원 저

https://doi.org/10.17340/jkna.2021.4.7

라쉬 측정 이론에 의한 파킨슨병 환자 대상 한국어판 병원불안-우울척도의 신뢰도 및 타당도 평가

최진혁 전성진 홍슬기 김아로 박지윤 양희준

울산대학교 의과대학 울산대학교병원 신경과

The Reliability and Validity of the Korean Version of Hospital Anxiety and Depression Scale Using Rasch Measurement Theory in Patients with Parkinson's Disease

Jin-Hyuk Choi, MD, Seongjin Jeon, MD, Seulgi Hong, MD, Ahro Kim, MD, Ji-Yun Park, MD, Hui-Jun Yang, MD

Department of Neurology, Ulsan University Hospital, University of Ulsan College of Medicine, Ulsan, Korea

Background: Depression and anxiety are prevalent and can cause suffering in patients with Parkinson's disease (PD). The Korean version of the Hospital Anxiety and Depression Scale (K-HADS) has been widely used to assess depression and anxiety symptoms in Korean patient with PD. The present study aimed to assess the reliability and validity of the K-HADS using Rasch measurement analysis.

Methods: A total of 106 PD patients (54 males, 52 females) who met the diagnostic criteria of the United Kingdom Brain Bank were recruited. Unidimensionality, the Rasch model fit, response category functioning, patient-item distribution, and the separation reliability of the K-HADS depression (K-HADS-D) and anxiety (K-HADS-A) subscales were statistically evaluated.

Results: The mean K-HADS-D and K-HADS-A scores were 8.08±4.69 (mean±standard deviation) and 5.44±4.18, respectively. Cronbach's α coefficients of the K-HADS-D and K-HADS-A were 0.82 and 0.83. The Rasch analysis revealed that the K-HADS-D and K-HADS-A showed unidimensionality and no disordered functioning was observed in the 4-point polytomous scale. However, both K-HADS-D and K-HADS-A exhibited suboptimal separation reliability, while the K-HADS-A showed inadequate scale targeting with floor effect.

Conclusions: The present study comprises the first validation of the K-HADS using the Rasch measurement model, suggesting that the K-HADS-D and K-HADS-A are clinimetrically acceptable and reliable scales for use in Korean patients with PD. However, the moderate person separation indices implicate the relatively low discriminatory ability of the K-HADS in our study patients.

J Korean Neurol Assoc 39(4):312-321, 2021

Key Words: Parkinson's disease, Rasch analysis, The Hospital Anxiety and Depression Scale, Validity, Reliability

서 론

특발성 파킨슨병(idiopathic Parkinson disease)은 명확한 발병 원인이 없이 진행성으로 나타나는 신경퇴행성 질환으로, 서동, 강직 및 안정 시 진전 등이 잘 알려진 파킨슨병의 주요 운동 증상들(cardinal motor symptoms)이다. 다수의 파킨슨병 환자에서는 운동 증상 외에도 우울 및 불안장애 등의 정동장애, 파킨슨치매, 환각, 환시 등의 인지장애, 후각저하, 수면이상, 자율신경장애 등의 비운동 증상들(non-motor symptom)이 광범위하게 동반된다.[1-4] 특히 우울 증상 및 불안 증상은 파킨슨병 환자에서 가장 흔한 정동장애로서 연구 집단 및 방법에 따라 다소 상이하지만 대체로 40% 이상의 파킨슨병 환자에서 우울 증상을 호소하며,[5-7] 불안 증상의

Received May 6, 2021 **Revised** September 23, 2021
Accepted September 23, 2021

Address for correspondence: Hui-Jun Yang, MD
Department of Neurology, Ulsan University Hospital, 877 Bangeojinsunhwando-ro, Dong-gu, Ulsan 44033, Korea
Tel: +82-52-250-8637 Fax: +82-52-250-7088
E-mail: yanghuijun@uuh.ulsan.kr

경우에는 약 12-57%의 환자에서 보고되고 있다.[8-10] 이러한 비율은 일반인구집단에 비해 현저히 높은 것으로, 파킨슨병에서의 우울 및 불안장애는 건강 관련 삶의 질에 부정적 영향을 미치는 것으로 알려져 있다.[2-4] 이에 2007년 및 2008년 국제파킨슨병및이상운동질환학회(International Parkinson and Movement Disorder Society) 산하 전문위원회(task force)에서는 파킨슨병에서 우울 및 불안을 평가하는 다양한 임상평정척도들의 신뢰도 및 타당도를 고전적 검사이론(classical test theory)에 입각해 비교 분석하여 보고한 바 있다.[5,8]

최근 고전적 검사이론의 여러 한계점들이 알려지면서,[11-14] 2013년 미국 국립보건원(National Institute of Health)이 주도하는 환자-자기평가 결과 측정정보체계(Patient-Reported Outcomes Measurement Information System) 및 2010년 유럽 주도의 건강측정도구에 대한 컨센서스 표준(COnsensus-based Standards for the selection of health Measurement INstruments) 등에서는 임상평정척도의 평가 시 고전적 검사이론에 따른 분석과 함께 문항반응이론(item response theory; IRT)에 의거한 추가적 보완을 권고하였다.[12,15] 문항반응이론의 분석 모형은 각각 라쉬 모형(Rasch model), 2-모수 IRT 모형(2-parameter IRT model) 및 3-모수 IRT 모형(3-parameter IRT model) 등으로 나뉘는데, 이 중 단일 모수(1-parameter)를 채용하고 있는 라쉬 모형은 연구에 필요한 표본 크기(sample size)가 상대적으로 적으면서 분석 시 요구되는 전제 조건이 단순하여 실제 임상현장에서 신뢰도 및 타당도에 대한 연구 시에 폭넓게 채택되고 있으며,[13,16,17] 파킨슨병과 관련하여서는 파킨슨병 성과척도-자율신경(Scales for Outcomes in Parkinson's disease-Autonomic),[18] 몬트리올 인지 평가(Montreal Cognitive Assessment),[19] 상태-특성불안검사(State-Trait Anxiety Inventory),[20] 파킨슨병이상운동척도(Parkinson Disease Dyskinesia Scale)[21] 등에서 라쉬 모형을 비롯한 문항반응이론에 의한 신뢰도 및 타당도 분석이 보고된 바 있다.

