World Journal of

Psychiatry

World Journal of
Psychiatry

Contents

ABOUT COVER

Associate editor of *World Journal of Psychiatry*, Dr. Giuseppe Lanza is Senior Academic Researcher and Assistant Professor at the University of Catania (Italy). After graduation with honors in Medicine, he trained in the School of Neurology and was awarded the international PhD from the same University. From 2013 to 2018, he worked as Consultant Neurologist at the Oasi Research Institute–IRCCS in Troina (Italy), a World Health Organization (WHO) collaborating center. He has co-authored more than 130 publications in peer-reviewed journals and conferences and reviewed for several international journals. More recently, he obtained a Master of Science in Clinical Research and the National Scientific Qualification as Full Professor. His ongoing research interests involve the translational neurosciences, particularly the application of non-invasive brain stimulation techniques in the diagnosis and treatment of neurological and neuropsychiatric disorders. (L-Editor: Filipodia)

AIMS AND SCOPE

The primary aim of *World Journal of Psychiatry* (*WJP*, *World J Psychiatr*) is to provide scholars and readers from various fields of psychiatry with a platform to publish high-quality basic and clinical research articles and communicate their research findings online.

WJP mainly publishes articles reporting research results and findings obtained in the field of psychiatry and covering a wide range of topics including adolescent psychiatry, biological psychiatry, child psychiatry, community psychiatry, ethnopsychology, psychoanalysis, psychosomatic medicine, *etc*.

INDEXING/ABSTRACTING

The *WJP* is now abstracted and indexed in Science Citation Index Expanded (SCIE, also known as SciSearch®), Current Contents/Clinical Medicine, Journal Citation Reports/Science Edition, PubMed, and PubMed Central. The 2020 edition of Journal Citation Reports® cites the 2019 impact factor (IF) for *WJP* as 3.545; IF without journal self cites: 3.545; Ranking: 46 among 155 journals in psychiatry; and Quartile category: Q2.

RESPONSIBLE EDITORS FOR THIS ISSUE

Production Editor: *Jia-Hui Li*; **Production Department Director:** *Yun-Xiaojian Wu*; **Editorial Office Director:** *Jia-Ping Yan*.

Observational Study

Development of a depression in Parkinson's disease prediction model using machine learning

Haewon Byeon

Haewon Byeon, Major in Medical Big Data, College of AI Convergence, Inje University, Gimhae 50834, Gyeonsangnamdo, South Korea

Corresponding author: Haewon Byeon, DSc, PhD, Professor, Major in Medical Big Data, College of AI Convergence, Inje University, Major in Medical Big Data, College of AI Convergence, Inje University, Gimhae 50834, Gyeonsangnamdo, South Korea. bhwpuma@naver.com

Abstract

BACKGROUND
It is important to diagnose depression in Parkinson's disease (DPD) as soon as possible and identify the predictors of depression to improve quality of life in Parkinson's disease (PD) patients.

AIM
To develop a model for predicting DPD based on the support vector machine, while considering sociodemographic factors, health habits, Parkinson's symptoms, sleep behavior disorders, and neuropsychiatric indicators as predictors and provide baseline data for identifying DPD.

METHODS
This study analyzed 223 of 335 patients who were 60 years or older with PD. Depression was measured using the 30 items of the Geriatric Depression Scale, and the explanatory variables included PD-related motor signs, rapid eye movement sleep behavior disorders, and neuropsychological tests. The support vector machine was used to develop a DPD prediction model.

RESULTS
When the effects of PD motor symptoms were compared using "functional weight", late motor complications (occurrence of levodopa-induced dyskinesia) were the most influential risk factors for Parkinson's symptoms.

CONCLUSION
It is necessary to develop customized screening tests that can detect DPD in the early stage and continuously monitor high-risk groups based on the factors related to DPD derived from this predictive model in order to maintain the emotional health of PD patients.

Core Tip: When the effects of Parkinson's disease (PD) motor symptoms were compared using "functional weight", the occurrence of levodopa-induced dyskinesia was the most influential risk factor in the diagnosis of depression in Parkinson's disease (DPD). These results can be used as baseline information to prevent DPD and establish management strategies. It is necessary to develop customized screening tests that can detect DPD patients in the early stage and continuously monitor high-risk groups based on the factors related to DPD derived from this predictive model in order to maintain the emotional health of PD. It is also necessary to develop customized programs for managing depression from the onset of PD.

INTRODUCTION

Parkinson's disease (PD) is a typical degenerative disease of the elderly with the second-highest incidence rate after Alzheimer's disease. The incidence rate of PD increases worldwide as the population ages. The Health Insurance Research and Assessment Service (2017)[1] reported that the number of patients with PD increased from 39265 in 2004 to 96499 in 2016, a 2.5-fold increase over 13 years, in South Korea. If the current trend in the aging population continues, the number of patients with PD will increase even further.

