SymTrak-8 as a Brief Measure for Assessing Symptoms in Older Adults

Patrick O. Monahan, PhD[1,2], Kurt Kroenke, MD[1,3,4], and Timothy E. Stump, MA[1]

[1]School of Medicine, Indiana University, 410 W. Tenth St., Suite 3000, Indianapolis, IN, USA; [2]Fairbanks School of Public Health, Indiana University, Indianapolis, IN, USA; [3]Center for Health Information and Communication, VA HSR&D, Washington DC, USA; [4]Regenstrief Institute, Inc., Indianapolis, IN, USA.

BACKGROUND: Patient- and caregiver-reported 23-item SymTrak scales were validated for monitoring clinically actionable symptoms and impairments associated with multiple chronic conditions (MCCs) in older adults. Items capture physical and emotional symptoms and impairments in physical and cognitive functioning. An abbreviated SymTrak is desirable when response burden is a concern.

OBJECTIVE: Develop and validate the 8-item SymTrak.

DESIGN AND PARTICIPANTS: Secondary analysis of SymTrak validation study; 600 participants (200 patient-caregiver dyads; 200 patients without an identified caregiver).

MAIN MEASURES: Demographic questions, SymTrak, and Health Utility Index Mark 3 (HUI3).

KEY RESULTS: SymTrak-8 demonstrated good fit to a one-factor model using confirmatory factor analysis (CFA). Concurrent criterion validity was supported by high standardized linear regression coefficients (STB) between baseline SymTrak-8 total score (independent variable) and baseline HUI3 preference-based overall HRQOL utility score (dependent variable; 0 = death, 1 = perfect health), after adjusting for demographics, comorbid conditions, and medications, with strength comparable to SymTrak-23 (STB = − 0.81 and − 0.84, respectively, for SymTrak-8 and SymTrak-23, when patient-reported; and − 0.60 and − 0.62, respectively, when caregiver-reported). Coefficient alpha (0.74; 0.76) and 24-h test–retest reliability (0.83; 0.87) were high for SymTrak-8 for patients and caregivers, respectively. The convergent correlation between brief and parent SymTrak scales was high (0.94). SymTrak-8 demonstrated approximate normality and a linear relationship with SymTrak-23 and HUI3. Importantly, a 3-month change in SymTrak-8 was sensitive to detecting the criterion (3-month reliable change categories; improved, stable, declined in HUI3 overall utility), with results comparable to SymTrak-23.

CONCLUSIONS: SymTrak-8 total score demonstrates internal reliably, test–retest reliability, criterion validity, and sensitivity to change that are comparable to SymTrak-23. Thus, patient- or caregiver-reported SymTrak-8 is a viable option for identifying and monitoring the aggregate effect of symptoms and functional impairments in patients with multimorbidity when response burden is a concern.

KEY WORDS: primary care; psychometrics; multimorbidity; aging; scale; symptoms.

J Gen Intern Med
DOI: 10.1007/s11606-020-06329-5

Prior Presentations *None.*

Supplementary Information *The online version contains supplementary material available at https://doi.org/10.1007/s11606-020-06329-5.*

Received April 25, 2020
Accepted October 16, 2020

INTRODUCTION

SymTrak-23 is a measure recently validated for identifying and monitoring clinically actionable symptoms and impairments in patients attending primary care with multimorbidity.[1,2] Multiple chronic conditions (MCCs),[3] complex health care needs,[4] and symptoms (physical, emotional, and cognitive)[5,6] are common clinical presentations among elderly patients. Symptoms account for over half of all US outpatient visits annually[7] and predict health care utilization and costs, quality of life, work disability, and mortality.[5,6,8,9]

SymTrak-23 was developed to target the most prevalent and disabling symptoms and functional impairments experienced by older adults, including SPADE symptoms (**s**leep disturbance, **p**ain, **a**nxiety, **d**epression, [low] **e**nergy/fatigue) as well as impairments related to mobility, cognition, and vision or hearing. Moreover, these domains were selected because they are potentially clinically actionable in terms of evaluation and management. Both *patient* and *caregiver* versions of SymTrak-23 were validated[1,2] because up to 57% of elderly patients are accompanied by an informal caregiver,[10] and some older patients are unwilling or unable to complete self-report measures, meriting proxy measures.[11]

The present paper develops and validates an abbreviated 8-item SymTrak scale (SymTrak-8).[1,2] Internal and test–retest reliability as well as criterion validity, convergent validity, and sensitivity to change were assessed. Our goal was to address the same domains covered by SymTrak-23. Thus, the total score for SymTrak-8 and SymTrak-23 is a measure of the overall burden of symptoms and functional impairments associated with multimorbidity in older adults. Secondarily, the total score can be used as an indicator of health-related quality of life (HRQOL) by health care systems and clinical researchers.