총 14개 항목으로 이뤄진 병원불안-우울척도(The Hospital Anxiety-Depression Scale, HADS)는 1983년 Zigmond와 Snaith에 의해 기존 신체적 질병(medical illness)이 있는 환자들이 병원을 방문하였을 때 보이는 불안 증상 및 우울 증상을 측정하기 위해 개발되었으며,[22] 국내에서는 Oh 등[23]이 체계적 번역 절차를 거쳐 한국어판 병원불안-우울척도(the Korean version of the Hospital Anxiety and Depression Scale, K-HADS)를 개발하고 정신건강의학과에서 우울 및 불안장애를 진단받은 환자군을 대상으로 고전적 검사이론에 기반한 신뢰도 및 타당도를 보고하였다. HADS는 기존의 임상평정척도들에 비해 비교적 문항 수가 적어 임상 현장에서 시행하기 용이하고 우울 및 불안을 동시에 측정할 수 있으며 특히 파킨슨병의 운동 증상과 유사한 신체적 증상(somatic symptom)에 대한 비중이 낮은 장점이 있다.[7-9] 해외에서는 Forjaz 등[16]이 라쉬 모형에 기반하여 파킨슨병 환자에서 HADS의 임상측정학적 특성을 분석한 바 있으나, 국내 파킨슨병 환자를 대상으로 하여 문항반응이론에 따라 K-HADS의 임상측정학적 특성을 조사한 연구는 현재까지 알려져 있지 않다. 이에 본 연구에서는 파킨슨병 환자에서 우울 및 불안 증상 평가를 위해 국내에서 널리 사용되고 있는 K-HADS에 대해 라쉬 모형에 의한 평가를 시행하고자 한다.

대상과 방법

1. 대상

2014년 2월부터 2020년 1월까지 울산대학교병원 신경과 외래에 내원한 환자 중 영국파킨슨병학회 뇌은행 진단 기준(UK Parkinson's Diseases Society Brain Bank Criteria)을 만족하는 특발성 파킨슨병 환자를 대상으로 하였다. 본 연구에 대해 울산대학교병원 임상연구심의위원회의 승인을 받고 고지된 동의(informed consent)를 수행하였으며, 연구의 절차 및 내용을 잘 이해하고 자발적으로 동의한 환자가 연구에 참여하였다. 단, 파킨슨병의 진단 시 40세 미만이거나, 레보도파 치료에 효과가 없는 파킨슨증후군 또는 이차성 파킨슨증 환자, 정신질환 진단 및 통계 편람(Diagnostic and Statistical Manual of Mental Disorders) 5판에 따라 치매가 진단된 경우 및 파킨슨병의 비운동 증상에 영향을 미칠 수 있는 중증의 내과적 질환이나 전신 상태로 인하여 연구를 진행하기 어려운 경우는 연구 대상에서 제외하였다.

2. 방법 및 도구

대상 환자들에 대하여 병력 청취 및 신경학적 검진을 통해 성별, 연령 및 파킨슨병 이환 기간, 복용 약물을 조사하였다. 파킨슨병 투여 약제에 대해서는 기존 연구에 따라 일일 레보도파 동등 약효량(levodopa equivalent daily dose)을 환산하였으며,[24] 항우울제 및 항불안제의 약물 투여 여부를 의무기록을 통하여 확인하였다. 운동 증상 및 비운동 증상의 측정에 대해서는 전반적인 질병의 정도는 호엔야척도로 평가하였으며 환자의 운동 증상 수준은 통합형 파킨슨병척도검사의 운동 부문 점수(Unified Parkinson's Disease Rating Scale motor part score)로 확인하였다. 비운동 증상 전반의 평가에는 총 30항목으로 구성된 비운동 증상 설문지(Non-Motor Symptoms Questionnaire, NMSQ)를 사용하였다.[1,4] NMSQ는 최

대 30점까지 측정될 수 있는 NMSQ 총점과 함께, 위장관계 증상, 배뇨 증상, 무의지/집중/기억장애, 환각/망상, 우울/불안, 성기능장애, 심혈관계 증상, 수면장애 및 기타 등의 9개 하위 범주의 점수를 각각 확인하였으며, 환자의 건강관련 삶의 질은 8개의 항목으로 구성된 파킨슨병 삶의 질 도구(the Short-form 8-item Parkinson's Disease Questionnaire)로 측정하였다.[2-4]

환자의 우울 및 불안 증상에 대해서는 위의 NMSQ와 더불어 K-HADS로 세분화하여 살펴보았다. K-HADS는 각각 7개의 우울 증상을 측정하는 짝수 문항인 우울 하위척도(the K-HADS depression subscale, K-HADS-D) 및 7개의 불안 증상을 평가하는 홀수 문항인 불안 하위척도(the K-HADS anxiety subscale, K-HADS-A)로 구성되어 있으며 점수가 높을수록 환자의 우울 또는 불안 증상 수준이 높다고 판단할 수 있다.[7,10] 문항별로 증상의 심한 정도를 0에서 3까지의 다분 리커트 척도로 측정하므로 각 하위척도의 점수의 범위는 최저 0점에서 최고 21점까지 계산된다. HADS 개발 시 Zigmond와 Snaith는 11점 이상을 우울 또는 불안의 증례 수준(case levels of depression or anxiety)에 준하는 것으로 제시하였고 8-10점을 경계치로, 7점 이하를 정상 대조로 간주한 바 있다.[22] 최근의 연구에 따르면 Mondolo 등[25]은 파킨슨병 환자를 대상으로 우울장애에 대한 HADS-D의 최적화 경계치를 11점 이상으로, 경도-중등도 이상의 불안장애에 대한 HADS-A의 경계치를 8점 이상으로 보고하였다.[10] 이에 본 연구에서도 K-HADS-D에 대해서는 11점을 절단점으로 하고 K-HADS-A에 대해서는 8점을 절단점으로 하였다.

3. 통계 분석과 자료 처리

연구 대상자의 인구학적 및 임상적 특성 평가는 SPSS version 21.0 (IBM Corporation, Somers, NY, USA) 통계 프로그램으로 빈도와 백분율 및 평균과 표준편차의 기술적 통계 분석을 시행하였다. 문항반응이론에 의한 타당도 및 신뢰도 분석은 Winsteps version 4.1.0 (Winsteps Inc., Chicago, IL, USA) 통계 프로그램으로 진행하였으며 절차는 다음과 같다.[17,19,20]

먼저 단일차원성(unidimensionality)이 성립하는지 확인하기 위하여 7개 문항으로 이뤄진 K-HADS-D 및 K-HADS-A 하위척도 각각에 대하여 잔차 주성분 분석(the principal component analysis of residual) 및 문항극성(item polarity) 분석을 시행하였다. 잔차 주성분 분석에서 라쉬 측정으로 예측되는 분산(variance explained by Rasch measure)이 40% 이상이며, 라쉬 측정으로 예측되지 않는 첫 번째 및 두 번째 잔차 분산의 고유값(eigenvalue)이 3.0 미만일 경우 그리고 문항 극성(item polarity) 분석에서 점-측정 상관계수(point measure correlation)가 0.30 이상일 경우 해당 임상평정척도가 단일차원성을 만족한다고 간주하였다.[11,19,26,27] 다음으로 문항 별 외적합 및 내적합 평균자승잔차로 개별 문항들이 라쉬 모형에 적합(Rasch model fit)한지 측정하였다. 외적합(outfit)은 환자의 전반적인 불안 또는 우울 증상 정도에 비하여 증상 수준이 낮은 문항을 시인(endorsement)하지 않거나 반대로 통상적으로 증상 수준이 높은 문항을 시인하는 등의 비일치 반응을 반영하는 지표로서 환자의 증상 정도에 비해 극단치(outlier)를 보이는 정도를 나타낸다. 내적합(infit)은 전술한 외적합 지표가 소수의 극단치에 따라 전반적인 적합도가 부적합을 보이는 가능성을 보정하고자 환자의 증상 정도에 인접한 비일치 반응에 대해 가중치를 둔 것이며, 이에 황자의 전반적인 증상 정도에 많이 떨어지지 않은 문항의 비일치 반응들을 좀 더 예민하게 반영하며, 외적합 또는 내적합의 평균자승잔차(mean square residual, MnSq)가 0.5 미만이거나 1.5를 초과할 경우 부적합(misfit)으로 해석하였다.[17,20,28,29] 평정척도 표적(scale targeting) 및 문항 난이도(item difficulty)는 원 점수를 로짓(logit)으로 변환하여 등간 척도로 도시한 환자-문항 분포도(patient-item distribution plot)를 사용하여 좌편의 환자 분포 및 우편의 문항 난이도 분포를 도시하고, 환자 및 문항 간 로짓 평균(mean logit)의 차이를 분석하였다.[13,14] 난이도의 의미는 해당 우울 또는 불안 증상 수준을 질의하는 문항에 대해 환자가 시인(endorsement)할 확률로서, 예컨대 난이도가 낮은 문항에서는 보다 많은 환자들이 해당 문항에서 질의하는 증상이 본인에게 있다고 시인하는 반면, 난이도가 높은 문항은 우울 또는 불안 증상 정도가 심한 환자들 만이 해당 문항 내용을 시인하게 된다.[11,29]

