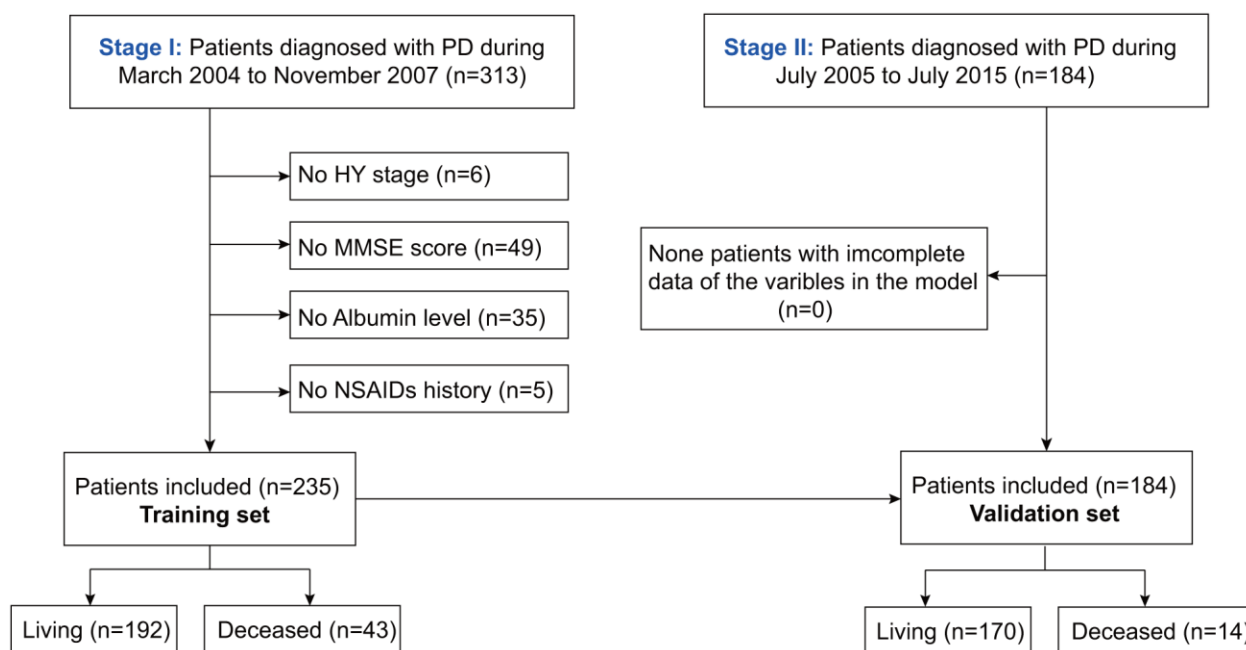

Figure 1. Flowchart of participant selection.

Table 1. Baseline characteristics of the training set and validation set.

Features	Training set (n=235)	Validation set (n=184)	P-value
Age (years)	69.25±9.12	72.97±8.06	<0.001
Sex			0.006
Male	100 (42.6%)	104(56.5%)	
Female	135 (57.4%)	80(43.5%)	
PD duration (years)	7.90±5.44	8.59±5.63	0.204
HY			0.547
≤3	148 (63.0%)	122 (66.3%)	
≥4	87 (37.0%)	62 (33.7%)	
MMSE	24.42±4.98	24.38±5.83	0.947
Albumin (mg/dL)	4.03±0.41	3.68±0.45	<0.001
CRP			-
<0.3mg/L	70 (29.8%)	-	
0.3-0.8mg/L	93 (39.6%)	-	
>0.8mg/L	72 (30.6%)	-	
NSAIDs			-
Non	195(83.0%)	-	
Current/habitual	40(17.0%)	-	
Vital status			0.002
Alive	192(81.7%)	170(92.4%)	
Deceased	43 (18.3%)	14 (7.6%)	

were identified as independent prognostic factors for PD by multivariate Cox regression analysis (Table 2).

Development of an individualized prediction model

Based on the Cox regression results, age, PD duration, and HY stage were incorporated into the nomogram (Figure 2A), which is an intuitive visualization model. According to the nomogram, age had the greatest influence on PD prognosis, followed by PD duration and HY stage. The total score is a sum of the individual scores of these three variables, and users finally obtain a specific probability of 4-, 6-, and 8-year survival.

Patient data included in the training set were clustered into high- (n = 117) and low-risk clusters (n = 118) according to the median risk score, following the risk score distribution shown in Figure 2B. The Kaplan–Meier survival curves exhibited significantly worse overall survival in the high-risk group (p=7.701e-7; Figure 2C). Training set patients' survival time and status are shown in Figure 2D.

Establishment of a web-based calculator

To facilitate the clinical application of our findings, we established a web-based calculator (https://tangyl.shinyapps.io/PDprognosis/) to predict the overall

survival of patients with PD according to the nomogram (Figure 3). For example, for patients aged 75 years with 22 years of PD duration and an HY stage 4-5, the 2079-day survival rate was approximately 45.0% (95% CI 22.6–76.0%).

Model performance in the training set

In the training set, the discrimination power of the model was assessed by concordance index (C-index) values and receiver operating characteristic (ROC) curves. The C-index of the model was 0.75. The area under the curve (AUC) values generated to predict 4-, 6-, and 8-year survival were 0.716, 0.783, and 0.814, respectively (Figure 4A), implying that the nomogram was efficient in predicting prognosis. Calibration plots based on the training set showed good consistency between nomogram prediction and actual observation (Figure 4B).