METHODS

Setting and Sample

The 600 participants (200 patient-caregiver dyads and 200 non-dyadic patients without an identified caregiver) were recruited from an academic-affiliated primary care network of clinics. The study was approved by the institutional review board and all participants provided written informed consent. Patient inclusion criteria were as follows: (1) age ≥ 65 years, (2) ≥ one primary care visit in the past 12 months, (3) ≥ one chronic condition according to medical records, and (4) for those participants who had an identified informal caregiver, the caregiver had to be ≥ 21 years of age and willing to participate in the study. Patient exclusion criteria were as follows: permanent residency at a long-term care facility or the presence of a severe mental illness such as bipolar disorder or schizophrenia.

Measures

A brief survey, consisting of demographics, SymTrak-23, and the Health Utility Index Mark 3 (HUI3)[12] was completed by participants ($N = 600$) by interview at baseline and 3 months post-baseline. A subsample ($n = 180$) consisting of 60 patient-caregiver dyads and 60 non-dyadic individual patients completed an interview 24 h after baseline. All interviews were telephone-administered by research assistants.

The HUI3 is a preference-based measure of HRQOL; the HUI3 overall utility score is a continuous variable ranging from 0 to 1, where 0 represents death and 1 represents perfect health.[12] SymTrak-23 consists of 23 items that measure symptoms and functional impairments. The item-response options are as follows: 0 = never, 1 = sometimes, 2 = often, 3 = always. Two general health items (Poor, Fair, Good, Very Good, Excellent), rated separately for physical and emotional health, were also administered to assess construct validity. All scales (SymTrak-23, HUI3, general health ratings) were self-reported by patients and proxy-reported by caregivers about their dyadic patient's health status.

Analysis

Item-level psychometrics were examined, including floor and ceiling effects, item-total correlations, and item-level validity correlations with the criterion of HUI3 overall utility score. Item-level 24-h test–retest reliability was estimated for ordinal items using weighted kappa with Fleiss–Cohen quadratic weights. Confirmatory factor analysis (CFA) of the hypothesized one-factor model was performed using MPLUS software (8th Edition)[13] with the nonlinear probit link for ordinal categorical items and weighted least square mean and variance adjusted (WLSMV) estimation. The following CFA fit indices and thresholds were used for indicating good fit[14]: root mean square error of approximation (RMSEA) < 0.06, comparative fit index (CFI) > 0.95, and standardized root mean square residual (SRMR) < 0.08. The chi-square goodness-of-fit test was also conducted.

Concurrent predictive criterion validity was assessed by using linear regression to test the association between the baseline SymTrak total scale score (independent variable) and the baseline HUI3 overall HRQOL utility score (dependent variable) while adjusting for covariates. The standardized regression coefficient (STB) was reported as an effect size. Cronbach's alpha was used to assess internal consistency reliability. The absolute-agreement intra-class correlation coefficient (ICC) was used to assess scale score test–retest reliability, while specifying occasions as a random effect.

Sensitivity to change was examined using baseline and 3-month data. The within-group mean change and standardized response mean (SRM) effect size (mean change divided by SD of change) were computed for SymTrak total scores, separately for each of three "reliable" HRQOL change groups. These HRQOL criterion groups (declined, stable, or improved) were formulated based on ± 1 standard error of measurement (SEM) in the HUI3 overall utility score. The relevant 1 SEM value (0.089) for HUI3 for this analysis was calculated using the temporal stability approach, based on 24-h test–retest reliability and the SD of 3-month change scores.[2] A general linear model (GLM) was used to compare SymTrak change scores. If the GLM omnibus F test (for the overall difference between all three HUI3 change groups) was significant, all pairwise differences (e.g., declined vs stable) were tested using Fisher's *protected* least significant difference method which controls experiment-wise type I error at 0.05 when the number of groups is only three.[15] All tests were conducted as two-sided using 0.05 alpha.

RESULTS

Participant Characteristics

Demographic characteristics of patients and caregivers have been previously reported.[2] Briefly, patients had a median age of 73 (range, 65 to 95), and 73% were women. Patients and caregivers were diverse demographically (e.g., 49% African American). Most caregivers were either a child (43%) or a spouse/partner (36%).

Content Validity

Content validity for the development of SymTrak-23 is described elsewhere.[2] The process of shortening a scale should be undertaken with domain experts to guide the narrowing of content and to assess redundancy and usefulness for the intended purpose.[16,17] Eight items (4 original and 4 that each contain bundled content from 2 original SymTrak-23 items) were considered ideal candidates by our content experts (all authors) for an abbreviated scale because each item (one for each domain) has the content validity for capturing 8 important domains: the SPADE symptom domains (**s**leep disturbance, **p**ain, **a**nxiety, **d**epression, and low **e**nergy), cognition, mobility, and vision/hearing problems (Table 1, bolded items).