K-HADS-D 및 K-HADS-A의 내적 신뢰도는 대상자 분리 신뢰도(person separation reliability) 및 문항 분리 신뢰도(item separation reliability)를 통해 평가하였다. 대상자 분리 신뢰도는 측정하는 증상의 심한 정도에 따라 연구 대상 환자들이 서로 분리되는 구성집단(strata)들로 세분할 수 있는지 확인하는 지표이며, 문항 분리 신뢰도는 평정척도내 각 문항들이 예를 들어 경증-중등도-중증 등의 위계에 따라 나눌 수 있는지 나타낸다. 각각 대상자 분리 신뢰도 및 문항 분리 신뢰도가 0.60 이상일 경우에는 수용 가능(acceptable)한 것으로, 0.80 이상인 경우에는 우수한 것으로 간주할 수 있다.[11,19,20] K-HADS의 4점 리커트 척도에 대한 응답 범주(response category) 분석은 라쉬 모형의 다범주 확장인 앤드리치의 평정척도 모형(Andrich's rating scale model)에 따라 우선 각 범주별로 10개 이상의 관찰이 있고, 응답 범주 확률 곡선이 위계적으로 배치되어 범주들의 평균 측정치가 단조적으로 증가하며, 각각의 응답 범주별 외적합 평균자승잔차가 2.0 미만일 경우 수용

가능하다고 판정하였다.[30,31]

전술한 문항반응이론에 의한 신뢰도 및 타당도 분석과 함께 고전적 검사이론에 따른 분석을 병행하였다 고전적 검사이론에 의한 한국어판 병원불안-우울척도의 신뢰도는 문항-전체 상관(item-total correlation) 및 크론바흐 alpha 계수(Cronbach's α coefficient)를 통해 확인하였으며, 각각 문항-전체 상관계수가 0.30 이상일 경우와 크론바흐 alpha 계수가 0.70 이상일 경우에 유의한 것으로 간주하였다.[5-8] 수렴 타당도(convergent validity)를 확인하기 위해서 NMSQ의 16번 문항(우울) 및 17번(불안) 응답 결과에 따라 K-HADS-D 및 K-HADS-A 점수 수준 간에 의미 있는 차이가 있는지 Mann-Whitney 검정을 통해 확인하였으며 $p<0.05$ 수준일 경우 통계적으로 유의한 것으로 판단하였다.[1,4]

결 과

본 연구의 대상자는 총 106명(남자 54명, 여자 52명)이었으며, 인구학적 및 임상적 특징에 대해서는 Table 1에 정리하였다. 환자들의 평균 연령은 64.52±8.91세(평균±표준편차)였고 평균 유병 기간은 1.22±1.11년이었으며 평균 Hoehn-Yahr척도는 2.17±0.41이었다. 평균 levodopa equivalent daily dose는 152.5±204.8 mg이었으며, 전체 대상 환자 중 파킨슨병의 진단은 받았으나 관련 약물 치료를 받지 않은 신규 파킨슨병(*de novo* Parkinson's disease) 환자는 56명(52.8%)이었다. 파킨슨병 운동 증상 및 비운동 증상을 살펴보면, Unified Parkinson's Disease Rating Scale 운동 증상 부문의 평균 점수는 19.13±8.51이었고 비운동 증상의 정도에 대한 NMSQ 평균 점수는 10.00±6.59였다. 우울 및 불안 증상에 대해서는 K-HADS-D와 K-HADS-A의 평균 점수는 각각 8.08±4.69 및 5.44±4.18이었으며, 척도별 절단점에 따라 환자들을 구분하였을 때, K-HADS-D의 경우 총 33명(31.1%)에서 11점 이상으로 평가되었고 K-HADS-A의 경우 27명(25.5%)에서 8점 이상으로 나타났다.

라쉬 분석에 따른 단일차원성을 보았을 때, K-HADS-D의 경우 라쉬 측정에 의해 설명되는 분산은 54.0%였으며, 라쉬 요소를 제외한 첫 번째 및 두 번째 잔차 분산의 고유값이 각각 1.66 및 1.46이었다. K-HADS-A에서는 잔차 주성분 분석에서 라쉬 측정에 의한 분산이 54.3%였고, 라쉬 요인을 제외한 첫 번째 및 두 번째 잔차 분산 고유값은 1.68 및 1.34였다. Table 2와 같이 K-HADS-D 문항들의 점-측정 상관계수 범위는 0.47-0.76으로 양호하였으며, K-HADS-A 점-측정 상관계수 범위도 0.58-0.79로 모든 문항에서 0.30 이상이었다. 이에 K-HADS-D 및 K-HADS-A 모두에서 단일차원성이 성립함을 확인하였다. K-HADS-D의 문항별 라쉬 모형 적합도를 보았을 때, 내적합 평균자승잔차의 범위는 0.64-1.26으로 K-HADS-D의 모든 문항에서 적합하였으나, 외적합의 경우 K-HADS-D 14번 문항에서 평균자승잔차가 1.59로 부적합을 보이는 것으로 판단되었다(Table 2). K-HADS-A의 경우 모든 문항들에서 내적합 및 외적합 모두 평균자승잔차가 0.5 미만이거나 1.5를 초과하지 않았다.