Although the primary symptom of PD is dyskinesia such as rigidity, it is highly likely that non-motor symptoms such as depression or cognitive impairment develop as PD progresses[2]. Among these symptoms, depressive symptoms are the most common non-motor symptoms of PD and previous studies[3-6] have reported that 35% to 75% of patients with PD suffer from depression, which is much higher than the prevalence of depression in adults in the local community (< 10%)[7]. Depression is known to adversely affect the quality of life in PD patients[8], and Parkinson's patients with depression have significantly higher anxiety symptoms, pessimism, suicidal thoughts, and self-condemnation compared to Parkinson's patients without depression[4,9]. Lee *et al*[10] analyzed 4362 patients with PD and reported that the elderly with PD had a 2-fold higher risk of suicide compared to healthy elderly. Even though patients with PD frequently experience depression[6,11], only 1% of them recognize that they have depression. These results suggest that it is necessary to diagnose depression in PD (DPD) as soon as possible[7,8]. Depression induced by PD requires social management and should not be considered a personal matter. As the symptoms of depression deteriorate as PD progresses, these symptoms, together with cognitive impairment, not only increase direct costs such as examinations and treatment but also raise indirect costs such as job loss due to disability and the care burden of supporting family members[12,13]. Ultimately, it causes unnecessary social expenditure at the national level. However, there are few studies on the characteristics and related factors of depression, a non-motor symptom of PD, compared to the motor symptoms of PD. Many medical practitioners are still more interested in the motor symptoms of PD than the non-motor symptoms. Moreover, depression is often misdiagnosed due to the non-motor and motor symptoms associated with PD[5]. It is important to diagnose depression in PD patients as soon as possible and identify predictors of depression to improve quality of life in PD patients.

It has been reported that the duration of PD, Hoehn and Yahr phase, age, activities of daily living, low cognitive function, and sleep behavior disorder affect DPD[5,14,15]. Depression frequently occurs during the early stages of PD[16]. Previous studies suspected that PD might cause depression along with dyskinesia due to dopamine

deficiency[6,17]. However, other studies showed that the administration of L-dopa (levodopa), a Parkinson's treatment, did not improve depression symptoms indisputably[18]. Therefore, it cannot be determined that depression is simply due to the effect of dopamine deficiency. On the other hand, many recent scholars have argued that depression in PD patients was caused by the complex interactions of multiple factors, rather than a single cause[19,20]. However, these previous studies[5,16,17] are limited in determining a risk factor, while considering multiple risk factors as each study used different confounding factors or covariates and used regression models to predict a risk factor, although they were effective in exploring individual risk factors. Additionally, regression analysis requires data which satisfy many assumptions such as normality, linearity, and homoscedasticity, and disease data are highly likely to violate these assumptions.

Recent studies have used various machine learning classifiers such as the support vector machine (SVM) and decision tree as statistical classification methods to identify multiple risk factors for diseases such as depression[21]. Of these, the SVM divides data into two groups linearly and explores the optimal boundary. SVM can be used for classifying nonlinear data, it has less probability of overfitting than the decision tree model, and it has high prediction accuracy even for small sample sizes[22,23]. The objectives of present study were to develop a model for predicting DPD based on the SVM, while considering sociodemographic factors, health habits, Parkinson's symptoms, sleep behavior disorders, and neuropsychiatric indicators as predictors and to provide baseline data for identifying DPD.

MATERIALS AND METHODS

Subjects

The present study was conducted by analyzing the Parkinson's Disease Epidemiology (PDE) Data provided by the National Biobank of Korea, the Centers for Disease Control and Prevention (CDC), and Republic of Korea (No. KBN-2019-005). This study was approved by the Research Ethics Review Board of the National Biobank of Korea (No. 2019-005) and the Korea-CDC (No. 2019-1327). The goal of the National Biobank and the structure of the data were described by Lee et al[24]. The PDE data used in this study were collected at 14 university hospitals from January to December, 2015, under the supervision of the Korea CDC. The PDE data consisted of health behaviors, sociodemographic factors, motor characteristics related to PD, disease history, neuropsychological test results and sleep behavior disorders. PD was diagnosed according to the idiopathic Parkinson's disease diagnosis criteria of the United Kingdom Parkinson's disease Society Brain Bank[25]. This study analyzed 223 of 335 patients who were 60 years or older with PD, after excluding 112 subjects who had a least one missing value in the Geriatric Depression Scale (GDS)[26], a depression screening test.

Measurement

Depression was measured using the 30 items of the GDS[26], and the threshold point of depression was 10 points. Explanatory variables included PD-related motor signs (*e.g.*, late motor complications, bradykinesia, tremor, postural instability, and rigidity), rapid eye movement (REM) and sleep behavior disorders, the Korean Mini Mental State Examination score[27], Korean Montreal Cognitive Assessment score[28], global Clinical Dementia Rating score (CDR)[29], Korean Instrumental Activities of Daily Living score (K-IADL)[30], Untitled Parkinson's Disease Rating total score (UPDRS)[31], UPDRS motor score[32], Hoehn and Yahr staging (H and Y staging)[33], and the Schwab and England Activities of Daily Living scale (Schwab and England ADL)[34].

SVM was used to develop a DPD prediction model. SVM is a MLalgorithm that finds an optimal decision boundary, in other words, a linear separation dividing the hyperplane (H-plane) optimally, by converting training data to a higher dimension through nonlinear mapping[35]. For example, A = [b, e] and B = [c, f] are non-linearly (non-lin) separable in two dimensions. When they are mapped in three dimensions, they have linearly separable characteristics. Therefore, when appropriate non-lin mapping is conducted at sufficiently large dimensions, data with two classes can always be separated in H-plane[23].