Table 1 SymTrak Item-Level Psychometrics (Baseline, Patients, $N = 400$)

SymTrak Items*	Test–retest weighted kappa†	% floor‡	% ceiling‡	Corr§ with 23-item total	Corr§ with HUI3
1. Feeling tired or having low energy	0.58	14.0	10.0	0.58	−0.57
2. Trouble falling asleep or trouble staying asleep	0.65	34.8	10.5	0.44	−0.32
3. Pain interfering with daily activities	0.64	29.0	17.5	0.62	−0.69
4. Pain in the back, arms, legs, or joints	0.74	14.0	26.8	0.64	−0.64
5. Foot pain or foot numbness	0.72	34.8	14.3	0.47	−0.50
6. Constipation or stomach problems	0.69	42.0	5.0	0.50	−0.40
7. Trouble with urination	0.72	70.8	3.8	0.32	−0.35
8. Shortness of breath	0.79	39.8	8.0	0.58	−0.50
9. Chest pain	0.72	72.8	0.8	0.39	−0.36
10. Trouble with vision	0.76	38.9	12.1	0.46	−0.43
11. Trouble with hearing	0.74	48.8	10.0	0.38	−0.51
10 or 11. Trouble with vision or hearing	0.70	22.8	6.0	0.46	−0.57
12. Trouble walking or trouble moving around	0.64	30.8	12.5	0.60	−0.60
13. Falling or tripping	0.65	74.3	0.5	0.37	−0.25
14. Less interest or less pleasure in doing things	0.54	43.4	4.0	0.62	−0.42
15. Feeling sad, down, or depressed	0.64	55.0	2.0	0.54	−0.29
14 or 15. Feeling sad, down, depressed, or having less interest in doing things	0.61	34.3	2.0	0.63	−0.39
16. Poor appetite or overeating	0.52	44.8	5.0	0.49	−0.31
17. Feeling nervous or anxious	0.53	49.5	3.0	0.50	−0.29
18. Worrying too much about different things	0.73	39.0	7.0	0.56	−0.32
17 or 18. Feeling nervous, anxious, or worrying too much	0.72	29.8	2.0	0.60	−0.35
19. Becoming easily annoyed or irritable	0.63	47.9	2.5	0.47	−0.24
20. Trouble taking medications in the right dose at the right time	0.39	76.0	2.5	0.27	−0.12
21. Trouble remembering appointments	0.73	69.0	2.0	0.42	−0.25
22. Trouble concentrating on things	0.66	45.3	2.5	0.54	−0.32
23. Memory Loss	0.61	39.3	2.3	0.47	−0.25
22 or 23. Trouble concentrating on things or memory loss	0.62	27.0	2.3	0.56	−0.33

Bolded items were selected for SymTrak-8 because of good item psychometrics and they represent prevalent, disabling, and undertreated domains of symptoms and functional impairment in elderly patents attending primary care, including **sleep disturbance, **p**ain interference, **a**nxiety, **d**epression, and low **e**nergy/fatigue (i.e., SPADE symptoms); cognition; mobility-related physical functioning; and impairments with vision or hearing. Each bolded italic bundled item combines the essential content of 2 original SymTrak-23 items*

†Sample size was restricted to those in the test–retest sample (122 patients)

‡The % floor and % ceiling = percentage of persons endorsing "never" or "always," respectively

§Values in the "Corr" columns are Spearman's rank correlation coefficients between each item and either the SymTrak-23 total score or the HUI3 overall utility score, which are measures of item discrimination and construct validity, respectively

Specifically, the SPADE symptom domains represent five of the most prevalent, chronic, disabling, and undertreated symptom domains; they also tend to co-occur in both the general population[18] and in clinical practice.[18–26] Cognitive impairment is an important 6th domain because of its prevalence in older adults, comorbidity and adverse effects on other illnesses, impact on poor adherence to prescribed therapies, and worse survival.[11] Mobility-related functioning (7th domain) has important health and social consequences in multimorbidity populations and has value as a predictor of nursing home placement and mortality.[11] The falling or tripping item had a high floor effect, a common limitation when measuring physical functioning;[11] therefore, the more general and prevalent item, walking or moving around, was selected. Finally, the 8th domain (vision or hearing loss) comprises two sensory impairments that are prevalent in geriatric primary care and are highly treatable,[27] often comorbid,[28] and associated with poorer HRQOL.[12,28,29]

Each of the 4 bundled items included the essential content from 2 original SymTrak-23 items according to established principles.[30] The score for each bundled item was calculated as the *average* of its two constituent item scores and then rounded to an integer (i.e., 0–3). In a sensitivity analysis, using the *maximum* produced similar psychometric results.