Figure 1의 환자-문항 분포도에 따르면 환자 로짓 평균(mean logit for patient)을 기준으로 하였을 때 K-HADS-D의 문항 로짓 평균(mean logit for item)과의 차이가 -0.73±0.66 (평균±표준오차)로 1.0 이내였다. 문항별로 보면 K-HADS-D 14번 문항이 가장 어려운 문항으로 나타났으며, 반면에 6번 문항은 가장 쉬운 문항이었다. 반면, K-HADS-A의 경우에는 환자-문항 로짓 평균 차이는

Table 1. Demographic and clinical characteristics of study patients (n=106)

Characteristic	Value
Age (years)	64.52±8.91
Male	54 (50.9)
Motor symptom duration (years)	1.22±1.11
LEDD (mg/day)	152.4±204.8
Antidepressant	16 (15.1)
Anxiolytics	14 (13.2)
Hoehn and Yahr stage	2.17±0.41
Stage 1	8 (7.5)
Stage 2	87 (82.1)
Stage 3	11 (10.4)
UPDRS motor part score	19.15±8.51
NMSQ total score	10.00±6.59
Gastrointestinal tract	2.06±1.83
Urinary symptom	1.25±0.79
Apathy/attention/memory	1.07±1.08
Hallucination/delusion	0.23±0.50
Depression/anxiety	0.82±0.84
Sexual dysfunction	0.69±0.86
Cardiovascular	0.70±0.81
Sleep disorders	1.85±1.66
Miscellaneous	1.35±1.23
K-HADS total score	13.53±8.00
PDQ-8 score	6.98±6.44

Values are presented as mean±standard deviation or number (%).
LEDD; levodopa equivalent daily dose, UPDRS; Unified Parkinson's Disease Rating Scale, NMSQ; Non-Motor Symptoms Questionnaire, K-HADS; the Korean version of the Hospital Anxiety and Depression Scale, PDQ-8; the Short-form 8-item Parkinson's Disease Questionnaire.

Table 2. Difficulty estimates and point measure correlation of K-HADS based on the item-response theory

Item	Difficulty	SE	Infit MnSq	Outfit MnSq	Point measure correlation	
					Observed	Expected
K-HADS-D subscale						
K-HADS-D2	-0.50	0.14	0.98	1.02	0.72	0.70
K-HADS-D4	0.80	0.15	0.64	0.59	0.73	0.63
K-HADS-D6	-0.90	0.14	0.85	0.82	0.76	0.71
K-HADS-D8	-0.40	0.14	1.26	1.31	0.59	0.70
K-HADS-D10	0.45	0.14	1.11	1.01	0.67	0.66
K-HADS-D12	-0.53	0.14	0.82	0.80	0.76	0.70
K-HADS-D14	1.09	0.16	1.22	1.59	0.47	0.61
K-HADS-A subscale						
K-HADS-A1	0.69	0.19	1.18	1.17	0.58	0.64
K-HADS-A3	-0.31	0.16	1.09	0.96	0.71	0.70
K-HADS-A5	-0.43	0.16	1.28	1.26	0.66	0.71
K-HADS-A7	-1.43	0.15	1.19	1.25	0.65	0.75
K-HADS-A9	1.69	0.23	0.94	0.75	0.60	0.55
K-HADS-A11	-0.07	0.16	0.85	0.78	0.77	0.69
K-HADS-A13	-0.15	0.16	0.62	0.61	0.79	0.69

The MnSq value of infit and outfit was expected to be 1, and the acceptable range of the MnSq value was 0.5-1.5.
K-HADS; the Korean version of Hospital Anxiety and Depression Scale, SE; standard error, MnSq; mean-square, K-HADS-D; K-HADS-depression, K-HADS-A; K-HADS-anxiety.

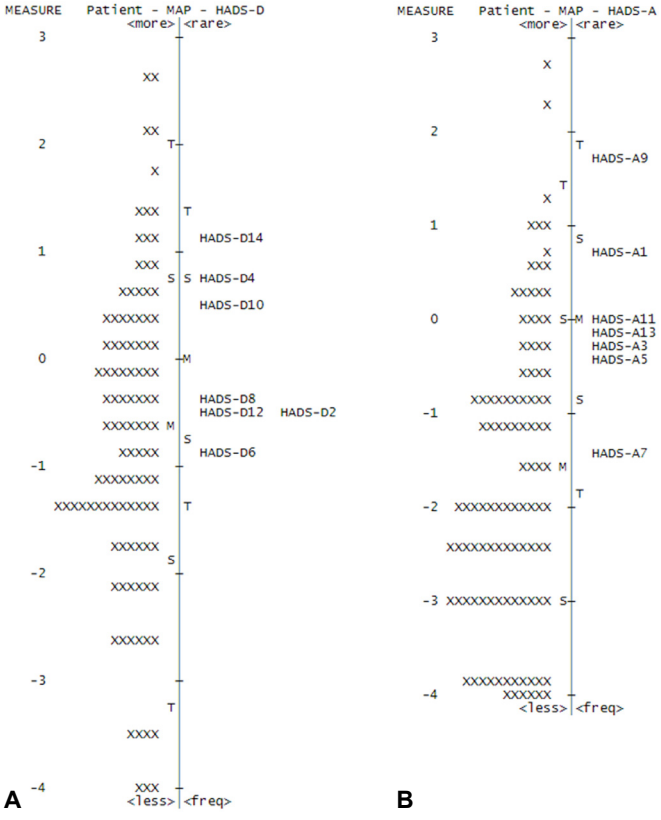

Figure 1. Patient-item distribution plot (wright map) for the Korean version of the Hospital Anxiety and Depression Scale depression (K-HADS-D) and anxiety (K-HADS-A) subscales in the patients with Parkinson's disease. The Wright map displayed the distribution of patients and K-HADS items side by side as calibrated on a linear logit scale (center vertical line). The patient estimates are located on the right side of the map and arranged from the lowest perceived levels of depression or anxiety (bottom) to the highest levels (top). As for the items on the lest side of the linear scale, the ordering goes from the least endorsed item (bottom) to the most endorsed (top). Each 'X' indicates one patient. (A) Wright map for the K-HADS-D. (B) Wright map for the K-HADS-A. M; mean of patient or item distribution, S; 1 standard deviation from the mean, T; 2 standard deviations.

-1.75±0.82로 파킨슨병 환자들이 경험하는 불안 정도보다 문항의 난이도가 대체로 높아 표적 이탈(off-target)로 해석할 수 있었으며, 파킨슨병 환자에서 가장 어려운 문항은 K-HADS-A 9번 문항이었고 가장 쉬운 문항은 7번 문항이었다. K-HADS-D의 대상자 분리 신뢰도는 0.76, 문항 분리 신뢰도는 0.96이었으며 K-HADS-A의 대상자 분리 신뢰도는 0.73, 문항 분리 신뢰도는 0.96으로, 두 척도 모두 문항 분리 신뢰도는 0.80 이상으로 우수하였다. 반면, 대상자 분리 신뢰도의 경우에는 0.60의 수용 기준치에는 도달하였으나 0.80 미만으로 측정되었다. K-HADS-D 및 K-HADS-A의 4점 리커트 척도가 적절히 기능하는지 평가하기 위해 앤드리치의 평정척도 모형을 적용하였을 때 모든 응답 범주의 외적합 평균자승잔차가 2.0 미만이었으며, 평균 측정치 또한 순서적으로 낮은 점수에서 높은 점수로 증가하였다(Table 3). 응답 범주 확률곡선을 도시한 결과에서도 마찬가지로 각각의 응답 범주들이 다른 범주들과 구분되는 영역에 있으면서 범주 간의 교차점이 단조적으로 증가하는 것을 확인할 수 있었다(Fig. 2).