SVM is very accurate as it can model complex nonlinear decision-making domains and tends to be overfitting less than other models, which is a major advantage of this method[36]. This study used R version 3.6.1 for statistical analyses. The prediction performance (accuracy) of eight SVM models was compared using four algorithms [

i.e., a radial basis function (Gauss function), a linear algorithm, a sigmoid algorithm, and a polynomial algorithm] and two types of SVM [C-SVM (C parameter) and Nu-SVM (Nu parameter)]. The prediction performance of the models was evaluated by considering overall accuracy, sensitivity, and specificity.

RESULTS

General characteristics of the 223 study subjects with PD were analyzed and are shown in Table 1. The mean age of the subjects was 71.7 years (SD = 6.1). The initial age at diagnosis of PD was 70.8 years (SD = 6.3) and the training period was 7.5 years (SD = 5.3). The percentage of non-smokers was 79.7%, right-handed subjects was 96.0%), and subjects without a family history of PD was 82.5%. It was found that 22.5% of the subjects had diabetes, 41.3% had hypertension, and 13.3% had hyperlipidemia. In terms of cognitive characteristics, 30.9% of the subjects had PDD, 61% had Mild Cognitive Impairment in Parkinson's disease (PD-MCI), and 8.1% of them had Parkinson's disease with cognitive impairment (PD-NC). The results of the GDS confirmed that 41.7% of the patients had depression. The distribution of neuropsychological test results is presented using a density plot (Figure 1).

Table 1 shows the general characteristics of the study subjects with depression and the related potential factors (influencing factors) of DPD. The results of the chi-square test showed that PD patients with depression and PD patients without depression were significantly different as assessed by the Korean Mini Mental State Examination, Korean Montreal Cognitive Assessment, Global CDR score, total score of UPDRS, sum of boxes in CDR, K-IADL, motor score of UPDRS, H and Y staging, and Chwab and England ADL ($P < 0.05$).

Comparing the accuracy of the DPD prediction model according to the SVM classification algorithm

The fit of the model varies by the kernel type of SVM. Therefore, our study compared the prediction accuracy of eight SVM models [(C-SVM or Nu-SVM) × (Gaussian kernel, linear, polynomial, or sigmoid algorithm)] to examine the performance of the models according to various kernel types (Table 2). The results of model fitting showed that the Gaussian Kernel-based Nu-SVM had the highest sensitivity (96.0%), specificity (93.3%), and mean overall accuracy (95%). On the other hand, although the polynomial-based C-SVM had the highest sensitivity (100%), it had the lowest specificity (20%) and the lowest mean overall accuracy (70%).

This study determined that the Gaussian algorithm-based Nu-SVM model, which had the highest sensitivity and overall accuracy, was the optimal model for predicting DPD and analyzed the importance of variables. The Gaussian algorithm-based Nu-SVM model utilized 34 support vectors and the "functional weight (importance of variables)" is presented in Table 3. Even though the functional weight of SVM is not a value simply for comparing the magnitudes of variables' influence or ranking the importance of variables, it is possible to compare the influence within the level of factors (*e.g.*, comparing the influence of gender) using it. It is also possible to understand whether the relationship between predictors and outcome variables is a risk factor or a preventive factor. The DPD prediction model revealed that the global CDR score, the sum of boxes in CDR, K-IADL, total UPDRS, motor UPDRS, age (≥ 75 years old), gender (female), education level (high school graduate or above), PD family history, smoking (21-40 packs per year), exposure to pesticides, postural instability, late motor complications (occurrence of levodopa-induced dyskinesia), late motor complications ("levodopa-induced dyskinesia" and occurrence of "wearing OFF"), and REM sleep behavior disorders were risk factors for depression. When the effects of PD motor symptoms were compared using "functional weight", late motor complications (occurrence of levodopa-induced dyskinesia) were the most influential risk factors for DPD.

DISCUSSION

Our study developed a depression prediction model using SVM for PD patients. This study used the hospital registry data and found that 41.7% of PD patients suffered from depression. Although it is difficult to compare directly, the results of this study were similar to the results of previous studies[4-6], showing that one in two PD patients

Table 1 General characteristics of the subjects with depression, *n* (%)

Characteristics	Depression		*P* value
	No (*n* = 130)	Yes (*n* = 93)	
Age			0.918
60-74 yr	83 (58.0)	60 (42.0)	
≥ 75 yr	47 (58.8)	33 (41.3)	
Sex			0.183
Male	62 (63.3)	36 (36.7)	
Female	68 (54.4)	57 (45.6)	
Education level			0.158
Middle school graduate and below	76 (54.7)	63 (45.3)	
High school graduate and above	54 (64.3)	30 (35.7)	
Handness			0.905
Right	125 (58.4)	89 (41.6)	
Left	2 (66.7)	1 (33.3)	
Both hands	2 (50.0)	2 (50.0)	
Family PD history			0.921
No	108 (58.7)	76 (41.3)	
Yes	9 (60.0)	6 (40.0)	
Pack year (smoking)			0.666
1-20 pack year	6 (42.9)	8 (57.1)	
21-40 pack year	5 (62.5)	3 (37.5)	
41-60 pack year	1 (50.0)	1 (50.0)	
61+ pack year	118 (59.3)	81 (40.7)	
Traumatic brain injury			0.852
No	118 (60.8)	76 (39.2)	
Yes	7 (63.6)	4 (36.4)	
Diabetes			0.214
No	103 (60.9)	66 (39.1)	
Yes	25 (51.0)	24 (49.0)	
Hypertension			0.246
No	71 (55.5)	57 (44.5)	
Yes	57 (63.3)	33 (36.7)	
Hyperlipidemia			0.220
No	114 (60.3)	75 (39.7)	
Yes	14 (48.3)	15 (51.7)	
Tremor			0.630
No	17 (54.8)	14 (45.2)	
Yes	107 (59.4)	73 (40.6)	
Rigidity			0.342
No	8 (72.7)	3 (27.3)	
Yes	120 (58.3)	86 (41.7)	
Bradykinesia			0.962