Item-Level Reliability and Validity

Item-level psychometric performance is shown for patient-reported data in Table 1. The 8 bolded items (i.e., SymTrak-8), initially selected based on content validity, demonstrated comparable psychometrics compared to the non-selected items, specifically, for test–retest reliability, item correlations with the original 23-item total, and criterion validity correlations with the HUI3 overall utility score (Table 1). The 4 bundled items demonstrated similar or better results for test–retest reliability, item-total correlations, and criterion validity (i.e., correlations with HUI3) compared to their original component items (Table 1). Importantly, bundled items exhibited lower ceiling effects and markedly lower floor effects than their original items. Results were similar for caregiver-reported data (Online Appendix 1).

Table 2 Confirmatory Factor Analysis (CFA) of One-Factor Model

SymTrak-8 (and item no. from SymTrak-23)	Patients ($n = 400$)	Caregivers ($n = 203$)
	Loadings	Loadings
1. Feeling tired or having low energy (no. 1)	0.68	0.71
2. Trouble falling asleep or trouble staying asleep (no. 2)	0.46	0.47
3. Pain interfering with daily activities (no. 3)	0.61	0.65
4. Trouble with vision or hearing (no. 10 or no. 11)	0.35	0.31
5. Trouble walking or trouble moving around (no. 12)	0.58	0.70
6. Feeling sad, down, depressed, or having less interest in doing things. (no. 14 or no. 15)	0.77	0.75
7. Feeling nervous, anxious, or worrying too much (no. 17 or no. 18)	0.66	0.62
8. Trouble concentrating on things or memory loss (no. 22 or no. 23)	0.57	0.53
Model fit statistics*		
Number of free parameters	32	32
Goodness-of-fit (GOF) chi-square (*df*)	75.8 (20)	41.8 (20)
GOF chi-square *p* value	< 0.0001	0.003
RMSEA (90% confidence interval)	0.084 (0.064, 0.104)	0.073 (0.042, 0.105)
CFI	0.950	0.964
SRMR	0.044	0.047

**The goodness-of-fit (GOF) chi-square test is a hypothesis test of the discrepancy between the observed and fitted covariance matrices; $p < 0.05$ indicates rejection of the GOF null hypothesis of perfect fit (i.e., rejection of hypothesis that all population residuals equal zero). Fit indices serve as effect sizes of model misfit and supplement the GOF chi-square test. Like the GOF chi-square test, the standardized root mean squared residual (SRMR) and the root mean square error of approximation (RMSEA) assess how well the hypothesized model reproduces the sample data; they are absolute fit indices because they do not assess incremental fit with respect to a reference model. The SRMR equals the square root of the mean of the squared residuals between the observed and fitted covariance matrices; values < 0.08 indicate good fit. The RMSEA measures the discrepancy between the fitted and population covariance matrices per degree of freedom. The RMSEA compensates for model complexity and ranges from 0 to 1; values < 0.06 indicate good fit. The comparative fit index (CFI) is an incremental fit index measuring the proportionate improvement provided by the fitted model compared to the reference or baseline model. The baseline model is typically, as in this CFA analysis, the "empty" no-factor model in which all observed variables are uncorrelated. The CFI ranges from 0 to 1; values > 0.95 indicate good fit. The RMSEA and CFI are based on the noncentrality parameter of the GOF chi-square statistic. Additional information about these fit indices and justification for these thresholds are described in Hu and Bentler (1999)*

Factorial Validity

The one-factor CFA model fit the 8 items reasonably well (Table 2). Perfect fit is often rejected by the chi-square test, as it was here, in non-small samples due to adequate power to detect minor misfit.[14] The magnitude of misfit was acceptable based on fit indices.[14] Two fit indices (CFI, SRMR) met or exceeded their "good fit" thresholds, and RMSEA was near its < 0.06 threshold. The loadings were above 0.45 for all items (except for the "vision or hearing" item for which loadings were above 0.30) for both patient- and caregiver-reported data (Table 2). MPLUS simulations, based on realistic threshold parameters estimated from the data, revealed power > 97% for one-factor CFA to detect 0.40 population loadings for both patient- and caregiver-reported data. The scree plot indicated that a single dominant factor explains SymTrak-8's inter-item correlations reasonably well, supporting the validity of using the SymTrak-8 total score as an overall measure of symptoms and functional impairments (Fig. 1).

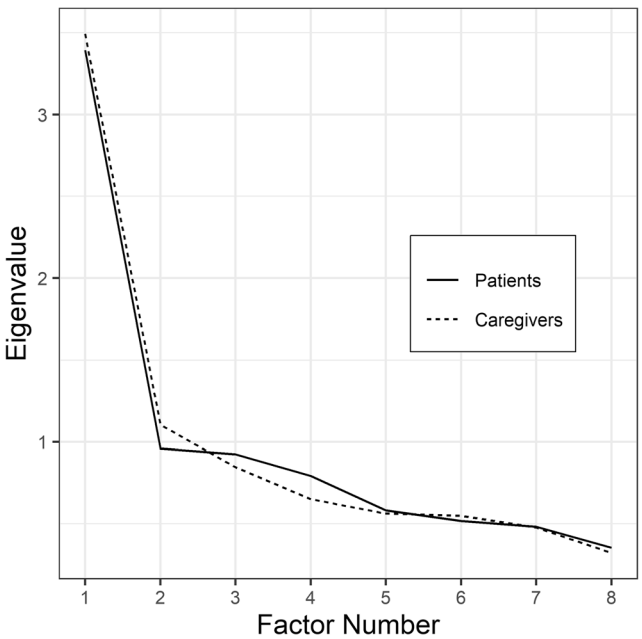

Fig. 1 Factor analysis scree plot.