전술한 라쉬 분석 결과와 상호 비교하기 위해 고전적 검사이론에 의한 신뢰도 및 타당도 분석을 진행하였다(Table 4). 문항 간의

Table 3. Threshold ordering test statistics for K-HADS

Response categories	Count	Average measure	Step calibration	Outfit MnSq
K-HADS-D subscale				
K-HADS-D category 0	241 (32.5)	-1.87	None	1.03
K-HADS-D category 1	241 (32.5)	-0.84	-1.40	0.97
K-HADS-D category 2	164 (22.1)	0.35	0.13	0.96
K-HADS-D category 3	96 (12.9)	1.20	1.27	1.13
K-HADS-A subscale				
K-HADS-A category 0	345 (46.5)	-2.80	None	0.98
K-HADS-A category 1	268 (36.1)	-1.14	-1.83	0.84
K-HADS-A category 2	78 (10.5)	0.30	0.72	0.88
K-HADS-A category 3	51 (6.9)	1.04	1.11	1.38

Values are presented as number (%).
The response category with an outfit MnSq of greater than 2 was considered to be misfit.
K-HADS; the Korean version of Hospital Anxiety and Depression Scale, MnSq; mean-square, K-HADS-D; K-HADS-depression, K-HADS-A; K-HADS-anxiety.

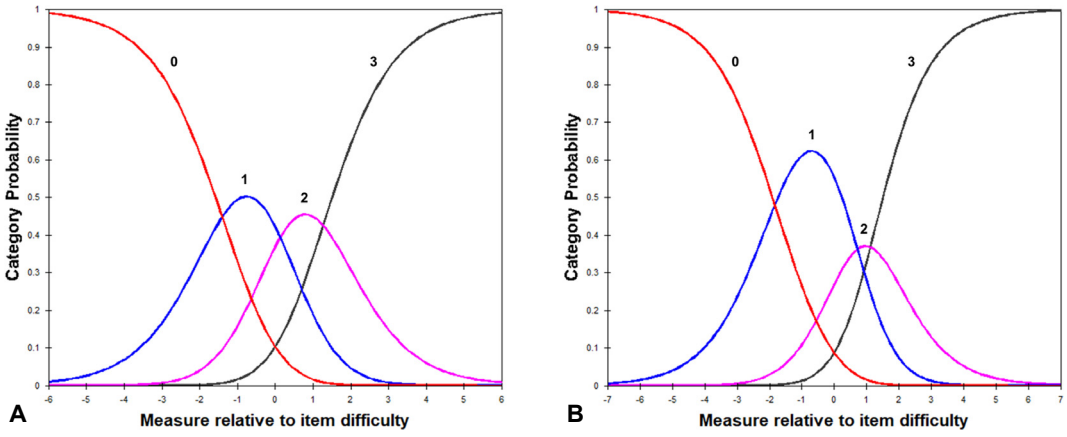

Figure 2. Probability curves of the four response categories (0, 1, 2, and 3) of the Korean version of the Hospital Anxiety and Depression Scale depression (K-HADS-D) and anxiety (K-HADS-A) subscales. Each category probability curve indicated the likelihood that a patient with depression or anxiety will endorse a response category. The threshold represented the point at which two adjacent response categories intersect and a patient had equivalent probability of endorsing one response category or the other at this intersection. Category probability curve for the K-HADS-D (A) or the K-HADS-A (B) showed ordered thresholds in the patients with Parkinson's disease.

Table 4. Clinimetric data of K-HADS based on the classical test theory

Item	Mean	Median [IQR]	Item-total correlation
K-HADS-D subscale			
K-HADS-D2	1.39	2 [0-2]	0.60
K-HADS-D4	0.77	1 [0-1]	0.68
K-HADS-D6	1.59	2 [1-2]	0.64
K-HADS-D8	1.35	1 [1-2]	0.43
K-HADS-D10	0.92	1 [0-2]	0.57
K-HADS-D12	1.41	1 [0-2]	0.67
K-HADS-D14	0.66	1 [0-1]	0.35
K-HADS-D total	8.08	8 [5-11]	Cronbach's α = 0.82
K-HADS-A subscale			
K-HADS-A1	0.53	0 [0-1]	0.49
K-HADS-A3	0.85	1 [0-1]	0.59
K-HADS-A5	0.90	1 [0-1]	0.55
K-HADS-A7	1.31	1 [1-2]	0.47
K-HADS-A9	0.30	0 [0-1]	0.57
K-HADS-A11	0.76	1 [0-1]	0.70
K-HADS-A13	0.79	1 [0-1]	0.75
K-HADS-A total	5.44	4 [2-8]	Cronbach's α=0.83

K-HADS; the Korean version of Hospital Anxiety and Depression Scale, IQR; interquartile range, K-HADS-D; K-HADS-depression, K-HADS-A; K-HADS-anxiety.

합치도를 평가하는 크론바흐 alpha 계수는 K-HADS-D에서 0.82였으며 K-HADS-A는 0.83으로 모두 0.70의 기준에 도달하였다. 문항-전체 상관에서는 7개 문항으로 구성된 K-HADS-D의 경우 0.35-0.68이었으며, 역시 K-HADS-A의 경우 0.47-0.75의 범위로 모든 문항에서 경계치인 0.30을 초과하였다. 다만, 앞서 문항별 라쉬 모형 적합도 분석에서 외적합 평균자승잔차가 부적합을 보인 K-HADS-D 14번 문항의 경우 문항-전체 상관계수에서도 0.35로 다른 문항들에 비해 상관 정도가 상대적으로 낮음을 알 수 있었다. 다음으로 NMSQ와의 수렴 타당성 평가에서는 NMSQ 16번 우울 문항 응답에 따라 K-HADS-D 점수가 유의하게 높았으며 (Mann-Whitney's U=849.0, p<0.000). 마찬가지로 NMSQ 17번 불안 문항에 따라 K-HADS-A 점수도 통계적으로 유의하게 높은 수준임을 알 수 있었다(Mann-Whitney's U=388.0, p<0.000).

고 찰

우울 및 불안장애는 파킨슨병의 여러 비운동 증상들 중에서도 빈번하게 발생하는 반면, 기존 연구들에 따르면 진료기관에 방문한 파킨슨병 환자에서 우울장애의 경우 약 46.2%, 불안장애의 경우 약 51.3%가 외래 방문 시 적절한 진단을 받지 못하는 것으로 조사되었다.[6,32] 이에 실제 임상 현장에서 파킨슨병 환자의 우울 및 불안을 측정하고 진단하기 위한 다양한 임상평정척도들이 제안된 바 있다.[7,9] 본 연구의 대상인 HADS 외에도, 파킨슨병의 우울 증상에 대해서는 Hamilton 우울평가척도, Montgomery-Åsberg 우울평가척도, Beck 우울검사, 노인우울척도(15-item Geriatric Depression Scale) 등이 사용되며,[5-7] 불안 증상에 대해서는 Beck 불안검사, Hamilton 불안평가척도, 파킨슨병불안척도(Parkinson's anxiety scale) 등이 이용되고 있다.[8-10] 그러나 2014년 Leentjens 등이 개발한 파킨슨병 불안척도[9] 외에는 앞서 말한 대부분의 우울 및 불안 평정척도들이 최초 구성 시에 파킨슨병 환자를 염두에 두기보다는 일반 고령층 환자 또는 정신건강의학과 진료 과정에서 정동장애가 의심되는 경우를 상정하여 설계된 경우가 많다.[5,8] 이에 파킨슨병 환자를 대상으로 한 우울 및 불안 평정척도의 신뢰도와 타당도의 확보에 대한 임상 현장의 요구가 늘어나고 있다.