Decision curve analysis (DCA) is a suitable method for evaluating alternative prognostic strategies that have advantages over other commonly used measures [15, 16]. Therefore, we applied DCA to evaluate the clinical usefulness of the prognostic nomogram. As shown in Figure 4C, for a >20% probability threshold, patients with PD would gain more benefit from this prognostic nomogram than the hypothetical treat-all or treat-none scenarios.

Table 2. Univariate and multivariate Cox regression analyses of prognostic factors in patients with PD in training set.

Variable	Univariate analysis		Multivariate analysis	
	HR (95%CI)	P-value	HR (95%CI)	P-value
Age	1.056(1.015-1.099)	0.007	1.049(1.003-1.098)	**0.038**
Sex				
Male	Reference		Reference	
Female	0.745(0.414-1.374)	0.357	0.679(0.373-1.302)	0.258
PD durations	1.085(1.036-1.136)	0.001	1.067(1.012-1.125)	**0.017**
HY				
≤3	Reference		Reference	
≥4	5.097 (2.682-9.686)	0.000	2.737(1.300-5.761)	**0.008**
MMSE	0.957(0.893-1.025)	0.208	1.062(0.978-1.153)	0.153
CRP				
<0.3(mg/L)	Reference		Reference	
0.3-0.8(mg/L)	1.053(0.443-2.504)	0.908	0.914(0.379-2.202)	0.840
>0.8(mg/L)	2.879(1.314-6.308)	0.008	1.840(0.806-4.198)	0.148
Albumin(mg/dL)	0.388(0.191-0.787)	0.009	0.537(0.252-1.145)	0.537
NSAIDs				
Non	Reference		Reference	
Current/ habitual	1.006(0.447-2.262)	0.989	0.874(0.375-2.034)	0.754

Model performance in the validation set

In the validation set, risk-scores were calculated by the same formula used for calculating the patient risk-scores in the training set. Then, patients were divided into two low- and high- risk groups according to the model's median score. As the longest follow-up time was 6.5 years in the validation set, we mainly assessed the prognostic value for 4- and 6-year survival probability. Survival analysis indicated that low-risk patients had significantly better prognosis than high-risk patients (p=1.414e-05, Figure 5A). Moreover, this new model showed good discriminative power in prognosis predictions, as reflected by an AUC for 4- and 6-year survival of 0.85 and 0.924, respectively (Figure 5B, 5C). The risk score distribution and the survival status of individuals in the low- and high-risk groups are shown in Figure 5D, 5E.

Moreover, we performed calibration plot analysis as shown in Figure 6A, 6B the results of which suggested that this model has good probability consistencies between prediction and observation in the validation set. Similarly, using this model to predict PD survival has more net benefits (Figure 6C, 6D).

DISCUSSION

PD is one of the most common neurodegenerative diseases and has heterogeneous clinical outcomes;

hence, more accurate predictive models are required to guide treatment. As we know, there are many models for predicting disease progression in patients with PD. However, the survival prediction models of PD are limited. To the best of our knowledge, this is the first study to develop a nomogram and web-based calculator for predicting the survival of patients with PD.

In the present study, we established a model based on several readily available variables. The multivariable model based on three features (age, PD duration, and HY stage) showed the best predictive power in both training and validation sets. In the training set, the AUC of 4-, 6-, and 8-year survival was 0.716, 0.783, and 0.814, respectively. In the validation set, the AUC of 4- and 6-year survival reached 0.85 and 0.924, respectively. Model performance was also evaluated via calibration curves and DCA in both datasets. Our results suggest that the model could be used as a cost-effective tool to predict PD prognosis and assist with clinical decision-making.

