


Psychometric validation of a patient-reported single-item assessment of 'Good Day Bad Day' in a neurogenic orthostatic hypotension population treated with droxidopa

Clément François ^a, Nicola Germain^a, Renata Majewska^b, Vanessa Taieb^c, L. Arthur Hewitt^d and Steven Kymes^e

^aHealth Economics and Outcomes Research, Creativ-Ceutical, Paris, France; ^bHealth Economics and Outcomes Research, Creativ-Ceutical S.A, Kraków, Poland; ^cHealth Economics and Outcomes Research, Creativ-Ceutical, London, UK; ^dMedical Affairs, Lundbeck, Deerfield, IL, USA; ^eHealth Economics and Outcomes Research, Lundbeck, Deerfield, IL, USA

ABSTRACT

Background

Symptoms of neurogenic orthostatic hypotension (nOH), including lightheadedness/dizziness, presyncope, syncope, and falls, can lead to impaired functional ability and reduced quality of life. Because the severity and frequency of nOH symptoms fluctuate, it may be difficult for patients to accurately quantify the effect of symptoms on their daily lives using available outcome measures. A new single-item instrument, the 'Good Day Bad Day,' was developed, and its psychometric validity was assessed in patients with nOH.

Methods

Data from a 6-month, prospective, observational cohort study of patients with nOH who were newly initiating droxidopa treatment were used. Patients were asked to quantify the number of good and bad days in the previous 7 days and responded to other validated patient-reported outcomes instruments. The concurrent and discriminant validities and the stability of the Good Day Bad Day instrument were assessed.

Results

A total of 153 patients were included in the analysis (mean [SD] age, 62.3 [17] years). Change in the number of good days moderately correlated with improvements in other patient-reported outcomes (rho value range, -0.38 to -0.61). When data were examined categorically (low vs high symptom severity), the mean number of good days was higher in subgroups representing low symptom severity across measures at 1, 3, and 6 months (all $P \leq 0.01$).

Conclusions

The Good Day Bad Day instrument provided good discrimination at baseline and over time and may aid in assessment of the effects of nOH symptoms on patients.

ARTICLE HISTORY


Received 13 April 2021
Revised 18 November 2021
Accepted 19 November 2021

KEYWORDS

Falls; functional impairment; patient-reported outcomes; single-item instrument; symptoms; validity

CONTACT Clément François  CFR@creativ-ceutical.com  Executive Vice President, Health Economics and Outcomes Research, Creativ-Ceutical, 215 Rue Faubourg St. Honoré, Paris 75008, France

[†]Employee of Creativ-Ceutical at the time of validation analyses.

 Supplemental data for this article can be accessed [here](#).

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

Neurogenic orthostatic hypotension (nOH) is a sustained fall in blood pressure on standing due to autonomic failure, with clinical signs and symptoms that include lightheadedness, dizziness, presyncope, syncope, and falls [1–4]. nOH can reduce a patient's mobility, negatively affect the ability to perform daily activities, and decrease independence and quality of life (QoL) [3,4]. In the clinical development trials of droxidopa, an approved treatment for nOH symptoms[5], patient-reported efficacy outcomes specifically related to nOH symptoms were evaluated using the 10-item Orthostatic Hypotension Questionnaire (OHQ), a validated instrument developed for nOH[6]. The OHQ has a 7-day recall period and consists of two parts, the 6-item Orthostatic Hypotension Symptom Assessment (OHSA), which assesses severity of symptoms, and the 4-item Orthostatic Hypotension Daily Activity Scale, which assesses the effect of symptoms on patients' daily activities[6]. Although the OHQ and its rigorous, detailed quantification of outcomes related to nOH may be useful, especially in clinical research environments, patients may find it challenging to answer multiple questions that each require recall and averaging of specific individual measures of experiences with nOH over the past week because of daily fluctuation of symptoms. As such, more global complementary measurements of how nOH symptoms affect patients' daily lives may be advantageous in nOH evaluation and clinical management. Further, because regulatory agencies (eg, US Food and Drug Administration [FDA], European Medicines Agency) are increasingly emphasizing the evaluation of patient-reported outcomes (PROs) [7–9], development of novel tools to capture patients' perspectives of diseases and treatments are of interest.