No	7 (58.3)	5 (41.7)	
Yes	121 (59.0)	84 (41.0)	
Postural instability			0.187
No	66 (64.1)	37 (35.9)	
Yes	59 (55.1)	48 (44.9)	
REM and RBD			0.081
No	58 (63.7)	33 (36.3)	
Yes	55 (51.4)	52 (48.6)	
Late motor complications			0.140
Only ON-OFF/Wearing OFF	12 (54.5)	10 (45.5)	
Only levodopa-induced dyskinesia	2 (22.2)	7 (77.8)	
Both ON-OFF/Wearing OFF and levodopa-induced dyskinesia are present	13 (59.1)	9 (40.9)	
Both ON-OFF / Wearing OFF and levodopa-induced dyskinesia are absent	99 (61.1)	63 (38.9)	
K-MMSE, mean ± SD	24.2 ± 4.6	21.5 ± 6.1	< 0.001
K-MoCA, mean ± SD	18.2 ± 6.1	13.9 ± 6.7	< 0.001
Global CDR score, mean ± SD	0.5 ± 0.3	0.8 ± 0.7	< 0.001
Sum of boxes in CDR, mean ± SD	2.3 ± 3.1	3.3 ± 4.0	0.041
K-IADL, mean ± SD	1.3 ± 3.0	3.2 ± 5.6	0.003
Total score of UPDRS, mean ± SD	34.5 ± 17.1	53.2 ± 26.1	< 0.001
Motor score of UPDRS, mean ± SD	22.7 ± 10.9	30.0 ± 13.5	< 0.001
H and Y staging, mean ± SD	2.2 ± 0.6	2.6 ± 0.9	0.002
Schwab and England ADL, mean ± SD	80.1 ± 14.5	67.3 ± 21.5	< 0.001

CDR: Clinical Dementia Rating; K-MMSE: Korean Mini-Mental State Examination; K-MoCA: Korean-Montreal Cognitive Assessment; K-IADL: Korean Instrumental Activities of Daily Living; UPDRS: Untitled Parkinson`s Disease Rating total score; H and Y: Hoehn and Yahr; ADL: Activities of Daily Living.

Table 2 Comparing the accuracy of depression in Parkinson's disease prediction model, %

Type of algorithm	Overall accuracy	Sensitivity	Specificity
C-SVM: Linear	92.5	92.0	93.3
C-SVM: Polynomial	70.0	100	20.0
C-SVM: Gaussian	87.5	96.0	73.3
C-SVM: Sigmoid	85.0	88.0	80.0
Nu-SVM: Linear	90.0	92.0	86.6
Nu-SVM: Polynomial	92.5	92.0	93.3
Nu-SVM: Gaussian	95.0	96.0	93.3
Nu-SVM: Sigmoid	82.5	92.0	66.6

SVM: Support vector machine.

had depression. Despite the frequent occurrence of depression among patients with PD, the Global Parkinson's Disease Survey Steering Committee (2002)[8] reported that only 1% of patients with PD recognized that they had depression. These results imply that, even though PD patients frequently experience depression, it is highly likely that many PD patients, their caregivers, and their medical practitioners do not find

Table 3 Functional weight (importance of variables)	
Variables	**Functional weight**
K-MMSE	-1.454
K-MoCA	-4.499
CDR (Global CDR score)	2.643
CDR (sum of boxes)	2.607
K-IADL	4.919
Total UPDRS	5.302
Motor UPDRS	0.978
H and Y staging	-2.190
Schwab and England ADL	-5.055
Age (60-74)	-3.638
Age (75+)	3.638
Gender (female)	0.345
Education (high school graduate and above)	1.150
Family PD history (yes)	0.232
Pack year (21-40)	0.032
Pack year (41-60)	0.000
Pack year (61+)	-1.032
Coffee (yes)	1.006
Pesticide exposure (currently not exposed but exposed previously)	-1.000
Pesticide exposure (currently exposed to pesticide)	2.264
Tremor (yes)	0.000
Rigidity (yes)	-1.000
Bradykinesia (yes)	0.675
Postural instability (yes)	1.117
LMC (levodopa-induced dyskinesia)	3.000
LMC (Both Wearing OFF and levodopa-induced dyskinesia are present)	0.264
LMC (Both Wearing OFF and levodopa-induced dyskinesia are absent)	-4.715
REM sleep behavior disorders	1.261
Number of support vectors: 34	

CDR: Clinical Dementia Rating; K-MMSE: Korean Mini-Mental State Examination; K-MoCA: Korean-Montreal Cognitive Assessment; K-IADL: Korean Instrumental Activities of Daily Living; UPDRS: Untitled Parkinson`s Disease Rating total score; H and Y: Hoehn and Yahr; ADL: Activities of Daily Living; REM: Rapid eye movement; PD: Parkinson's disease.