Currently, researchers are exploring novel biomarkers for diagnosing and predicting mortality in patients with PD. While many studies have been conducted, few of these markers have been applied for predicting survival in patients with PD. Therefore, we choose easily obtainable variables that are convenient to apply in clinical practice. In the present study, we identified three independent risk factors, including age, PD
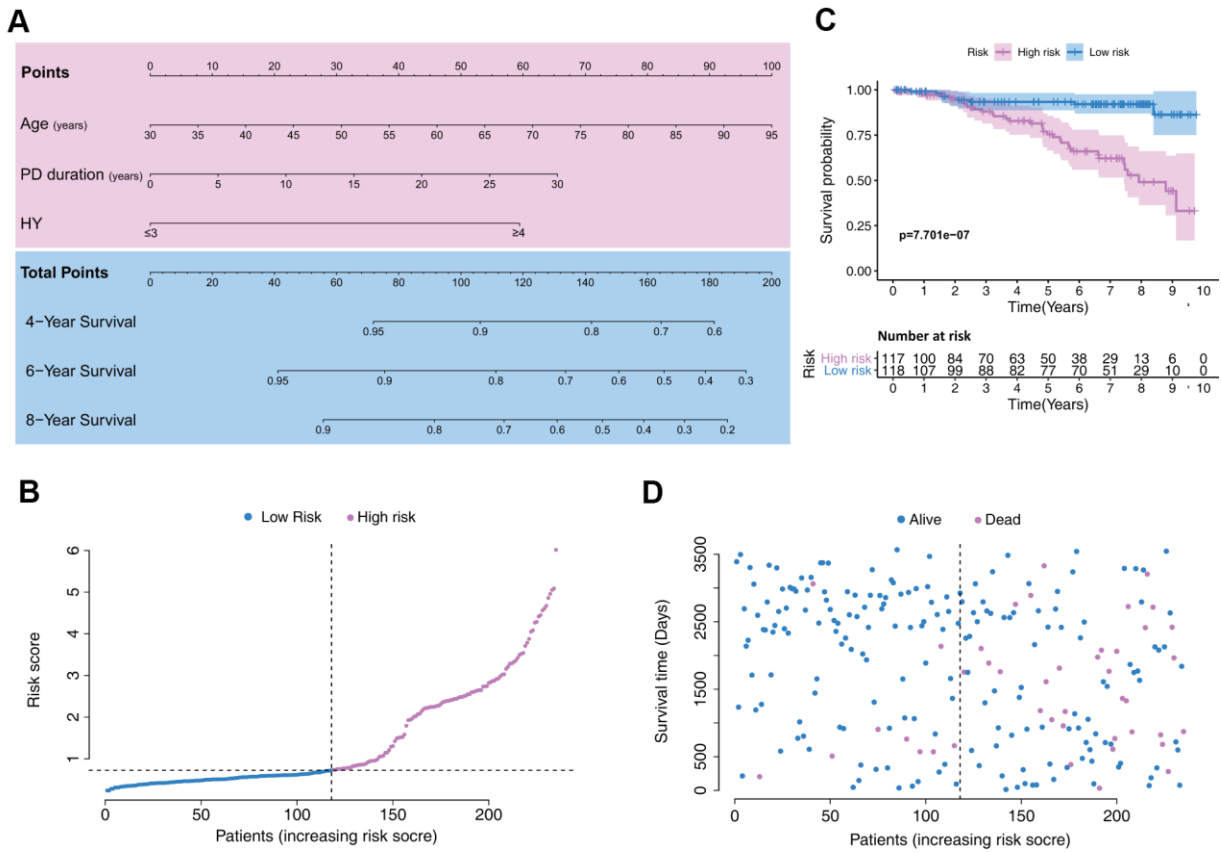
Figure 2. Establishment of a model with clinical indices to predict Parkinson's disease survival. (**A**) A clinical feature model was used to develop a nomogram. (**B**) Distribution of the risk scores in the low- and high-risk groups. (**C**) Kaplan–Meier survival analysis between low- and high-risk groups. (**D**) Patient distribution in the low- and high-risk groups based on survival status.
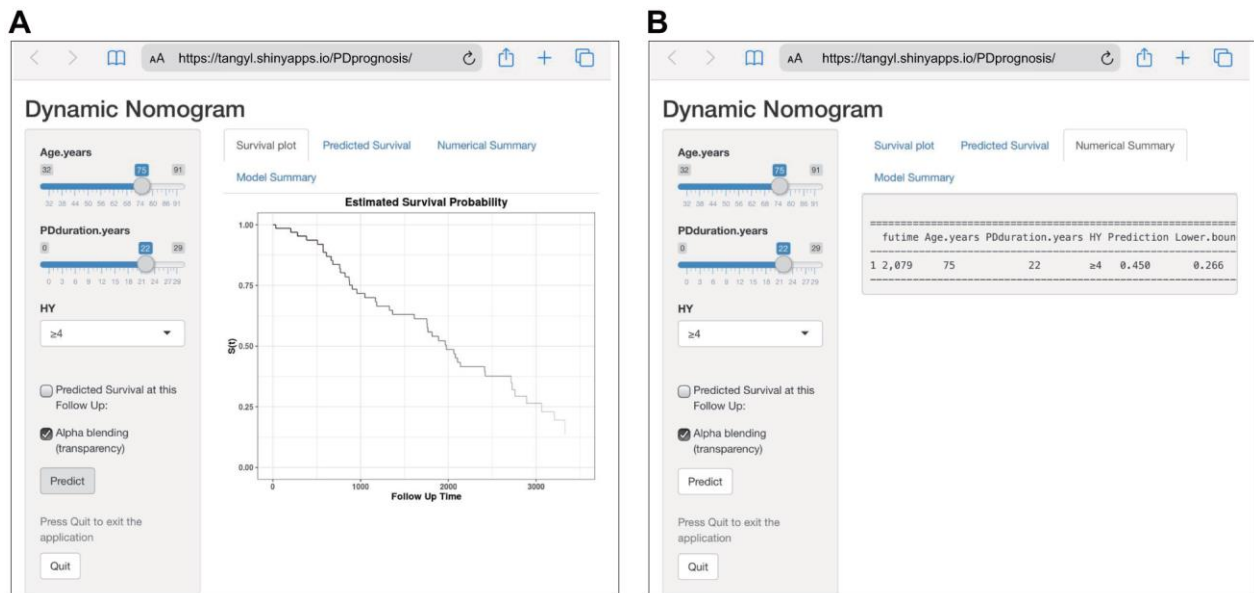
Figure 3. Construction of a web-based calculator for predicting Parkinson's disease survival based on the model (https://tangyl.shinyapps.io/PDprognosis/). (**A**) Web survival rate calculator. (**B**) 95% confidence interval of the web survival rate calculator.