Within this context, we developed a new single-item measure for patients with nOH, the 'Good Day Bad Day' instrument. This new PRO scale measures patients' global perception of the effect of nOH symptoms on their lives by asking how many 'good days' and 'bad days' they had over the past week. Here, we describe the psychometric validation analyses conducted to show that the newly developed Good Day Bad Day instrument is clinically relevant and accurately reflects patients' experience with nOH.

Material and methods

Study design and population

Validation of the Good Day Bad Day instrument used data from a 6-month, non-interventional, US-based prospective cohort study of patients newly initiating droxidopa for the treatment of nOH[10]. Study participants consisted of adults (≥ 18 years old) with a confirmed diagnosis of nOH and an underlying diagnosis of primary autonomic failure (Parkinson's disease, multiple-system atrophy, pure autonomic failure), dopamine β -hydroxylase deficiency, or non-diabetic autonomic neuropathy who were enrolled in the NORTHERA® (Lundbeck, Deerfield, IL) pharmacy hub (HUB), a specialty pharmacy used for the distribution of droxidopa. Patients who were non-ambulatory or confined to a wheelchair or diagnosed with dementia, Alzheimer's disease (AD), schizophrenia, or other psychiatric disorders were excluded. All enrolled participants provided consent. Two persons with nOH were contracted to be interviewed twice to support the development of the protocol and the questionnaire by ensuring patient perspective was well considered. They filled in all the questionnaires selected in the Case Report Form and provided comments. Patients commented that it was sometimes difficult to respond to the questionnaire as they had fluctuations of good days and bad days, which led to the addition of the Good day Bad Day questions. As they did not describe days as either good or bad, the dichotomization was kept in the Good Day Bad Day questionnaire.

Study outcomes

In addition to the Good Day Bad Day item (Table 1), study outcomes consisted of the following validated PRO assessments:

- (1) OHSA Item 1[6], which measures 'dizziness, lightheadedness, feeling faint, or feeling that you are about to black out,' the cardinal symptoms of nOH and the primary outcome of the pivotal studies of droxidopa leading to FDA approval;
- (2) Short Falls Efficacy Scale–International (FES-I) [11], which measures feeling of falling;
- (3) Sheehan Disability Scale (SDS) [12], which evaluates functional impairment experienced at work/school as well as in social and family life;
- (4) 8-item Short-Form Health Survey (SF-8) [13], which evaluates physical and mental health;

Table 1. Patient-reported outcomes assessments.

Patient-Reported Outcome	Instrument	Assessment	Recall	Scoring
Good and bad days	Single question	Good or bad days	Past week	<ul style="list-style-type: none"> Number of declared good days
Orthostatic dizziness or lightheadedness	OHSA Item 1 ⁶	Severity of dizziness/lightheadedness	Past week	<ul style="list-style-type: none"> Scored 0–10 Higher value = worse symptoms
Fear of falling	FES-I (7 Items) ¹¹	Concern about the possibility of falling during 7 activities	At time of survey	<ul style="list-style-type: none"> Each activity scored 1–4 Higher value = more concern
Functional impairment	SDS ¹²	Functional impairment in work/school, social, family life domains	Past week	<ul style="list-style-type: none"> Each domain scored 0–10 Higher value = greater impairment
Health-related QoL	SF-8 ¹³	8 items measuring health domains	Past month	<ul style="list-style-type: none"> Each item scored from 1–5 Higher values = poorer outcomes
Depressive symptoms	PHQ-9 ¹⁴	9 items measuring severity of depression symptoms	Past 2 weeks	<ul style="list-style-type: none"> Each item scored 0–3 Higher values = greater symptoms

Abbreviations: FES-I, Short Falls Efficacy Scale–International; OHSA, Orthostatic Hypotension Symptom Assessment; PHQ-9, Patient Health Questionnaire-9; QoL, quality of life; SDS, Sheehan Disability Scale; SF-8, 8-item Short-Form Health Survey.

(5) Patient Health Questionnaire-9 (PHQ-9) [14], which screens for depression.