Criterion Validity

The SymTrak-8 total score was a strong cross-sectional predictor of overall HRQOL utility (HUI3), with results comparable to the SymTrak-23 total score (Table 3). After adjusting for patients' baseline demographics (analysis 1), a 1 SD increase in SymTrak-8 and SymTrak-23, respectively, were associated with a 0.83 SD and 0.84 SD decrease in HUI3 utility score. After also adjusting for baseline medical comorbidities and medications (analysis 2), the abbreviated SymTrak remained nearly as strongly associated with HUI3 as the original SymTrak (STB = − 0.81 and − 0.84, respectively).

Table 3 SymTrak as a Concurrent Predictor of the Health Utility Index (HUI3) Criterion

SymTrak scale	Standardized beta;* analysis 1 (adjusted for demographics)†		Standardized beta;* analysis 2 (adjusted for demographics, comorbidities, medications)‡	
	Patients	Caregivers	Patients	Caregivers
	(*N* = 390)	(*N* = 177)	(*N* = 301)	(*N* = 105)
SymTrak-23 total score	−0.84	−0.70	−0.84	−0.62
SymTrak-8 total score	−0.83	−0.70	−0.81	−0.62

All demographics in analysis 1 and analysis 2 for Table 3 were characteristics of patients and reported by patients. The SymTrak and HUI3 instruments were self-reported by patients in the Patients models and proxy-reported by caregivers (about their dyadic patient's health status) in Caregivers models. The sample size for each model was based on available data; e.g., the Caregivers model in analysis 2 required patient-reported demographics, comorbidities, and medications, as well as caregiver-reported scale scores (SymTrak and HUI3) from intact dyads
**Standardized regression coefficient (beta) for each column and row is from a separate multivariable linear regression model in which the dependent variable is baseline HUI3 overall utility score and the primary independent variable is baseline SymTrak total score, adjusted for covariates*
†For analysis 1, models were adjusted for demographics including age, sex, race (Black vs White; Other vs White), highest level of education (high school [HS] graduate vs < HS; some college or higher vs < HS), total household income ($15,000–$30,000 vs < $15,000; > $30,000 vs < $15,000; unknown vs < $15,000), and marital status (married/living together vs not)
‡For analysis 2, models were adjusted for these same demographics as well as the number of medical comorbidities and the number of medication classes

Results for SymTrak-8 and SymTrak-23 were also strong for caregiver-reported data, with comparable strength for the two scales (Table 3). The SymTrak-8 total score had an approximately linear relationship with the HUI3 overall HRQOL utility score for both patients and caregivers (Online Appendix 2).

Scale Distribution Features and Reliability

Very few participants scored the lowest (0.3 to 2.5%) or highest (0%) possible total score for SymTrak-8 or SymTrak-23, indicating no problems with floor or ceiling effects (Table 4). The percentage of respondents missing any item was low (i.e., ≤ 4%). The SymTrak-8 total score was approximately normally distributed (Online Appendix 3). Coefficient alpha for SymTrak-8 was 0.74 and 0.76 for patient- and caregiver-reported scores, respectively, and test–retest reliability was high, with ICCs of 0.83 and 0.87 (Table 4).

Convergent Validity

The Pearson correlation between SymTrak-8 and SymTrak-23 total scores was 0.94, separately for caregivers and patients. Because the two scales contain overlapping item content, the Pearson correlation was also calculated between the SymTrak-8 total score and the "remaining" total score, the latter formed by summing 11 items from SymTrak-23 after excluding the 12 items that contributed content to SymTrak-8. As expected, the correlation between the brief and parent scales was reduced after excluding overlapping items but was still of moderate-to-high magnitude (0.75, separately for patients and caregivers). The relationships were markedly linear between SymTrak-8 total and both the SymTrak-23 total and the "remaining" total (Online Appendix 4). A generally monotonic decreasing pattern of overall symptom severity (measured by SymTrak-8 or SymTrak-23 total score) was observed across general health ratings (Online Appendix 5).