duration, and HY stage, for predicting PD survival time. It has been established that age is a contributing factor in the progressive decline of dopamine transporter binding in healthy aging subjects [17]. PD is one of the most prevalent age-related neurodegenerative disorders, affecting 2–3% of the global population ≥65 years of age [18]. With an increase in age, the occurrence of PD also increases [19]. Thus, it has been regarded as the greatest risk factor for PD [20], and older age at baseline is associated with more rapid disease progression. Moreover, disease duration is also associated with PD progression [21, 22]. Nevertheless, there is evidence that age of PD patients contributes to disease severity, independent of the duration of the disease [23]. The HY scale is a widely used clinical standard for assessing PD stage, which is based on the level of clinical disability [24]. While not necessarily intimating pathophysiologic correlation, this method of staging is practical and
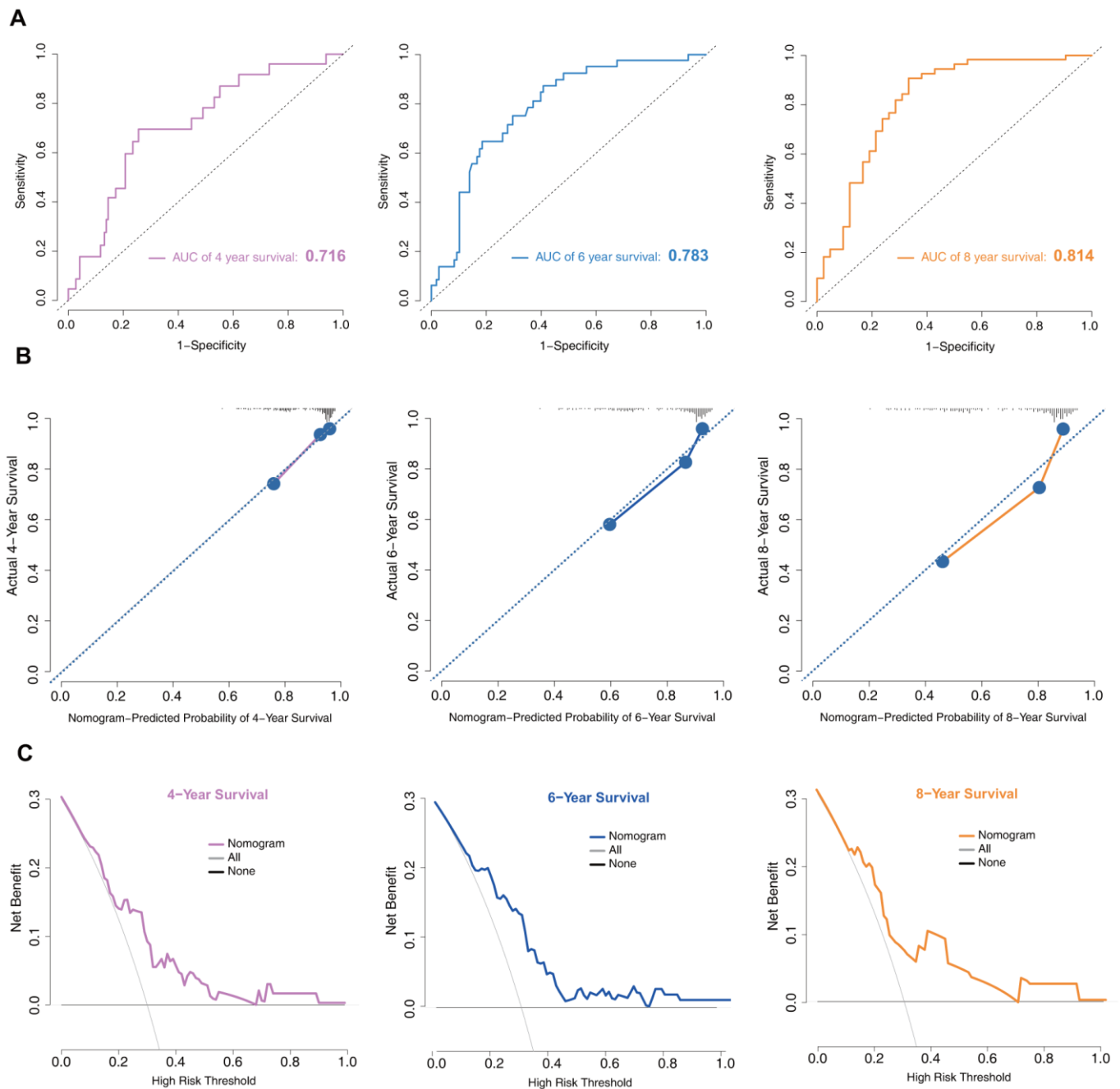

Figure 4. Model discrimination and performance in the training set. (**A**) Receiver operating characteristic curves for nomogram-based prognostic prediction. (**B**) Calibration plot examining estimation accuracy. (**C**) DCA assessing clinical utility.

allows for reproducible assessments by independent examiners of the general functional level of the patient. Combining these three clinical parameters, we developed a simple and accurate model.

Previous studies investigated single prognostic factors for PD [12, 25]. However, multivariable predictive models are known to show increased accuracy over individual risk factors. Therefore, unifactorial models are not appropriate for multifactorial diseases such as PD. Recently, studies have reported two models composed of several traditional factors [26, 27]. Macleod et al. reported a PD mortality model by combining four variables (age, sex, severity of axial features, and Charlson comorbidity index) with a moderate discrimination power (AUC=0.75 in the training set) [27]. Velseboer et al.'s study identified a 3-variable predictor model with higher patient age, higher UPDRSME axial score, and lower animal fluency score, all giving a higher probability of an unfavorable outcome, with an AUC value was 0.765 [26]. However, there were several flaws in these two studies. First, they used a small sample size with less than 400 participants. Second, their models could only predict the survival probability for a single time-point.