The Good Day Bad Day item asked, ‘Over the past week, how many “Good Days” did you have?’ and ‘Over the past week, how many “Bad Days” did you have?’ All outcomes were self-reported by patients via an online, paper, or telephone interview assessment at baseline and at 1, 3, and 6 months after initiation of droxidopa.

Psychometric validation and statistical analyses

Concurrent validity

Concurrent validity was assessed by Spearman rank correlation coefficients (ρ) between good days and the validated PRO assessments (OHSA Item 1, FES-I, SDS, SF-8, and PHQ-9) at baseline and each visit. Spearman absolute ρ coefficients were interpreted as follows: very weak (0.00–0.19), weak (0.20–0.39), moderate (0.40–0.59), strong (0.60–0.79), or very strong (0.80–1.0) monotonic correlation.

Discriminant validity

Discriminant validity was assessed by comparing the number of good days among the following defined PRO outcome subgroups at baseline and at 1, 3, and 6 months with the subgroups having low or high responses based on the following categorizations: OHSA Item 1; median score at baseline used as cutoff; FES-I; score ≤ 10 vs > 10 , and PHQ-9 score ≤ 5 vs > 5 .

The number of good days in the subgroups defined above was summarized with standard descriptive statistics (eg, mean, standard deviation). Student *t* test or

Mann-Whitney test, if needed, were used to compare the number of good days among subgroups.

Change in number of good days over time

The validation analysis for the change in number of good and bad days was based on observations where the sum of declared good and bad days equaled 7, and only the number of declared good days was taken into consideration. Observations where the number of good and bad days did not equal 7 were not included in the analysis.

Patterns of change over time

The pattern of change over time for good days and OHSA Item 1 or FES-I was assessed by categorizations of stability, improvement, or deterioration of each outcome. Changes in each outcome from the preceding visit (ie, month 1 vs baseline, month 3 vs month 1, month 6 vs month 3) were calculated. To compare changes, the following subjective categorization scheme was used. For good days, changes were categorized as deterioration (≥ 2 -day decrease), stability (≤ 1 -day change), or improvement (≥ 2 -day increase). For OHSA Item 1 (in which higher score indicates greater symptom severity), changes were categorized as deterioration (≥ 2 -unit score increase), stability (≤ 1 -unit score change), or improvement (≥ 2 -unit score decrease). For FES-I (in which higher score indicates greater concern about falling), changes were categorized as deterioration (≥ 3 -unit score increase), stability (≤ 2 -unit score change), or improvement (≥ 3 -unit score decrease).

The strength of association was assessed using Goodman and Kruskal's gamma (ie, gamma coefficient), which is suitable for use with ordinal variables and varies from -1 to 1 [15]. Agreement between variables was checked using weighted kappa with Cicchetti-Allison weights[16]. Kappa agreement was interpreted as follows: values <0.2 indicate poor agreement, values ranging from 0.2 to <0.4 indicate fair agreement, values ranging from 0.4 to <0.6 indicate moderate agreement, values from 0.6 to <0.8 indicate good agreement, and values ≥ 0.8 indicate very good agreement[17].

Results

Demographics and baseline characteristics

Patient demographics and baseline characteristics are shown in Table 2. At baseline, 164 patients completed the good day/bad day item. After exclusion of 10 patients with missing data (ie, total reported good and bad days over the past week was <7) and 1 patient who reported a sum of >7 good and bad days over the past week, 153 patients were included in the analysis at baseline (mean [SD] age, 62.3 [17.0] years). After 1,3 and 6 months of treatment, 12% ($n = 19/153$), 20% ($n = 31/153$) and 21% ($32/153$) of patients were lost to follow-up.

Concurrent validity

Correlation of good days with the validated PRO assessments ranged from weak (ρ , -0.38 for FES-I at baseline and at 1 month) to strong (ρ , -0.61 for the SDS Work/School at 6 months; Table 3). In general, moderate correlations between increases in good days and improvements in validated PRO measures of nOH

Table 2. Patient demographic and baseline characteristics.