Table 4 SymTrak Scale Score Distribution Features and Reliability

Scale characteristics	Patients (*N* = 400)		Caregivers (*N* = 203)	
	SymTrak-23	SymTrak-8	SymTrak-23	SymTrak-8
Baseline scale score distribution				
Mean	17.9	8.3	18.3	8.6
Median	17.0	8.0	19.0	9.0
SD	9.2	4.0	9.9	4.4
Possible score range	0–69	0–24	0–69	0–24
Observed score range	0–48	0–20	0–50	0–20
% floor (lowest possible score)	1 (0.3%)	4 (1%)	2 (1%)	5 (2.5%)
% ceiling (highest possible score)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
n (%) missing any item	4 (1%)	0 (0%)	8 (4%)	2 (1%)
n (%) not computable (> 50% items missing)*	0 (0%)	0 (0%)	1 (0.5%)	1 (0.5%)
Reliability				
Cronbach's Coefficient Alpha	0.85	0.74	0.86	0.76
Test–retest (24-h) reliability (ICC)	0.87	0.83	0.91	0.87

ICC absolute-agreement intra-class correlation coefficient, with occasions specified as a random effect
**Only one caregiver had their scale score set to missing due to missing more than 50% of SymTrak-8 or SymTrak-23 items (i.e., the rule we recommend when computing SymTrak scale scores)*

Table 5 Sensitivity to Change of SymTrak Scales

	HRQOL 3-month reliable change groups*						*p* values			
	HUI3 declined (*D*)		HUI3 stable (*S*)		HUI3 improved (*I*)					
SymTrak scales†	*M* (SD)	SRM	*M* (SD)	SRM	*M* (SD)	SRM	Omnibus	*D* vs. *S*	*D* vs. *I*	S vs. I
Patient SymTrak-23	− 1.07 (5.30)	− 0.20	0.56 (5.17)	0.11	2.03 (6.01)	0.34	< 0.001	0.021	< 0.0001	0.046
Patient SymTrak-8	− 0.52 (2.61)	− 0.20	0.24 (2.70)	0.09	1.12 (2.93)	0.38	< 0.0001	0.031	< 0.0001	0.017
Caregiver SymTrak-23	− 4.33 (7.33)	− 0.59	− 0.53 (6.17)	− 0.09	2.25 (6.36)	0.35	< 0.0001	0.003	< 0.0001	0.023
Caregiver SymTrak-8	− 1.64 (3.18)	− 0.52	− 0.52 (2.88)	− 0.18	0.75 (3.12)	0.24	< 0.001	0.057	< 0.001	0.025

M (SD) mean (standard deviation), SRM standardized response mean. The SRM is an effect size calculated as the mean change for SymTrak total score (baseline minus 3 months) divided by SD of change scores. A positive effect size indicates a decrease in the SymTrak symptom total score which represents improvement in patients' symptoms. A negative effect size represents a worsening of symptoms

**HUI3 overall utility score was used to measure preference-based health-related quality of life (HRQOL), for which a higher score is better. The HUI3 declined, stable, and improved groups were defined using a threshold of ± 1 three-month standard error of measurement (SEM) of change on HUI3 overall utility score (i.e., a change of 0.089)*

†Each row of values is from a separate general linear model (GLM) where, for the purpose of this analysis, HUI3 reliable change groups was the independent variable and change score of the SymTrak total score was the dependent variable. The number of individuals in the HUI3 Declined, Stable, and Improved groups were 113, 128, and 99, respectively, for patient-reported data, and 45, 66, and 53, respectively, for caregiver-reported data

Sensitivity to Change

The SRM represents the magnitude of change in SymTrak-8 (and SymTrak-23 for comparative purposes) over 3 months (Table 5). The SRM was approximately centered near 0, as hypothesized, for the HUI3 Stable group, and in the anticipated negative and positive direction, respectively, for the groups of patients that declined or improved in HUI3-based HRQOL.

For patient- and caregiver-reported data, the change scores for both SymTrak-8 and SymTrak-23 were sensitive to detecting overall differences between the three HRQOL change groups (Table 5, omnibus F test, $p < 0.001$). For pairwise differences, the scales performed comparably, in general. When self-reported, both scales markedly differentiated the declined and improved groups ($p < 0.0001$), and also significantly ($p < 0.05$) distinguished the other two pairwise group differences. When proxy-reported by caregivers, both SymTrak-8 and SymTrak-23 significantly differentiated the declined and improved groups, and the stable and improved groups. However, there was one nuanced difference between the abbreviated and original scales. Caregiver-reported SymTrak-23 demonstrated stronger separation ($p = 0.003$) of the declined and stable groups than the marginal separation observed for caregiver-reported SymTrak-8 ($p = 0.057$; Table 5).

DISCUSSION

The SymTrak-8 Patient and Caregiver Forms (Online Appendices 6 and 7) have identical items and differ only in the opening stem. Consistent with SymTrak-23 findings,[2] a good fit to the hypothesized one-factor model was confirmed for SymTrak-8. The total score for SymTrak-8, like that of SymTrak-23,[2] was approximately normally distributed and showed an adequate range as well as negligible floor and ceiling effects. Internal and test–retest reliability were also good (> 0.70)[31] for SymTrak-8, although as expected were slightly higher for SymTrak-23.