Our study has several advantages when compared with previous studies. A major strength of this study was that the nomogram was developed and validated in two complete cohorts of patients diagnosed with PD of sufficient sample size. This is important as the samples of previous studies may have been too small, which might lead to inaccurate outcome predictions. We tested the nomogram's discrimination power by an ROC analysis, finding that the AUC of the 8-year survival prediction in the training set reached 0.814, while that of the 6-year survival prediction in the validation set reached 0.924. The prediction accuracies significantly increased when compared to the previous study [26, 27]. Further, all three variables (age, PD duration, and HY stage) in our nomogram are easily available. Therefore, medical personnel can quickly perform an assessment without laboratory tests or instrumental examinations. Moreover, to facilitate clinical application, we established a web-based calculator to provide a free prediction service. It is convenient to provide the individual mortality percentages at different time-points. In addition, it cuts the cost and time needed for assessments both for PD patients and medical staff.
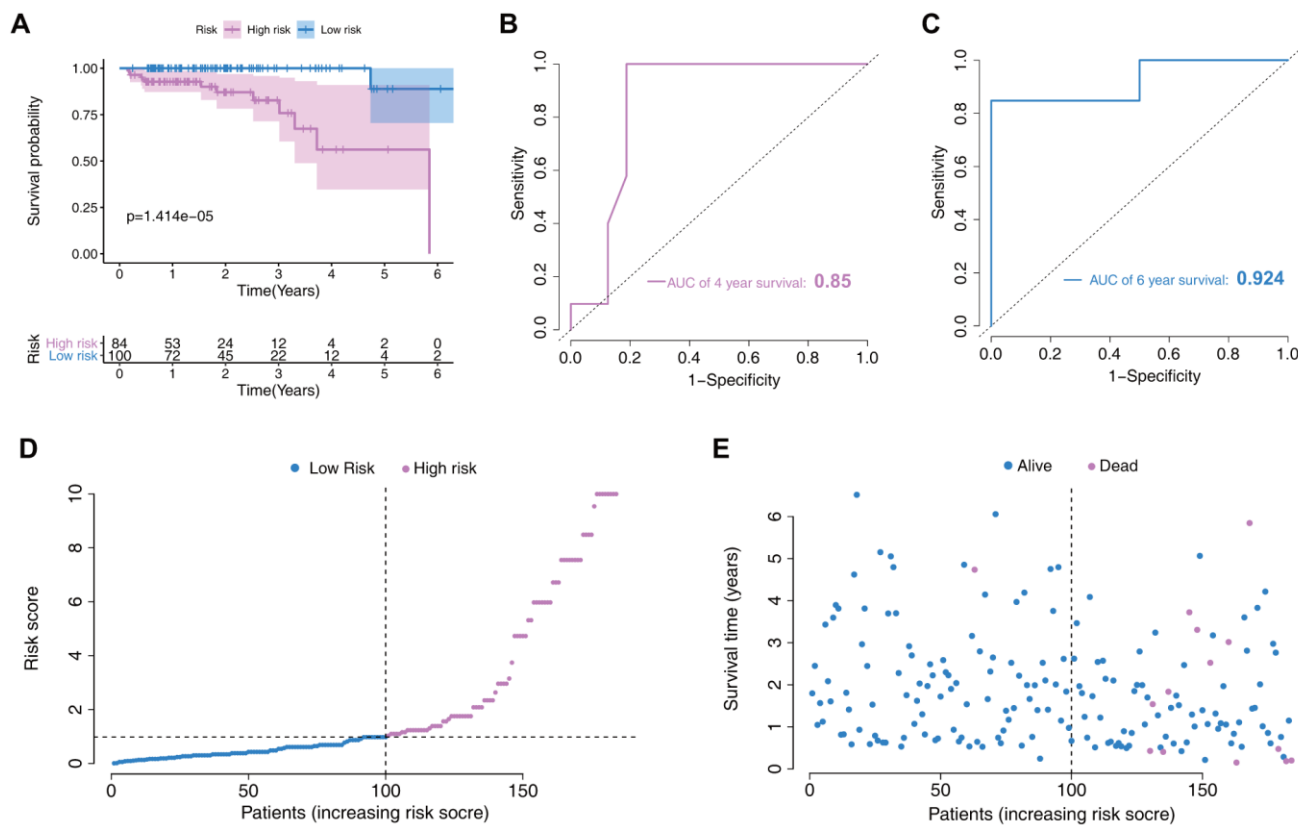

Figure 5. Nomogram verification in the validation set. (**A**) Kaplan–Meier survival analysis between the low- and high-risk groups. (**B**) Receiver operating characteristic (ROC) curves for 4-year prediction. (**C**) ROC curves for 6-year prediction. (**D**) Distribution of the risk scores in low- and high-risk groups. (**E**) Patient distribution in the low- and high-risk groups based on survival status.

There are certain limitations to this study. First, eight probable prediction variables were included in the analysis, while other features, such as dementia, psychotic, and other symptoms, which may function as independent risk factors for PD survival, failed to be considered in this model. Second, as racial differences may contribute to different PD outcomes [28], and participants in this study only included those of Asian ethnicity, the conclusions from this study and this predictor might not be generalizable. Third, the primary outcome measure was patient survival; other outcomes,

including postural instability or dementia, should be assessed in future studies.