Variable ^a	Patients(N = 153)
Mean (SD) age, y	62.3 (17.0)
Women, n (%)	77 (51.7)
Race, n (%)	
Non-Hispanic white	131 (87.9)
Mean (SD) good days	3.3 (2.1)
Mean (SD) OHSA Item 1 score	5.8 (2.7)
Mean (SD) FES-I (7 items) score	17.1 (5.4)
Mean (SD) SDS scores	
Work/School	6.7 (2.9)
Social	6.2 (2.9)
Family	6.3 (2.9)
Mean (SD) SF-8 summary score	26.3 (6.7)
Mean (SD) PHQ-9 score	11.6 (6.7)

^aData based on all patients who completed the individual assessment; $n = 88$ (for SDS work)- $n = 147$ to 153 for other variables.

Abbreviations: FES-I, Short Falls Efficacy Scale-International; OHSA, Orthostatic Hypotension Symptom Assessment; PHQ-9, Patient Health Questionnaire-9; SDS, Sheehan Disability Scale; SF-8, 8-item Short-Form Health Survey.

Table 3. Correlation between reported good days and other patient-reported outcomes.

	Patient-Reported Outcome	Spearman's Correlation Coefficient (ρ)	P Value
Baseline	OHSA Item 1, $n = 152$	-0.52	<0.0001
	FES-I (7 items), $n = 151$	-0.38	<0.0001
	SDS		
	Work/School, $n = 88$	-0.55	<0.0001
	Social, $n = 147$	-0.54	<0.001
	Family, $n = 149$	-0.53	<0.0001
	SF-8, $n = 151$	-0.53	<0.0001
	PHQ-9, $n = 149$	-0.47	<0.0001

Abbreviations: FES-I, Short Falls Efficacy Scale-International; OHSA, Orthostatic Hypotension Symptom Assessment; PHQ-9, Patient Health Questionnaire-9; SDS, Sheehan Disability Scale; SF-8, 8-item Short-Form Health Survey.

symptoms, functional impairment, and depressive symptoms were identified. The complete list of Spearman correlation coefficients between number of good days and other scales is presented in Supplemental Table S1.

Discriminant validity

During the study, there was an overall significant increase in the mean (SD) number of good days from 3.3 (2.1) days at baseline to 4.5 (2.1) days at 6 months ($P < 0.001$; Figure 1). At all assessment points, the mean number of good days was higher in patient groups with a lower degree of nOH symptoms (OHSA Item 1 scores of dizziness/lightheadedness), concern about falling (FES-I scores), and depressive symptoms (PHQ-9 scores) versus those with higher scores (Figure 1). In patients with low versus high OHSA Item 1 scores, the mean differences in the number of good days ranged from 1.4 to 2.1 days ($P \leq 0.002$). Similar statistically significant mean differences in good days per week were observed for low versus high concern about fear of falling (FES-I scores, 1.2–1.9 days; $P \leq 0.01$) and low versus high levels of depressive symptoms (PHQ-9 scores, 1.5–2.0 days; $P \leq 0.0003$).

Pattern of change in number of good days over time

There was a fair level of agreement in the pattern of change in the number of good days and OHSA Item 1 scores over time (kappa, 0.213–0.382; Supplemental Table S2). However, similar analyses of good days and FES-I scores showed poor to fair agreement (kappa, 0.129–0.246; Supplemental Table S3).