depression symptoms in the patients or treat them as symptoms due to aging, and, consequently, the patients do not receive appropriate evaluation or treatment. Therefore, it is necessary to develop an education program in order to manage depression from the onset of PD so that PD patients and their caregivers can receive the correct information regarding depression and receive effective treatments. The results of this study showed that the global CDR score, the sum of boxes in CDR, K-IADL, total UPDRS, motor UPDRS, age (≥ 75 years old), gender (female), education level (≥ high school graduate), PD family history, smoking (21-40 packs per year), exposure to pesticides, postural instability, late motor complications (occurrence of levodopa-induced dyskinesia), late motor complications (occurrence of "wearing OFF" and "levodopa-induced dyskinesia"), and REM sleep behavior disorders were the major predictors of DPD. A number of studies[5,14,15] exploring the risk factors for DPD reported that daily living ability, sleep behavior disorders, cognitive level, and Hoehn
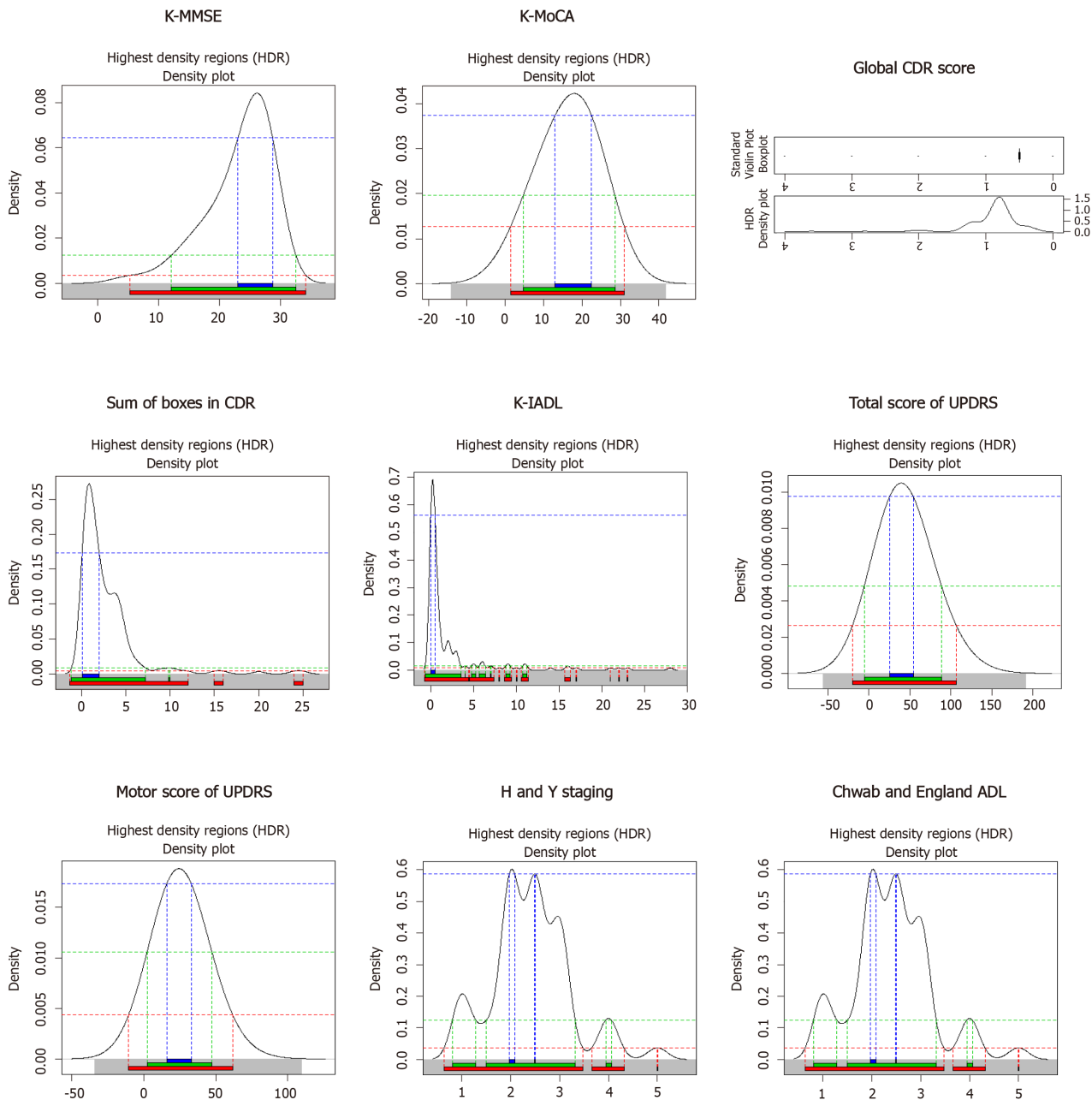
Figure 1 The distribution of neuropsychological tests. CDR: Clinical Dementia Rating; K-MMSE: Korean Mini-Mental State Examination; K-MoCA: Korean-Montreal Cognitive Assessment; K-IADL: Korean Instrumental Activities of Daily Living; UPDRS: Untitled Parkinson`s Disease Rating total score; H and Y: Hoehn and Yahr; ADL: Activities of Daily Living.

and Yahr[33] stages, as well as environmental factors (*e.g.*, social stigma and social participation), were key influencing factors of depression and these results supported the results of our study. In particular, sleep behavior disorder is known to be the most representative factor for predicting the risk of depression[37] and has a high correlation with depression[38]. It is known that sleep behavior disorders are caused by dysfunction in the autonomic nervous system of patients with PD[39]. Bae *et al*[40] evaluated factors related to DPD using structural equation modeling, and reported that a sleeping issue best predicted DPD.