최근까지 신뢰도 및 타당도 연구에 주로 사용된 고전적 검사이론은 관찰된 평정척도의 점수(observed rating scale score)가 진점수(true score)와 오차(error)의 합이며 동일한 대상에 대해 여러 차례 실행할 경우 관찰된 점수들의 기대값이 진점수로 수렴하다는 가정에 근거한다.[12-14] 즉, 고전적 검사이론에 따른 신뢰도는 해당 평정척도가 반복 실행에서 얼마나 안정적으로 측정되는지 평가하고 타당도는 평정척도가 실제 임상적 속성을 반영하는 정도에 대해 기준 타당도(criterion validity), 구성 타당도(construct validity), 내용 타당도(content validity)로 각각 나누어 확인하게 된다. 전술한 고전적 검사이론의 통계적 개념은 직관적이고 간명하여 연구자들에게 익숙하다는 장점이 있으나, 반면 문항의 난이도에 따라 환자의 점수가 상호 의존적으로 바뀌므로 실제 환자의 증상 수준에 대한 산출이 과대 추정 또는 과소 추정될 수 있다. 또한, 자료의 등간성이 유지되지 않고 임상에서 자주 이용되는 다분 리커트 척도의 기능에 대해 제한적인 해석만 가능하다는 문제들이 지적되었다.[11,14,17]

위와 같은 고전적 검사이론의 한계를 보완하기 위해 개발된 현대적 검사이론(modern test theory)인 문항반응이론은 비선형 문항반응함수(non-linear item response function)에 따른 환자의 개별 문항 응답에 대한 확률적 모형화(probabilistic modeling)에 이론적 기반을 둔다. 즉, 문항반응함수는 해당하는 문항에 대해 특정한 정도의 증상을 가진 환자가 특정한 응답을 할 확률을 추정하는데, 예를 들면 불안 증상이 경한 환자의 경우 K-HADS의 11번 문항 '나는 가만히 있지 못하고 안절부절한다(I feel restless as I have to

be on the move)' 항목의 답변에서 4점 리커트 척도에 대한 응답이 0점(전혀 그렇지 않다) 또는 1점(가끔 그렇다)으로 낮게 나올 가능성이 큰 반면, 보다 불안 증상이 심하다면 리커트 척도 중 2점(자주 그렇다) 또는 3점(매우 그렇다)의 범주를 시인할 가능성이 높다고 확률적으로 기대할 수 있다.[29] 이와 같이 단조적(monotonic) 비선형 문항반응함수는 각 문항별로 환자가 특정 범주를 시인할 확률을 구하여 해당 확률을 로짓 점수로 변환하며,[13] 개별 문항 각자의 고유한 로지스틱 함수에 따른 문항 특성 곡선(item characteristic curve)에 기반하여 서열 척도(ordinal scale)를 등간 척도(interval scale)의 성격을 가진 로짓으로 변환하므로 자료의 등간성이 확보되며 문항 난이도와 환자의 증상 정도가 상호 독립적(mutually independent)으로 추정하므로 문항반응이론에 따른 환자 증상 수준은 표본집단 또는 문항 구성에 따라 불변하게 된다.[17] 이와 함께 다분 응답 범주(polytomous response category)에 대해서 적합도 지표(fit index)가 제공되므로 리커트 척도의 기능에 대한 세부적인 통계 분석이 가능한 장점이 있다.[11,14,30] 본 연구에서는 106명의 파킨슨병 환자의 우울 및 불안 증상에 대해 K-HADS를 사용하여 평가한 자료를 바탕으로 고전적 검사이론에 따른 내적 신뢰도 및 수렴 타당도를 측정하였고, 이와 함께 문항반응이론에 따른 라쉬 분석을 시행하여 단일차원성, 개별 문항의 라쉬 모형 적합도, 분리 신뢰도 및 다분 응답 범주 기능을 분석하였다.

우선 고전적 검사이론에 따른 K-HADS의 신뢰도 및 타당도 분석을 살펴보면 단일 실행의 신뢰성을 의미하는 크론바흐 alpha 계수가 K-HADS-D에서 0.82, K-HADS-A에서 0.83으로 양호하였으며 문항-전체 상관계수의 분포 범위도 모두 적합하였다. 또한 NMSQ의 우울 및 불안 점수와의 수렴 타당도 역시 통계적으로 유의한 차이를 보였다. 이는 이전 Oh 등[23]의 K-HADS 표준화 연구에서 크론바흐 alpha 계수가 K-HADS-D에서 0.86, K-HADS-A에서 0.89로 중등도 이상의 내적 신뢰도를 보였던 결과와 일관된 결과이다. 해외에서는 Rodriguez-Blázquez 등[33]이 387명의 파킨슨병 환자를 대상으로 한 연구에서 적합한 검사-재검사 신뢰도를 보였고 HADS-D(크론바흐 alpha=0.83) 및 HADS-A(크론바흐 alpha=0.81)에서 적절한 내적 일치도를 보고한 바 있다.

다만, 상기 고전적 검사이론에 의한 신뢰도 및 타당도 결과에 비해 라쉬 분석에서는 K-HADS-D 또는 K-HADS-A의 일부 제한점들이 밝혀졌다. 우선 문제가 되는 것은 K-HADS-D 및 K-HADS-A는 모두 상대적으로 낮은 대상자 분리 신뢰도를 보였다는 점이다. 두 척도의 대상자 분리 신뢰도는 각각 0.76 및 0.73으로서 0.60의 수용 수준에는 도달하였으나, 모두 0.80 이하로 나타났으며 이는 K-HADS-D 또는 K-HADS-A가 대상 환자를 증상 정도에 따라 경증 집단, 중등도 집단이나 중증 집단 등의 분리된 집단으로 세분하는 데 채택하기는 어렵다는 점을 의미한다.[11,13,19] 상기의 대상자 분리 신뢰도의 문제는 K-HADS의 서열 척도를 등간 변환하여 도시한 환자-문항 분포도와 연관되는데 Figure 1-A에서 살펴보면 K-HADS-D의 평정척도 표적은 환자 및 문항 평균 로짓 간 차이가 1.0 이내로 대체로 적합하다고 볼 수 있으나 측정 범위가 협소(narrow)하였는데, 예컨대 K-HDAS-D 4번 및 14번 등의 일부 문항은 환자들의 우울 증상 수준에 비해 난이도가 상대적으로 높으며, 반대로 경한 우울 증상을 변별하기 위한 문항이 미흡함을 확인할 수 있다. 특히 K-HADS-A의 경우에는 상당수의 파킨슨병 환자들이 임상평정척도의 측정 범위를 벗어나 바닥효과(floor effect)를 보이고 있어 평정척도 표적 이탈이 있다고 평가되었다(Fig. 1-B). 문항 난이도의 평균 및 환자 증상 정도 범위가 근접하거나 일치할 때 평정척도의 변별력이 높다고 해석할 수 있으므로, 파킨슨병 환자에서 가벼운 정도의 불안 증상은 K-HADS-A로 파악하기 어려울 수 있으며, 위와 같은 표적 이탈 또는 협소한 측정 범위의 문제는 민감도가 높은 다른 임상평정척도를 병행하거나 비교적 낮은 정도의 우울 또는 불안 증상까지 포괄할 수 있도록 기존 K-HADS 문항을 수정 또는 새 문항의 추가가 필요하다는 점을 시사한다.