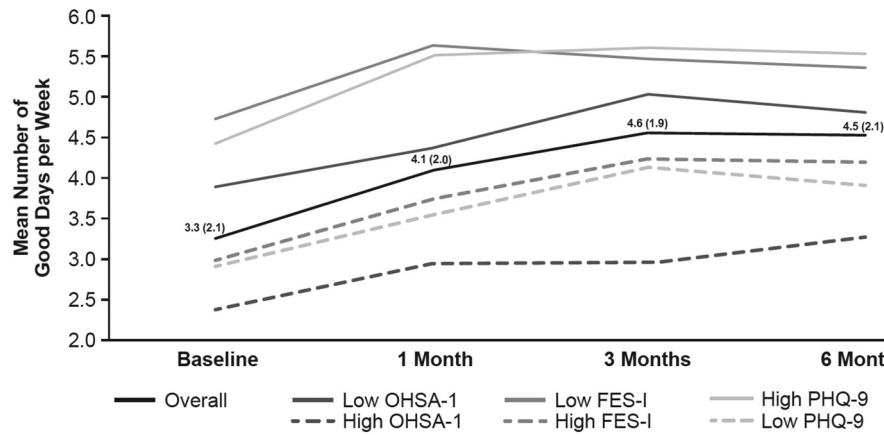


Figure 1. Mean (SD) number of good days at each visit stratified by selected subgroups. Low OHSA Item 1, score ≤ 6 ; high OHSA Item 1, score > 6 . Low FES-I, score 7–10; high FES-I, score 11–28. Low symptoms, PHQ-9 score ≤ 5 ; high symptoms, PHQ-9 score > 5 . **Abbreviations:** FES-I, Short Falls Efficacy Scale–International; OHSA, Orthostatic Hypotension Symptom Assessment; PHQ-9, Patient Health Questionnaire-9.

Loss to follow-up analysis

There were no significant differences in the number of good days at baseline, 1 month, and 3 months between those observed only at 1, 3, or 6 months and those who completed the study. Analyses showed that loss to follow-up did not correlate with number of good days, and the odds of discontinuation decreased by 13% (odds ratio, 0.874; 95% CI: 0.729–1.048) with each additional good day, which indicates a low risk of bias as a result of study discontinuation during any of the treatment periods.

Discussion

This psychometric validation showed that the Good Day Bad Day single-item measure of patients' overall experience of nOH symptoms and impairment is a valid and sensitive instrument. This new PRO instrument was developed to measure patients' overall experiences with nOH because daily fluctuations in nOH symptoms and their associated impact on daily living can make comprehensive assessment of the effects of nOH potentially challenging. The Good Day Bad Day item offers a simple global assessment of overall experience that does not require patients to report and 'average out' multiple nOH symptoms and their effects. We believe that this new tool can be used in conjunction with other validated PRO assessments, such as the OHQ, to help clinicians better understand the effects of nOH symptoms on patients' lives and potentially guide treatment approaches. However, we would recommend to

provide further guidance on the instructions and specify that days have to be either good or bad, and that the total of Good days and Bad days should not exceed seven.

In this validation, increases of approximately 1 good day per week from baseline were measured. In a condition like nOH, the ability to detect the patient's experience of 1 additional good day per week is likely clinically meaningful due to the profound symptomatic burden associated with nOH. Survey data have shown that 87% of patients reported that nOH symptoms negatively affected their ability to perform daily activities, and 42% indicated that nOH symptoms 'robbed them of their independence.'³ Thus, it is likely that even 1 more good day per week represents a substantial decrease in nOH disease burden and may represent a clinically meaningful change. However, that will be the subject of future studies.

To our knowledge, the data reported here are the first to validate the use of a 'good day' measure by correlation with other validated PRO assessments in patients with nOH. However, our approach of assessing a global 'good day' is supported by other studies. In patients with AD, reports of a good or bad day (as noted in health records based on patient and caregiver descriptions) were qualitatively associated with cognitive state, functional ability, and mood (ie, good days were associated with more positive features and bad days with more negative features)[18]. Another qualitative study examining reports of good and bad days in patients with chronic obstructive pulmonary disease

found that they were attuned to their symptoms and could chronicle their good and bad days within the context of their own lives[19]. These findings reinforce the relationship of patient perception of good/bad days with severity and impact of symptoms and underscore the importance of the development and validation of our novel Good Day Bad Day single-item assessment for evaluation of nOH.

The potential utility of the Good Day Bad Day instrument is not exclusively limited to nOH. With amended wording, the Good Day Bad Day single-item measure can be easily adapted for use in other conditions in which patients also experience daily fluctuations in their symptoms. A similar approach has been used for other instruments, such as the Work Productivity and Activity Impairment questionnaire [20], which has been specifically adapted and validated for conditions such as Crohn's disease, irritable bowel syndrome, and asthma [21–23]. Additionally, the Good Day Bad Day instrument could also be extended to understand the burden on caregivers of patients with nOH or other diseases. A study of informal caregivers of patients with neurologic disorders (eg, dementia, AD) reported significant daily fluctuations in caregiving burden and stress, which may contribute to negative caregiver outcomes[24]. In addition, such a tool would be easy to use in an Apps or similar with a diary and reminder to fill in the Good Day Bad Day on a daily frequency.