Regarding convergent validity, SymTrak-8 serves as a brief and reasonable approximation to its parent scale; they explain 88% of the variance in each other's total score ($r = 0.94$) and their relationship is remarkably linear. The moderate-to-high correlation ($r = 0.75$) between SymTrak-8 total and the "remaining" total of SymTrak-23 supports construct validity. Specifically, the magnitude of shared variance (56%) suggests adequate conceptual similarity, which supports the use of SymTrak-8, while the amount of non-shared variance (44%) indicates sufficient uniqueness, which justifies SymTrak-23 for capturing additional information when response burden is not a concern. It should be noted that SymTrak-23 is itself relatively short compared to traditional instruments that capture multiple domains of symptoms and functional impairments.

HUI3-based criterion validity was previously reported for SymTrak-23 after adjusting for patient demographics.[2] The present analyses found that both SymTrak-8 and SymTrak-23 demonstrate excellent HUI3-related criterion validity even after also adjusting for comorbidities and medications. SymTrak-8's monotonic relationship with physical and emotional general health ratings were consistent with findings for SymTrak-23.[2] Because the total score for both scales is a unidimensional measure of overall symptom burden and functional impairment, an analysis of even greater relevance, not previously published for SymTrak-23, is the relationship between these SymTrak total scores and the *composite* general health rating, for which both scales demonstrated strong monotonic relationships. This provides further support for the construct validity of both SymTrak-8 and SymTrak-23.

The SymTrak-8 total score also revealed good sensitivity to change. Its effect size (i.e., SRM) and significance for detecting HUI3-based HRQOL change were similar to those of SymTrak-23. However, SymTrak-23 was slightly more sensitive to detecting declining versus stable HRQOL when the patients' symptoms were reported by caregivers, making SymTrak-23 possibly preferred over SymTrak-8 when longitudinal monitoring is captured through proxies, especially

during periods of suspected declining health status and when the longer scale is not burdensome to proxy respondents.

Multimorbidity, an alternative term for MCCs, has been an area of increasing research.[32–35] Particularly relevant in older adults, multimorbidity causes substantial impairment as well as increased health care utilization and costs.[36–38] The substantial adverse impact of SPADE and other symptoms, as well as physical and cognitive functional impairments, which are captured by SymTrak-8, has been demonstrated among patients with multimorbidity in numerous studies.[39–53] Interventions to optimize care in patients with multimorbidity have been recently emphasized,[54,55] and to this end, SymTrak-8 and SymTrak-23 may also be useful.

A strength of this study was that we assessed several different aspects of reliability and validity.[16] In particular, when scales are intended for longitudinal application, sensitivity to change is an essential source of validity evidence to be evaluated.[16,56] Furthermore, only 20% of published articles pertaining to shortening scales tend to use confirmatory instead of exploratory factor analysis.[16] Moreover, test–retest reliability is often not assessed due to the challenges of rapidly re-administering measures. Furthermore, when available, published retest subsamples are often smaller (e.g., $n = 30$) than those in our study (120 patients and 60 caregivers).

Limitations

Generalizability of SymTrak-8 and SymTrak-23 should be studied for younger (age < 65) patients with MCCs. Sensitivity to change for these abbreviated and parent scales should be investigated over a longer interval than 3 months to allow more time for (1) responsiveness to treatments prescribed during routine primary care or (2) accumulating deleterious effects of MCCs. Responsiveness to interventions in randomized controlled trials for populations with multimorbidity would further substantiate sensitivity to change.

All scales in this study were researcher-administered by telephone. Importantly, psychometrics were also strong for patient- and caregiver-reported SymTrak-23 when it was self-administered in clinics by paper and pencil during the pilot study, including high Cronbach's alpha, high usability ratings, and brief administration time (average of 3 min).[1] SymTrak-8 was developed using the same data set used to develop and validate SymTrak-23.[1,2] The psychometrics of SymTrak-8 and SymTrak-23 should be further investigated in an independent sample.

Research and Clinical Implications

The SymTrak-8 total score, like that of SymTrak-23,[2] is approximately normally distributed and has a remarkably linear relationship when used in linear regression to predict the HUI3-based HRQOL overall utility score. These are advantages in parametric models for satisfying normality when SymTrak total score is used as an outcome measure (dependent variable) or for satisfying linearity when SymTrak total score is used as an independent variable to predict HRQOL utility. Future research could determine whether SymTrak's linear relationship with HUI3 holds with other HRQOL utility questionnaires such as the EQ-5D.[57,58]