다음으로 K-HADS-D 및 K-HADS-A에 대해 단일차원성을 평가하였을 때 잔차 주성분 분석과 문항 극성 평가에서 모두 단일차원성을 지지하는 결과를 보였다. 이는 앞서 고전적 분석에 따라 내적 합치도를 의미하는 크론바흐 alpha 계수 및 문항-전체 상관이 모두 적합하였던 것과 일관된 결과이다.[23,33] 단, 세부 문항별 라쉬 모형 적합도 분석을 보았을 때 K-HADS-D 14번 문항인 '나는 책이나 라디오, 텔레비전 프로그램을 즐긴다(I can enjoy a good book or radio or TV program)'의 경우 외적합 평균자승잔차에서 라쉬 모형 부적합을 보였는데, 이는 Forjaz 등[16]이 파킨슨병 환자를 대상으로 문항반응이론에 따른 분석을 수행하였을 때 HADS-A 문항들은 대체로 라쉬 모형에 적합함을 확인하였으나, HADS-D에서는 라쉬 측정 모형에 부적합한 문항이 관찰되었던 결과와 유사하다. 다만 본 연구에서는 K-HADS-D 14번 문항이 부적합을 보였으나 Forjaz 등[16]의 연구에서는 주로 HADS-D 2번 문항에서 부적합이 발생하였는데, 이러한 차이가 발생한 이유로는 두 연구 간의 환자군 상이성 또는 연구 대상 도구의 차이에 의할 것으로 사료된다.

마지막으로 앤드리치의 다분 평정척도 모형에 따른 응답 범주 분석을 살펴보면, 응답 범주 확률곡선, 평균 측정치 및 단계 조정치가 위계적으로 증가하였고 모든 응답 범주의 외적합 평균자승잔차가 2.0 미만이었으므로, K-HADS-D 및 K-HADS-A의 다분

응답 범주의 기능이 적절한 것으로 해석할 수 있었다. 일부 연구들에서 파킨슨병 환자 등 고연령 대상자에서 우울 또는 불안 증상을 평가할 경우, HADS를 비롯하여 Hamilton 우울평가척도, Montgomery-Åsberg 우울평가척도, Beck 불안검사 및 파킨슨불안척도 등에서 채용하는 다범주 리커트 척도에 의한 평가보다는 노인우울척도와 같이 이분형 응답에 의한 평가가 보다 단순하고 적절하다는 의견이 있었다.[7,9,11] 그러나 이번 연구 결과에서는 다분 응답 범주가 포함되는 리커트 척도가 파킨슨병 환자에서도 대체로 적합한 기능을 수행하는 것을 통계적으로 확인할 수 있었다.

본 연구의 제한점은 다음과 같다. 먼저, 고전적 검사이론에 따른 신뢰도 분석에서는 크론바흐 alpha 계수 등 단일 실행의 내적 일치도를 측정하는 방법과 함께, 동일한 연구 대상에 대해 반복 실행하여 측정할 수 있는 검사-재검사 신뢰도 및 척도의 반응도(responsiveness)가 주요한 부분을 이룬다.[11,12] 본 연구에서는 1회의 면접 및 검진에 의해 진행되었고 재검사를 시행하지 않아 검사-재검사 신뢰도를 평가하지 못하였으며 척도 반응도 또한 확인하지 못한 한계점이 있다. 다음으로 이번 연구의 대상 환자는 평균 유병 기간이 1.22년으로 비교적 초기의 파킨슨병 환자들로 구성되었고 증상 정도를 나타내는 평균 Hoehn-Yahr척도 역시 2.17로 대체로 낮았다. K-HADS-A에서 나타난 바닥효과나 K-HADS-D의 협소한 측정 범위 문제 그리고 K-HADS-D와 K-HADS-A의 문항 난이도가 대체로 환자의 불안 또는 우울 증상 정도에 비해 높았던 점은 연구 대상의 운동 증상 및 비운동 증상이 심하지 않은 초기 파킨슨병 환자군에서 시행되었다는 점이 일부 원인이 되었을 수 있다.[10,33] 마지막으로, 본 연구는 대조군을 포함하지 않았으며 연구 대상자에서 K-HADS의 시행과 병행하여 우울 및 불안장애에 대한 정신건강의학과 진료 등의 추가적인 정동장애 진단 절차를 수행하지 않았다. 기존 Oh 등[23]의 표준화 연구에서는 임상 진료 및 정신질환 진단 및 통계 편람 4판의 진단기준에 따라 정신건강의학과 의사 2인 이상이 동의하여 우울장애를 진단받은 74명의 환자 및 불안장애가 진단된 66명의 환자를 표준 진단 집단(gold standard group)으로 하여 189명의 정상 대조군에 대해 우울 및 불안장애 선별을 위한 기준 타당도를 조사하였으나 이번 연구에서는 NMSQ와의 수렴 타당도 이외에 기준 타당도에 대한 연구는 수행하지 않았다는 한계점이 있다.

본 연구는 국내 파킨슨병 환자에서 K-HADS에 대해 최초로 시도된 문항반응이론 기반 분석 연구이다. 단일 병원 신경과에 방문한 106명의 파킨슨병 환자들의 우울 및 불안 증상을 K-HADS로 측정한 후 해당 자료에 대해 고전적 검사이론에 따른 신뢰도 및 타당도 분석과 함께 라쉬 분석을 진행하였으며, K-HADS-D 및 K-HADS-A가 전반적으로 적합한 단일차원성, 문항 극성 및 내적 신뢰도, 수렴 타당도를 보이는 임상평정척도임을 보여주었다. 반면, K-HADS-D의 경우 라쉬 모형에 부적합한 문항이 확인되었고, K-HADS-A에서는 문항 표적의 이탈 및 바닥효과를 보이는 것으로 조사되었으며, 두 척도 모두 상대적으로 낮은 대상자 분리 신뢰도를 보였다. 본 연구에서 확인된 임상평정척도 상의 제한점들에 대해 향후 검토 및 문항 개선이 이뤄질 경우, 파킨슨병 환자의 우울 및 불안장애의 정량적 측정 및 효과적인 치료 계획 수립에 도움이 될 것으로 기대한다.