Limitations of our study include that the total number of good and bad days did not add up to 7 for a small proportion of patients (7%); future work should aim to modify this question so that an accurate 7-day report is captured. Also, our validity assessments of the Good Day Bad Day measure did not assess reliability through repeated administration. Although the current study only captured data from patients who continued treatment, our analyses suggest that loss to follow-up was not correlated with number of good days reported. Finally, no qualitative assessment was performed to assess patients' personal criteria for what qualified a day as 'good' or 'bad' and whether responses were solely related to the severity of nOH symptoms on a given day. Interestingly, subsequent data analyses suggest that experiencing a fall may contribute to whether an individual had a good or bad day. In these analyses, we found that having more falls was associated with fewer good days (r , -0.119 to -0.363) and a significant inverse association between the number of good days and the number of falls reported by patients ($P < 0.0001$). Although this association with falls is intriguing and worthy of additional investigation, it is likely that there are other factors (eg, comorbidities,

interaction with family members or caregivers) that influence patients' perception of good/bad days. As such, the Good Day Bad Day item is analogous to other validated, widely accepted PROs that broadly capture QoL and/or functional outcomes (e.g., SF-8,¹³ SDS¹²). However, future research including patient interviews would be beneficial for providing greater understanding of what a good/bad day means to patients in the context of nOH and investigating the meaningfulness of having 1 additional good day each week to patients with nOH, as well as providing further insight regarding the potential possibilities for optimizing the scoring classification to achieve better predictive results and further strengthen the usefulness of this instrument.

Conclusions

This study found that the newly developed Good Day Bad Day single-item instrument provided good discrimination at baseline and longitudinally. The Good Day Bad Day item can aid in assessment of the effects of nOH symptoms on patients' lives, may reduce respondent burden, and has the potential to be adapted to other diseases and conditions. The clinical and personal significance of patient-reported changes in good days merits future research.

Abbreviations

AD: Alzheimer's disease; FES-I: Short Falls Efficacy Scale–International; nOH: Neurogenic orthostatic hypotension; OHQ: Orthostatic Hypotension Questionnaire; OHSA: Orthostatic Hypotension Symptom Assessment; PHQ-9: Patient Health Questionnaire-9; PRO: patient-reported outcome; QoL: quality of life; rho: Spearman rank correlation coefficients; SDS: Sheehan Disability Scale; SF-8: Short-Form Health Survey-8; US FDA: USA Food and Drug Administration

Acknowledgments

The authors thank Christine Sullivan of Lundbeck (Deerfield, IL, USA) for conducting additional analyses. The authors received editorial assistance from Srividya Ramachandran, PhD, Lauren Stutzbach, PhD, and Patrick Little, PhD, employees of ICON (North Wales, PA, USA), which was supported by Lundbeck.

Disclosure statement

CF is an employee of Creativ-Ceutical, was formerly employed by Lundbeck, and is a shareholder of Lundbeck. RM is an employee of Creativ-Ceutical, and NG and VT were employees of Creativ-Ceutical at the time of the validation analyses. LAH and SK are employees of Lundbeck. This work was supported

by Lundbeck; the data reported were derived from a study that was supported by Lundbeck. The sponsor participated in the design of this study, data analysis and interpretation, and in the preparation of the manuscript.