ADL, indicating the limitation of physical functions, is also related to early depression[41], as social relations can decrease due to limited ADL resulting in depression[42]. If depression persists, the quality of life in PD patients is highly likely to decrease[43] and they have a high risk of suicide[10]. Therefore, it is necessary to develop customized screening tools that can detect high-risk groups sensitively and conduct continuous monitoring in order to prevent depression and maintain emotional health.

Another finding of our study was that the prediction accuracy of the Gaussian kernel-based Nu-SVM was the highest when comparing eight SVM classification

algorithms [(C-SVM or Nu-SVM) × (Gaussian kernel, linear, polynomial, or sigmoid algorithm)]. The performance of non-lin SVM is affected by the kernel function and the parameters constituting it[44]. Of these, the Gaussian kernel is an algorithm that maps in a specific space with infinite dimension and it had high predictive accuracy in a previous study[25]. The results of this study suggest that, in the case of binary disease data with a small sample size, developing a prediction model using Gaussian kernel-based Nu-SVM will have higher predictive accuracy than SVM models based on other algorithms.

The present study was meaningful because it developed an SVM-based DPD prediction model using national Parkinson's registry data while considering sociodemographic factors, health habits, Parkinson's symptoms, and sleep behavior disorders as predictors in addition to neuropsychological indicators. The limitations of this study are as follows. First, it is difficult to generalize the results of this study as hospital registry data collected using convenience sampling was used. Future studies are needed to apply systematic sampling at the stage of recruiting subjects to minimize selection bias. Secondly, the sample size was small. Thirdly, causality could not be identified because it was a cross-sectional study. Additional longitudinal studies are required to prove causality. Fourthly, biomarkers and Parkinson's treatments related to depression were not investigated. In order to predict depression more sensitively, it is necessary to develop a predictive model including biomarkers in addition to cognitive tests and neuropsychological tests.

CONCLUSION

The results of our study can be used as baseline information to prevent DPD and establish management strategies. It is necessary to develop customized screening that can detect DPD in the early stage and continuously monitor high-risk groups based on the factors related to DPD derived from this predictive model in order to maintain the emotional health of PD patients. It is also necessary to develop customized programs for managing depression from the onset of PD.

ARTICLE HIGHLIGHTS

Research background
It is important to diagnose depression in Parkinson's disease (PD) patients as soon as possible and identify predictors of depression to improve the quality of life in PD patients.

Research motivation
It has been reported that the duration of PD, Hoehn and Yahr phase, age, activities of daily living, low cognitive function, and sleep behavior disorder affect depression in PD. However, these previous studies are limited in determining a risk factor, while considering multiple risk factors as each study used different confounding factors or covariates and used regression models to predict a risk factor, although they were effective in exploring individual risk factors.

Research objectives
The objectives of our study were to develop a model for predicting depression in Parkinson's disease (DPD) based on the support vector machine while considering sociodemographic factors, health habits, Parkinson's symptoms, sleep behavior disorders, and neuropsychiatric indicators as predictors and to provide baseline data for identifying DPD.

Research methods
The data used in this study was collected at 14 university hospitals from January to December, 2015, under the supervision of the Korea Centers for Disease Control. The data consisted of health behaviors, sociodemographic factors, motor characteristics related to PD, disease history, neuropsychological test results and sleep behavior disorders.

Research results

When the effects of PD motor symptoms were compared using "functional weight", late motor complications (occurrence of levodopa-induced dyskinesia) were the most influential risk factors for DPD.

Research conclusions

It is necessary to develop customized screening that can detect DPD in the early stage and continuously monitor high-risk groups based on the factors related to DPD derived from this predictive model in order to maintain the emotional health of PD patients.

Research perspectives

It is also necessary to develop customized programs for managing depression from the onset of PD.

ACKNOWLEDGEMENTS

REFERENCES

1   **Health Insurance Research Assessment Service**.   (2019, September 29). Ministry for health, welfare and family affairs. Available from: http://www.mohw.go.kr

2   **Sagna A**, Gallo JJ, Pontone GM. Systematic review of factors associated with depression and anxiety disorders among older adults with Parkinson's disease. *Parkinsonism Relat Disord* 2014; **20**: 708-715 [PMID: 24780824 DOI: 10.1016/j.parkreldis.2014.03.020]

3   **Aarsland D**, Påhlhagen S, Ballard CG, Ehrt U, Svenningsson P. Depression in Parkinson disease--epidemiology, mechanisms and management. *Nat Rev Neurol* 2011; **8**: 35-47 [PMID: 22198405 DOI: 10.1038/nrneurol.2011.189]

4   **Cummings JL**. Depression and Parkinson's disease: a review. *Am J Psychiatry* 1992; **149**: 443-454 [PMID: 1372794 DOI: 10.1176/ajp.149.4.443]

5   **Wichowicz HM**, Sławek J, Derejko M, Cubała WJ. Factors associated with depression in Parkinson's disease: a cross-sectional study in a Polish population. *Eur Psychiatry* 2006; **21**: 516-520 [PMID: 16531018 DOI: 10.1016/j.eurpsy.2006.01.012]