REFERENCES

1. Chaudhuri KR, Martinez-Martin P, Schapira AH, Stocchi F, Sethi K, Odin P, et al. International multicenter pilot study of the first comprehensive self-completed nonmotor symptoms questionnaire for Parkinson's disease: the NMSQuest study. *Mov Disord* 2006;21:916-23.
2. Martinez-Martin P, Rodriguez-Blazquez C, Kurtis MM, Ray Chaudhuri K, on behalf of the NMSS Validation Group. The impact of non-motor symptoms on health-related quality of life of patients with Parkinson's disease. *Mov Disord* 2011;26:399-406.
3. Kwon DY, Koh SB, Lee JH, Park HK, Kim HJ, Shin HW, et al. The KMDS-NATION study: Korean movement disorders and quality of life in Parkinson's disease. *J Clin Neurol* 2016;12:393-402.
4. Sauerbier A, Jitkritsadakul O, Titova N, Klingelhoefer L, Tsuboi Y, Carr H, et al. Non-motor symptoms assessed by non-motor symptoms questionnaire and non-motor symptoms scale in Parkinson's disease in selected Asian populations. *Neuroepidemiology* 2017;49:1-17.
5. Schrag A, Barone P, Brown RG, Leentjens AFG, McDonald WM, Starkstein S, et al. Depression rating scales in Parkinson's disease: critique and recommendations. *Mov Disord* 2007;22:1077-1092.
6. Williams JR, Hirsch ES, Anderson K, Bush AL, Goldstein SR, Grill S, et al. A comparison of nine scales to detect depression in Parkinson disease: which scale to use? *Neurology* 2012;78:998-1006.
7. Torbey E, Pachana NA, Dissanayaka NN. Depression rating scales in Parkinson's disease: a critical review updating recent literature. *J Affect Disord* 2015;184:216-224.
8. Leentjens AF, Dujardin K, Marsh L, Martinez-Martin P, Richard IH, Starkstein SE, et al. Anxiety rating scales in Parkinson's disease: critique and recommendations. *Mov Disord* 2008;23:2015-2025.
9. Dissanayaka NNW, Torbey E, Pachana NA. Anxiety rating scales in Parkinson's disease: a critical review updating recent literature. *Int Psychogeriatr* 2015;27:1777-1784.
10. Mondolo F, Jahanshahi M, Granà A, Biasutti E, Cacciatori E, Di Benedetto P. Evaluation of anxiety in Parkinson's disease with some commonly used rating scales. *Neurol Sci* 2007;28:270-275.
11. Forjaz MJ, How to evaluate validation data. In: Sampaio C, Goetz CG, Schrag A. *Rating scales in Parkinson's disease: clinical practice and research*. New York: Oxford University Press, 2012;16-41.
12. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res*

2010;19:539-549.
13. Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Rheum* 2007;57:1358-1362.
14. Cappelleri JC, Jason Lundy J, Hays RD. Overview of classical test theory and item response theory for the quantitative assessment of items in developing patient-reported outcomes measures. *Clin Ther* 2014;36:648-662.
15. Riley WT, Pilkonis P, Cella D. Application of the National Institutes of Health Patient-reported Outcome Measurement Information System (PROMIS®) to mental health research. *J Ment Health Policy Econ* 2011;14:201-208.
16. Forjaz MJ, Rodriguez-Blázquez C, Martinez-Martin P. Rasch analysis of the hospital anxiety and depression scale in Parkinson's disease. *Mov Disord* 2009;24:526-532.
17. da Rocha NS, Chachamovich E, de Almeida Fleck MP, Tennant A. An introduction to Rasch analysis for psychiatric practice and research. *J Psychiatr Res* 2013;47:141-148.
18. Forjaz MJ, Ayala A, Rodriguez-Blazquez C, Frades-Payo B, Martinez-Martin P. Assessing autonomic symptoms of Parkinson's disease with the SCOPA-AUT: a new perspective from Rasch analysis. *Eur J Neurol* 2010;17:273-279.
19. Kletzel SL, Hernandez JM, Miskiel EF, Mallinson T, Pape TL. Evaluating the performance of the montreal cognitive assessment in early stage Parkinson's disease. *Parkinsonism Relat Disord* 2017;37:58-64.
20. Yang HJ, Ahn JH, Lee J, Lee WK, Lee J, Kim Y. Measuring anxiety in patients with early-stage Parkinson's disease: Rasch analysis of the state-trait anxiety inventory. *Front Neurol* 2019;10:49.
21. Katzenschlager R, Schrag A, Evans A, Manson A, Carroll CB, Ottaviani D, et al. Quantifying the impact of dyskinesias in PD: the PDYS-26: a patient-based outcome measure. *Neurology* 2007;69:555-563.
22. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983;67:361-370.
23. Oh SM, Min KJ, Park DB. A study on the standardization of the hospital anxiety and depression scale for Koreans: a comparison of normal, depressed and anxious groups. *J Korean Neuropsychiatr Assoc* 1999;38:289-296.
24. Schade S, Mollenhauer B, Trenkwalder C. Levodopa equivalent dose conversion factors: an updated proposal including opicapone and safinamide. *Mov Disord Clin Pract* 2020;7:343-345.
25. Mondolo F, Jahanshahi M, Granà A, Biasutti E, Cacciatori E, Di Benedetto P. The validity of the hospital anxiety and depression scale and the geriatric depression scale in Parkinson's disease. *Behav Neurol* 2006;17:109-115.
26. Chiang KS, Green KE, Cox EO. Rasch analysis of the geriatric depression scale-short form. *Gerontologist* 2009;49:262-275.
27. Li CY, Romero S, Bonilha HS, Simpson KN, Simpson AN, Hong I, et al. Linking existing instruments to develop an activity of daily living item bank. *Eval Health Prof* 2016;41:25-43.
28. Khadka J, Gothwal VK, McAlinden C, Lamoureux EL, Pesudovs K. The importance of rating scales in measuring patient-reported outcomes. *Health Qual Life Outcomes* 2012;10:80.
29. Kreijns K, Bijker M, Weidlich J. A Rasch Analysis Approach to the Development and Validation of a Social Presence Measure. In: Khine MS. *Rasch Measurement: Applications in quantitative educational research*. Singapore: Springer Nature, 2020;197-221.
30. Linacre JM. Optimizing rating scale category effectiveness. *J Appl Meas* 2002;3:85-106.
31. Janssen KC, Phillipson S, O'Connor J, Johns MW. Validation of the Epworth sleepiness scale for children and adolescents using Rasch analysis. *Sleep Med* 2017;33:30-35.
32. Shulman LM, Taback RL, Rabinstein AA, Weiner WJ. Non-recognition of depression and other non-motor symptoms in Parkinson's disease. *Parkinsonism Relat Disord* 2002;8:193-197.
33. Rodriguez-Blázquez C, Frades-Payo B, Forjaz MJ, de Pedro-Cuesta J, Martinez-Martin P; Longitudinal Parkinson's Disease Patient Study Group. Psychometric attributes of the hospital anxiety and depression scale in Parkinson's disease. *Mov Disord* 2009;24:519-525.