ORCID

Clément François  <http://orcid.org/0000-0003-2278-3352>

References

- [1] Freeman R, Abuzinadah AR, Gibbons C, et al. Orthostatic hypotension: JACC state-of-the-art review. *J Am Coll Cardiol*. 2018;72(11):1294–1309.
- [2] Freeman R, Wieling W, Axelrod FB, et al. Consensus statement on the definition of orthostatic hypotension, neurally mediated syncope and the postural tachycardia syndrome. *Clin Auton Res*. 2011;21(2):69–72.
- [3] Claassen DO, Adler CH, Hewitt LA, et al. Characterization of the symptoms of neurogenic orthostatic hypotension and their impact from a survey of patients and caregivers. *BMC Neurol*. 2018;18(1):125.
- [4] Merola A, Romagnolo A, Rosso M, et al. Autonomic dysfunction in Parkinson's disease: a prospective cohort study. *Mov Disord*. 2018;33(3):391–397.
- [5] Northera® (droxidopa). Full Prescribing Information. Deerfield IL: Lundbeck NA Ltd; 2017.
- [6] Kaufmann H, Malamut R, Norcliffe-Kaufmann L, et al. The Orthostatic Hypotension Questionnaire (OHQ): validation of a novel symptom assessment scale. *Clin Auton Res*. 2012;22(2):79–90.
- [7] US Food and Drug Administration, US Department of Health and Human Services. Guidance for Industry. Patient-reported outcome measures: use in medical product development to support labeling claims. cited 2020 Jul 21. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-reported-outcome-measures-use-medical-product-development-support-labeling-claims>
- [8] US Food and Drug Administration. CDER patient-focused drug development. cited 2020 May 4. <https://www.fda.gov/drugs/development-approval-process-drugs/cder-patient-focused-drug-development>
- [9] European Medicines Agency. Terms of reference for the EMA/FDA cluster on patient engagement. cited 2020 May 4. https://www.ema.europa.eu/en/documents/other/terms-reference-european-medicines-agency/food-drug-administration-cluster-patient-engagement_en.pdf
- [10] François C, Shiao CA, Biaggioni I, et al. Six-month use of droxidopa for neurogenic orthostatic hypotension. *Mov Disord Clin Pract*. 2019;6(3):235–242.
- [11] Kempen GI, Yardley L, van Haastregt JC, et al. The Short FES-I: a shortened version of the falls efficacy scale-international to assess fear of falling. *Age Ageing*. 2008;37(1):45–50.
- [12] Leon AC, Olfson M, Portera L, et al. Assessing psychiatric impairment in primary care with the Sheehan Disability Scale. *Int J Psychiatry Med*. 1997;27(2):93–105.
- [13] Ware JE Jr, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36). I. Conceptual framework and item selection. *Med Care*. 1992;30(6):473–483.
- [14] Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16(9):606–613.
- [15] Göktaş A, İçi Ö. A comparison of the most commonly used measures of association for doubly ordered square contingency tables via simulation. *Metodološki zvezki*. 2011;8(1):17–37.
- [16] Tang W, Hu J, Zhang H, et al. Kappa coefficient: a popular measure of rater agreement. *Shanghai Arch Psychiatry*. 2015;27(1):62–67.
- [17] Altman DG. Some common problems in medical research. In: Altman DG, editor. *Practical statistics for medical research*. London UK: Chapman & Hall/CRC; 1991. p. 396–439.
- [18] Rockwood K, Fay S, Hamilton L, et al. Good days and bad days in dementia: a qualitative chart review of variable symptom expression. *Int Psychogeriatr*. 2014;26(8):1239–1246.
- [19] Williams V, Ryan S. "I just know": exploring self-knowledge in chronic obstructive pulmonary disease. *Palgrave Commun*. 2017;3(1):17089.
- [20] Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. *Pharmacoeconomics*. 1993;4(5):353–365.
- [21] Reilly MC, Gerlier L, Brabant Y, et al. Validity, reliability, and responsiveness of the work productivity and activity impairment questionnaire in Crohn's disease. *Clin Ther*. 2008;30(2):393–404.
- [22] Reilly MC, Bracco A, Ricci JF, et al. The validity and accuracy of the Work Productivity and Activity Impairment questionnaire-Irritable Bowel Syndrome version (WPAI:IBS). *Aliment Pharmacol Ther*. 2004;20(4):459–467.
- [23] Chen H, Blanc PD, Hayden ML, et al. Assessing productivity loss and activity impairment in severe or difficult-to-treat asthma. *Value Health*. 2008;11(2):231–239.
- [24] Pihet S, Moses Passini C, Eicher M. Good and bad days: fluctuations in the burden of informal dementia caregivers, an experience sampling study. *Nurs Res*. 2017;66(6):421–431.