6   **Ranoux D**. [Depression and Parkinson disease]. *Encephale* 2000; **26** Spec No 3: 22-26 [PMID: 10858928]

7   **Tandberg E**, Larsen JP, Aarsland D, Cummings JL. The occurrence of depression in Parkinson's disease. *Arch Neurol* 1996; **53**: 175-179 [PMID: 8639068 DOI: 10.1001/archneur.1996.00550020087019]

8   **Global Parkinson's Disease Survey (GPDS) Steering Committee**. Factors impacting on quality of life in Parkinson's disease: results from an international survey. *Mov Disord* 2002; **17**: 60-67 [PMID: 11835440 DOI: 10.1002/mds.10010]

9   **Richard IH**, Schiffer RB, Kurlan R. Anxiety and Parkinson's disease. *J Neuropsychiatry Clin Neurosci* 1996; **8**: 383-392 [PMID: 9116473 DOI: 10.1176/jnp.8.4.383]

10   **Lee T**, Lee HB, Ahn MH, Kim J, Kim MS, Chung SJ, Hong JP. Increased suicide risk and clinical correlates of suicide among patients with Parkinson's disease. *Parkinsonism Relat Disord* 2016; **32**: 102-107 [PMID: 27637284 DOI: 10.1016/j.parkreldis.2016.09.006]

11   **McDonald WM**, Richard IH, DeLong MR. Prevalence, etiology, and treatment of depression in Parkinson's disease. *Biol Psychiatry* 2003; **54**: 363-375 [PMID: 12893111 DOI: 10.1016/s0006-3223(03)00530-4]

12   **Huse DM**, Schulman K, Orsini L, Castelli-Haley J, Kennedy S, Lenhart G. Burden of illness in Parkinson's disease. *Mov Disord* 2005; **20**: 1449-1454 [PMID: 16007641 DOI: 10.1002/mds.20609]

13   **Johnson SJ**, Kaltenboeck A, Diener M, Birnbaum HG, Grubb E, Castelli-Haley J, Siderowf AD. Costs of Parkinson's disease in a privately insured population. *Pharmacoeconomics* 2013; **31**: 799-806 [PMID: 23907717 DOI: 10.1007/s40273-013-0075-0]

14   **Havlikova E**, van Dijk JP, Nagyova I, Rosenberger J, Middel B, Dubayova T, Gdovinova Z, Groothoff JW. The impact of sleep and mood disorders on quality of life in Parkinson's disease patients. *J Neurol* 2011; **258**: 2222-2229 [PMID: 21614432 DOI: 10.1007/s00415-011-6098-6]

15   **Lawrence BJ**, Gasson N, Kane R, Bucks RS, Loftus AM. Activities of daily living, depression, and quality of life in Parkinson's disease. *PLoS One* 2014; **9**: e102294 [PMID: 25025280 DOI: 10.1371/journal.pone.0102294]

16   **Brooks DJ**, Doder M. Depression in Parkinson's disease. *Curr Opin Neurol* 2001; **14**: 465-470 [PMID: 11470962 DOI: 10.1097/00019052-200108000-00006]

17   **Eskow Jaunarajs KL**, Angoa-Perez M, Kuhn DM, Bishop C. Potential mechanisms underlying anxiety and depression in Parkinson's disease: consequences of l-DOPA treatment. *Neurosci Biobehav Rev* 2011; **35**: 556-564 [PMID: 20615430 DOI: 10.1016/j.neubiorev.2010.06.007]

18   **Nègre-Pagès L**, Grandjean H, Lapeyre-Mestre M, Montastruc JL, Fourrier A, Lépine JP, Rascol O; DoPaMiP Study Group. Anxious and depressive symptoms in Parkinson's disease: the French cross-sectionnal DoPaMiP study. *Mov Disord* 2010; **25**: 157-166 [PMID: 19950403 DOI: 10.1002/mds.22760]

19   **Goodarzi Z**, Mele B, Guo S, Hanson H, Jette N, Patten S, Pringsheim T, Holroyd-Leduc J. Guidelines for dementia or Parkinson's disease with depression or anxiety: a systematic review. *BMC Neurol* 2016; **16**: 244 [PMID: 27887589 DOI: 10.1186/s12883-016-0754-5]

20   **Reijnders JS**, Ehrt U, Weber WE, Aarsland D, Leentjens AF. A systematic review of prevalence studies of depression in Parkinson's disease. *Mov Disord* 2008; **23**: 183-9; quiz 313 [PMID: 17987654 DOI: 10.1002/mds.21803]

21   **Byeon H**. Development of Depression Prediction Models for Caregivers of Patients with Dementia Using Decision Tree Learning Algorithm. *Int J Gerontol* 2019; **13**: 314-319 [DOI: 10.6890/IJGE.201912_13(4).0010]

22   **Lin HT**, Lin CJ. A study on sigmoid kernels for SVM and the training of non-PSD kernels by SMO-type methods. *Neural Comput* 2003; **3**: 1-32

23   **Byeon H**. Predicting the Swallow-Related Quality of Life of the Elderly Living in a Local Community Using Support Vector Machine. *Int J Environ Res Public Health* 2019; **16** [PMID: 31684165 DOI: 10.3390/ijerph16214269]