In conclusion, we identified three variables and developed a novel nomogram and a web-based calculator to predict survival in patients with PD. These results may help further improve clinical decision-making and individualized treatment for PD patients. Furthermore, since this model could distinguish patients with high PD risk, it could help to closely follow-up on those patients.
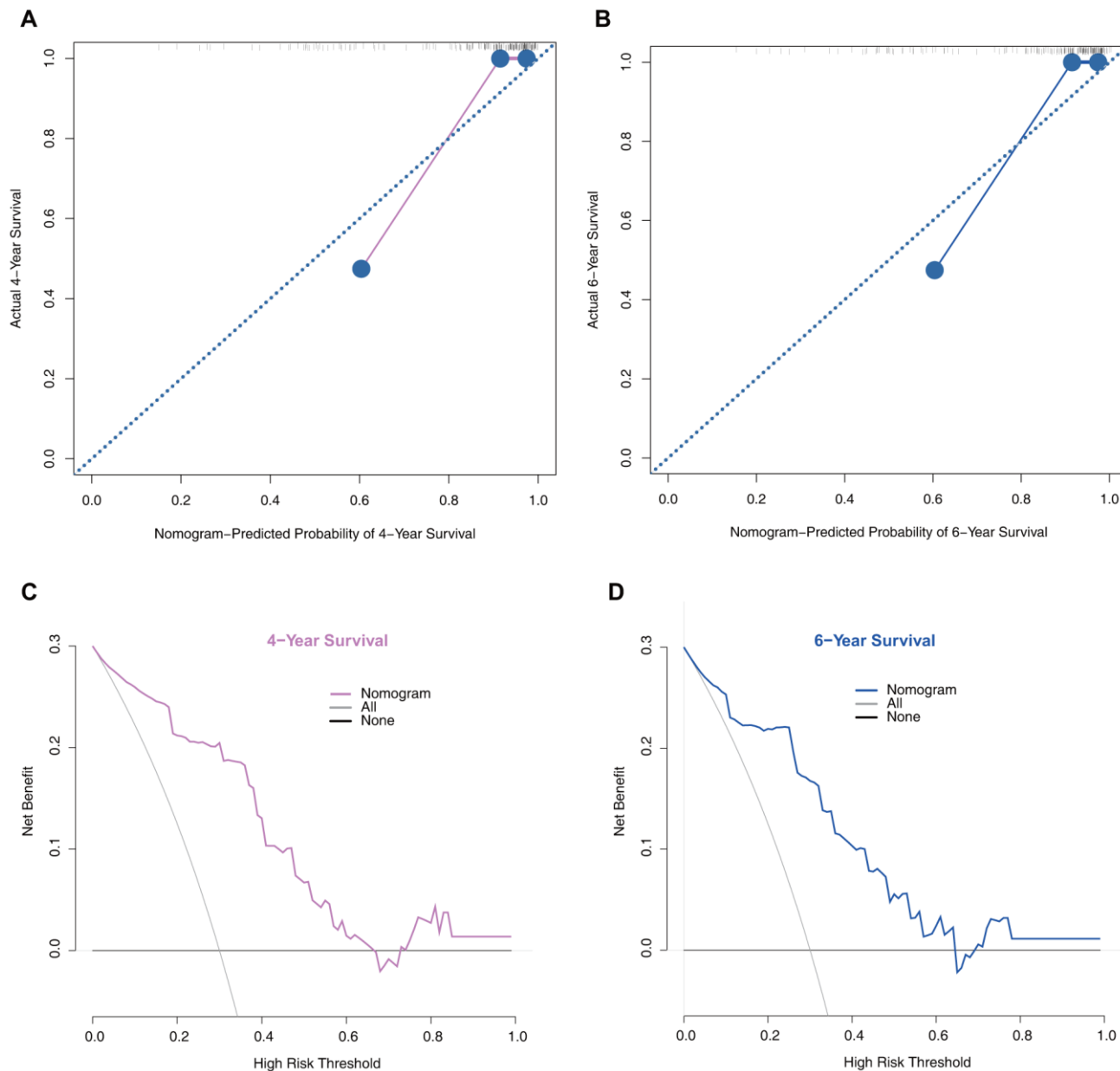
Figure 6. Model performance in the validation set. (**A**) 4-year calibration plot examining the estimation accuracy. (**B**) 6-year calibration plot examining the estimation accuracy. (**C**) 4-year decision curve analyses assessing clinical utility. (**D**) 6-year decision curve analyses assessing clinical utility.

MATERIALS AND METHODS

Patients

Raw data were extracted from a non-profit repository named the Dryad Digital Repository (http://www.datadryad.org/). This cohort study was conducted at the Utano National Hospital Parkinson's Disease Centre. The study design comprised two stages. The first stage (training set) retrospectively identified patients with PD from March 2004 to November 2007 [14]. The enrollment criteria included PD patients free of any infection. The second stage (validation set) enrolled PD patients from July 2005 to July 2015 [29]. These patients underwent brain magnetic resonance imaging to exclude other neurologic disorders. Moreover, patients were excluded if they were undergoing tube feeding, had undergone tracheostomy, or had other diseases that could cause dysphagia. All PD patients were diagnosed according to the United Kingdom Parkinson's disease Brain Bank Clinical Diagnostic Criteria. The study was approved by the Bioethics Committee of Utano National Hospital (No.28-15). Detailed information on the original studies is shown in Supplementary Table 1.

Data collection

We performed a secondary analysis based on data from the above two stages. Eight probable prediction variables were selected, including age, sex, PD duration, HY stages, Mini-Mental State Examination (MMSE) scores, serum albumin, CRP, and nonsteroidal anti-inflammatory drug (NSAIDs) history. Moreover, survival time and survival status of each patient was extracted.

Statistical analysis

To obtain a subset of predictors, univariate and multivariate Cox proportional hazards analyses were performed to select the optimal predictors from the risk factors in the training set. Variables with a p-value <0.05 in multivariable Cox regression were to act as independent predictors [30]. By using the selected variables based on the Cox regression results, a nomogram was developed using the "rms", "survival", and "foreign" R packages. In addition, the "DynNom" and "survival" R packages were used to construct a web-based calculator for predicting PD survival. The C-index and ROC curve were used to evaluate the discrimination ability of the model. Nomogram performance was assessed by a calibration plot and DCA, both in the training and validation cohorts.

Normal distribution data are expressed as means ± standard deviation, and non-normal data are expressed as medians (interquartile range).

Differences between the training and validation sets were analyzed using chi-square tests for categorical variables and t-tests for continuous variables. The Kaplan–Meier method and the log-rank test were used to estimate survival. All analyses were performed with the software R (Version 3.6.2; http://www.Rproject.org). A value of p-value <0.05 was considered statistically significant.