24   **Lee JE**, Kim JH, Hong EJ, Yoo HS, Nam HY, Park O. National Biobank of Korea: Quality control Programs of Collected-human Biospecimens. *Osong Public Health Res Perspect* 2012; **3**: 185-189 [PMID: 24159512 DOI: 10.1016/j.phrp.2012.07.007]

25   **Hughes AJ**, Daniel SE, Kilford L, Lees AJ. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *J Neurol Neurosurg Psychiatry* 1992; **55**: 181-184 [PMID: 1564476 DOI: 10.1136/jnnp.55.3.181]

26   **Cho MJ**, Bae JN, Suh GH, Hahm BJ, Kim JK, Lee DW, Kang MH. Validation of geriatric depression scale, Korean version (GDS) in the assessment of DSM-III-R major depression. *J Korean Neuropsychiatr Assoc* 1999; **38**: 48-63

27   **Kang Y**, Na DL, Hahn S. A validity study on the Korean Mini-Mental State Examination (K-MMSE) in dementia patients. *J Korean Neurol Assoc* 1997; **15**: 300-308

28   **Kang Y**, Park J, Yu KH, Lee BC. The validity of the Korean-Montreal Cognitive Assessment (K-MoCA) as a screening test for both MCI and VCI. *The frontal lobes* 2010 [DOI: 10.3389/conf.fnins.2010.14.00148]

29   **Choi SH**, Na DL, Lee BH, Hahm DS, Jeong JH, Yoon SJ, Yoo KH, Ha CK, Han IW. Estimating the validity of the Korean version of expanded clinical dementia rating (CDR) scale. *J Korean Neurol Assoc* 2001; **19**: 585-591

30   **Kang SJ**, Choi SH, Lee BH, Kwon JC, Na DL, Han SH. The reliability and validity of the Korean Instrumental Activities of Daily Living (K-IADL). *J Korean Neurol Assoc* 2002; **20:** : 8-14

31   **Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease**. The Unified Parkinson's Disease Rating Scale (UPDRS): status and recommendations. *Mov Disord* 2003; **18**: 738-750 [PMID: 12815652 DOI: 10.1002/mds.10473]

32   **Richards M**, Marder K, Cote L, Mayeux R. Interrater reliability of the Unified Parkinson's Disease Rating Scale motor examination. *Mov Disord* 1994; **9**: 89-91 [PMID: 8139610 DOI: 10.1002/mds.870090114]

33   **Hoehn MM**, Yahr MD. Parkinsonism: onset, progression and mortality. *Neurology* 1967; **17**: 427-442 [PMID: 6067254 DOI: 10.1212/wnl.17.5.427]

34   **Gillingham FJ**, Donaldson MC. Schwab and England activities of daily living. In: Third symposium of Parkinson's disease. ES Livingstone: Edinburgh, 1969

35   **Cuentas S**, Peñabaena-Niebles R, Garcia E. Support vector machine in statistical process monitoring: a methodological and analytical review. *Int J Adv Manuf Technol* 2017; **91**: 485-500 [DOI: 10.1007/s00170-016-9693-y]

36   **Furey TS**, Cristianini N, Duffy N, Bednarski DW, Schummer M, Haussler D. Support vector machine classification and validation of cancer tissue samples using microarray expression data. *Bioinformatics* 2000; **16**: 906-914 [PMID: 11120680 DOI: 10.1093/bioinformatics/16.10.906]

37   **Manni R**, Sinforiani E, Zucchella C, Terzaghi M, Rezzani C. A sleep continuity scale in Alzheimer's disease: validation and relationship with cognitive and functional deterioration. *Neurol Sci* 2013; **34**: 701-705 [PMID: 22622870 DOI: 10.1007/s10072-012-1118-6]

38   **Happe S**, Schrödl B, Faltl M, Müller C, Auff E, Zeitlhofer J. Sleep disorders and depression in patients with Parkinson's disease. *Acta Neurol Scand* 2001; **104**: 275-280 [PMID: 11696021 DOI: 10.1034/j.1600-0404.2001.00024.x]

39   **Yoritaka A**, Ohizumi H, Tanaka S, Hattori N. Parkinson's disease with and without REM sleep behaviour disorder: are there any clinical differences? *Eur Neurol* 2009; **61**: 164-170 [PMID: 19129703 DOI: 10.1159/000189269]

40   **Bae ES**, Chun SM, Kim JW, Kang CW. A prediction model for depression in patients with Parkinson's disease. *Korean J Health Educ Promot* 2013; **30:** : 139-151 [DOI: 10.14367/kjhep.2013.30.5.139]

41   **Yang Y**, George LK. Functional disability, disability transitions, and depressive symptoms in late life. *J Aging Health* 2005; **17**: 263-292 [PMID: 15857959 DOI: 10.1177/0898264305276295]

42   **Valderas JM**, Alonso J. Patient reported outcome measures: a model-based classification system for research and clinical practice. *Qual Life Res* 2008; **17**: 1125-1135 [PMID: 18836850 DOI: 10.1007/s11136-008-9396-4]

43   **Fan JY**, Chang BL, Wu YR. Relationships among Depression, Anxiety, Sleep, and Quality of Life in Patients with Parkinson's Disease in Taiwan. *Parkinsons Dis* 2016; **2016**: 4040185 [PMID: 27293956 DOI: 10.1155/2016/4040185]

44   **Steinwart I**, Christmann A. Support vector machines, Springer Science Business Media: Manhattan, New York, 2008