Data sharing

Data are available from the Dryad Digital Repository (http://www.datadryad.org/).

Abbreviations

PD: Parkinson's disease; HY: Hoehn and Yahr stage; AUC: area under the curve; CRP: C-reactive protein; MMSE: Mini-Mental State Examination; NSAIDs: nonsteroidal anti-inflammatory drug; C-index: concordance index; ROC: receiver operating characteristic; DCA: decision curve analyses.

AUTHOR CONTRIBUTIONS

Tang Yunliang collected the data. Feng Zhen designed the study. Wang Jiao performed statistical analysis of the data. Tang Yunliang and Wang Jiao prepared the manuscript. Chen Gengfa, Ye Wen, and Yan Nao searched the relevant literature, and Feng Zhen secured funding for the project. All authors have read and approved the final manuscript.

CONFLICTS OF INTEREST

All authors declare no potential conflicts of interest.

REFERENCES

1. Mor DE, Tsika E, Mazzulli JR, Gould NS, Kim H, Daniels MJ, Doshi S, Gupta P, Grossman JL, Tan VX, Kalb RG, Caldwell KA, Caldwell GA, et al. Dopamine induces

soluble α-synuclein oligomers and nigrostriatal degeneration. Nat Neurosci. 2017; 20:1560–68. https://doi.org/10.1038/nn.4641 PMID:28920936

2. Dorsey ER, Bloem BR. The Parkinson pandemic-A call to action. JAMA Neurol. 2018; 75:9–10. https://doi.org/10.1001/jamaneurol.2017.3299 PMID:29131880

3. Deng H, Wang P, Jankovic J. The genetics of Parkinson disease. Ageing Res Rev. 2018; 42:72–85. https://doi.org/10.1016/j.arr.2017.12.007 PMID:29288112

4. Dickson DW. Neuropathology of Parkinson disease. Parkinsonism Relat Disord. 2018 (Suppl 1); 46:S30–33. https://doi.org/10.1016/j.parkreldis.2017.07.033 PMID:28780180

5. Dong W, Qiu C, Jiang X, Shen B, Zhang L, Liu W, Zhang W, Chen J. Can the executive control network be used to diagnose Parkinson's disease and as an efficacy indicator of deep brain stimulation? Parkinsons Dis. 2020; 2020:6348102. https://doi.org/10.1155/2020/6348102 PMID:32148755

6. Peng Q, Zhong S, Tan Y, Zeng W, Wang J, Cheng C, Yang X, Wu Y, Cao X, Xu Y. The rodent models of dyskinesia and their behavioral assessment. Front Neurol. 2019; 10:1016. https://doi.org/10.3389/fneur.2019.01016 PMID:31681132

7. Jahanshahi M, Ardouin CM, Brown RG, Rothwell JC, Obeso J, Albanese A, Rodriguez-Oroz MC, Moro E, Benabid AL, Pollak P, Limousin-Dowsey P. The impact of deep brain stimulation on executive function in Parkinson's disease. Brain. 2000; 123:1142–54. https://doi.org/10.1093/brain/123.6.1142 PMID:10825353

8. Homayoun H. Parkinson disease. Ann Intern Med. 2018; 169:ITC33–48. https://doi.org/10.7326/AITC201809040 PMID:30178019

9. Li T, Le W. Biomarkers for Parkinson's disease: how good are they? Neurosci Bull. 2020; 36:183–94. https://doi.org/10.1007/s12264-019-00433-1 PMID:31646434

10. Fricova D, Harsanyiova J, Kralova Trancikova A. Alpha-synuclein in the gastrointestinal tract as a potential biomarker for early detection of Parkinson's disease. Int J Mol Sci. 2020; 21:8666. https://doi.org/10.3390/ijms21228666 PMID:33212934

11. Farotti L, Paolini Paoletti F, Simoni S, Parnetti L. Unraveling pathophysiological mechanisms of Parkinson's disease: contribution of CSF biomarkers. Biomark Insights. 2020; 15:1177271920964077. https://doi.org/10.1177/1177271920964077 PMID:33110345

12. Majbour NK, Vaikath NN, Eusebi P, Chiasserini D, Ardah M, Varghese S, Haque ME, Tokuda T, Auinger P, Calabresi P, Parnetti L, El-Agnaf OM. Longitudinal changes in CSF alpha-synuclein species reflect Parkinson's disease progression. Mov Disord. 2016; 31:1535–42. https://doi.org/10.1002/mds.26754 PMID:27548849

13. Chung SJ, Lee HS, Yoo HS, Lee YH, Lee PH, Sohn YH. Patterns of striatal dopamine depletion in early Parkinson disease: prognostic relevance. Neurology. 2020; 95:e280–90. https://doi.org/10.1212/WNL.0000000000009878 PMID:32616674

14. Sawada H, Oeda T, Umemura A, Tomita S, Kohsaka M, Park K, Yamamoto K, Sugiyama H. Baseline c-reactive protein levels and life prognosis in Parkinson disease. PLoS One. 2015; 10:e0134118. https://doi.org/10.1371/journal.pone.0134118 PMID:26218286

15. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making. 2006; 26:565–74. https://doi.org/10.1177/0272989X06295361 PMID:17099194

16. Fitzgerald M, Saville BR, Lewis RJ. Decision curve analysis. JAMA. 2015; 313:409–10. https://doi.org/10.1001/jama.2015.37 PMID:25626037

17. Eusebio A, Azulay JP, Ceccaldi M, Girard N, Mundler O, Guedj E. Voxel-based analysis of whole-brain effects of age and gender on dopamine transporter SPECT imaging in healthy subjects. Eur J Nucl Med Mol Imaging. 2012; 39:1778–83. https://doi.org/10.1007/s00259-012-2207-8 PMID:22890804

18. Poewe W, Seppi K, Tanner CM, Halliday GM, Brundin P, Volkmann J, Schrag AE, Lang AE. Parkinson disease. Nat Rev Dis Primers. 2017; 3:17013. https://doi.org/10.1038/nrdp.2017.13 PMID:28332488

19. Parveen R, Shamsi TN, Fatima S. Nanoparticles-protein interaction: role in protein aggregation and clinical implications. Int J Biol Macromol. 2017; 94:386–95. https://doi.org/10.1016/j.ijbiomac.2016.10.024 PMID:27746352

20. McDonald C, Gordon G, Hand A, Walker RW, Fisher JM. 200 years of Parkinson's disease: what have we learnt from James Parkinson? Age Ageing. 2018; 47:209–14.

https://doi.org/10.1093/ageing/afx196
PMID:29315364

21. Kordower JH, Olanow CW, Dodiya HB, Chu Y, Beach TG, Adler CH, Halliday GM, Bartus RT. Disease duration and the integrity of the nigrostriatal system in Parkinson's disease. Brain. 2013; 136:2419–31.
https://doi.org/10.1093/brain/awt192
PMID:23884810

22. Claassen DO, Dobolyi DG, Isaacs DA, Roman OC, Herb J, Wylie SA, Neimat JS, Donahue MJ, Hedera P, Zald DH, Landman BA, Bowman AB, Dawant BM, Rane S. Linear and curvilinear trajectories of cortical loss with advancing age and disease duration in Parkinson's disease. Aging Dis. 2016; 7:220–29.
https://doi.org/10.14336/AD.2015.1110
PMID:27330836

23. Szewczyk-Krolikowski K, Tomlinson P, Nithi K, Wade-Martins R, Talbot K, Ben-Shlomo Y, Hu MT. The influence of age and gender on motor and non-motor features of early Parkinson's disease: initial findings from the Oxford Parkinson Disease Center (OPDC) discovery cohort. Parkinsonism Relat Disord. 2014; 20:99–105.
https://doi.org/10.1016/j.parkreldis.2013.09.025
PMID:24183678

24. Hoehn MM, Yahr MD. Parkinsonism: onset, progression and mortality. Neurology. 1967; 17:427–42.
https://doi.org/10.1212/wnl.17.5.427
PMID:6067254

25. Macleod AD, Taylor KS, Counsell CE. Mortality in Parkinson's disease: a systematic review and meta-analysis. Mov Disord. 2014; 29:1615–22.
https://doi.org/10.1002/mds.25898
PMID:24821648

26. Velseboer DC, de Bie RM, Wieske L, Evans JR, Mason SL, Foltynie T, Schmand B, de Haan RJ, Post B, Barker RA, Williams-Gray CH. Development and external validation of a prognostic model in newly diagnosed Parkinson disease. Neurology. 2016; 86:986–93.
https://doi.org/10.1212/WNL.0000000000002437
PMID:26888991

27. Macleod AD, Dalen I, Tysnes OB, Larsen JP, Counsell CE. Development and validation of prognostic survival models in newly diagnosed Parkinson's disease. Mov Disord. 2018; 33:108–16.
https://doi.org/10.1002/mds.27177 PMID:28976022

28. Nuytemans K, Manrique CP, Uhlenberg A, Scott WK, Cuccaro ML, Luca CC, Singer C, Vance JM. Motivations for participation in Parkinson disease genetic research among hispanics versus non-hispanics. Front Genet. 2019; 10:658.
https://doi.org/10.3389/fgene.2019.00658
PMID:31379924

29. Tomita S, Oeda T, Umemura A, Kohsaka M, Park K, Yamamoto K, Sugiyama H, Sawada H. Video-fluoroscopic swallowing study scale for predicting aspiration pneumonia in Parkinson's disease. PLoS One. 2018; 13:e0197608.
https://doi.org/10.1371/journal.pone.0197608
PMID:29874285

30. Abougergi MS, Peluso H, Saltzman JR. Thirty-day readmission among patients with non-variceal upper gastrointestinal hemorrhage and effects on outcomes. Gastroenterology. 2018; 155:38–46.e1.
https://doi.org/10.1053/j.gastro.2018.03.033
PMID:29601829

Supplementary Table

Supplementary Table 1. Details of the original studies.

	Stage I (training set)	Stage II (validation set)
N	313	184
Performed time	March 2004 to November 2007	July 2005 and July 2015
Enrollment criteria	PD were diagnosed according to the United Kingdom Parkinson's disease Brain Bank Clinical Diagnostic Criteria	PD were diagnosed according to the United Kingdom Parkinson's disease Brain Bank Clinical Diagnostic Criteria
Exclude criteria	Any infection	Other neurologic disorders; Receiving tube feedings; Had a tracheostomy; Other diseases that could cause dysphagia; Observed for fewer than six months after VFSS;
Data collection	CRP, age, sex, PD duration, HY, MMSE, albumin, NSAID use, follow- up time	Age, sex, PD duration, UPDRS-3 score, HY stage, BMI, MMSE, Serum albumin, Patients consuming processed diets, follow-up time
Study design	Kaplan-Meier curves for the cumulative incidence of death were obtained after dividing patients into two groups according to clinical features. The log-rank test was used to determine the associations between life prognosis and clinical factors.	The patients were divided into two groups: those who developed aspiration pneumonia within six months after VFSS (cases) and those who did not develop aspiration pneumonia (controls)
Outcome	Survival time	The development of aspiration pneumonia; Survival time

VFSS: Video-fluoroscopic swallowing study.