Although SymTrak-23 can serve as the full measure in certain research and clinical settings, SymTrak-8 may have a broader reach in busy primary care practice settings, as well as research studies, when (1) response burden is a concern or (2) the aggregate effect of symptoms and deficits is either a secondary outcome or a covariate. Moreover, its sensitivity to change can be valuable in monitoring treatment outcomes in trials or practice. The cumulative effect of symptoms and functional impairments tapped by SymTrak total score (8- or 23-item) is relevant to intervention research and clinical practice among multimorbidity populations, given that some common treatments tend to synergistically affect multiple conditions and symptoms simultaneously.[20,21,59,60]

Corresponding Author: *Patrick O. Monahan, PhD; School of Medicine, Indiana University, 410 W. Tenth St., Suite 3000, Indianapolis, IN 46202-3002, USA (e-mail: pmonahan@iu.edu).*

Funders *This work was supported by a National Institute on Aging R01 award to P.O.M. (R01 AG043465). The sponsor had no role in study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the article for publication. The views expressed in this article are those of the authors and do not necessarily represent the views of the National Institute on Aging.*

Compliance with Ethical Standards:

Conflict of Interest: *All authors have no financial or non-financial interests. Patrick O. Monahan is Chief Technology Officer and has 3% equity ownership (valued at $3000) in a for-profit company called RestUp. The purpose of RestUp is to use internet and mobile technology to connect caregivers and care seekers. The RestUp caregivers are paid hourly, as 1099 contractors, by care seekers, and RestUp earns its income by receiving a percentage of each hour worked. The present paper has no overlap with the RestUp company; the SymTrak tool developed in the paper is not used in the RestUp company; and none of the activities of the RestUp company are involved in any way with this paper or the SymTrak tool.*

REFERENCES

1. **Monahan PO**, **Kroenke K**, **Callahan CM**, **et al**. Development and feasibility of SymTrak, a multi-domain tool for monitoring symptoms of older adults in primary care. J Gen Intern Med. 2019;34(6):915-922.
2. **Monahan PO**, **Kroenke K**, **Callahan CM**, **et al**. Reliability and validity of SymTrak, a multi-domain tool for monitoring symptoms of older adults with multiple chronic conditions. J Gen Intern Med. 2019;34(6):908-914.
3. **Schneider KM**, **O'Donnell BE**, **Dean D**. Prevalence of multiple chronic conditions in the United States' Medicare population. Health Qual Life Outcomes. 2009;7:82.
4. **Leroy L**, **Bayliss E**, **Domino M**, **et al**. The Agency for Healthcare Research and Quality Multiple Chronic Conditions Research Network: overview of research contributions and future priorities. Med Care. 2014;52 Suppl 3:S15-22.

5. **Sha MC**, **Callahan CM**, **Counsell SR**, **Westmoreland GR**, **Stump TE**, **Kroenke K**. Physical symptoms as a predictor of health care use and mortality among older adults. Am J Med. 2005;118(3):301-306.
6. **Callahan CM**, **Weiner M**, **Counsell SR**. Defining the domain of geriatric medicine in an urban public health system affiliated with an academic medical center. J Am Geriatr Soc. 2008;56(10):1802-1806.
7. **Kroenke K.** Studying symptoms: sampling and measurement issues. Ann Intern Med. 2001;134(9 Pt 2):844-853.
8. **Kessler RC**, **Zhao S**, **Katz SJ**, et al. Past-year use of outpatient services for psychiatric problems in the National Comorbidity Survey. Am J Psychiatry. 1999;156:115-123.
9. **Kroenke K.** Patients presenting with somatic complaints: epidemiology, psychiatric comorbidity and management. Int J Methods Psychiatr Res. 2003;12:34-43.
10. **Adelman RD**, **Greene MG**, **Ory MG**. Communication between older patients and their physicians. Clin Geriatr Med. 2000;16(1):1-24, vii.
11. Working Group on Health Outcomes for Older Persons with Multiple Chronic Conditions. Universal health outcome measures for older persons with multiple chronic conditions. J Am Geriatr Soc. 2012;60(12):2333-2341.
12. **Horsman J**, **Furlong W**, **Feeny D**, **Torrance G**. The Health Utilities Index (HUI): concepts, measurement properties and applications. Health Qual Life Outcomes. 2003;1:54.
13. **Muthen LK**, **Muthen BO**. Mplus User's Guide. Eighth Edition. Los Angeles, CA: Muthen & Muthen; 1998-2017.
14. **Hu L-T**, **Bentler PM**. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling. 1999;6(1):1-55.
15. **Liu SF**, **Tu DS**. On the applications of Fisher's least significant difference (LSD) procedure in three-arm clinical trials with survival endpoints. Drug Inf J. 2008;42(1):81-91.
16. **Coste J**, **Guillemin F**, **Pouchot J**, **Fermanian J**. Methodological approaches to shortening composite measurement scales. J Clin Epidemiol. 1997;50(3):247-252.
17. **Goetz C**, **Coste J**, **Lemetayer F**, **et al**. Item reduction based on rigorous methodological guidelines is necessary to maintain validity when shortening composite measurement scales. J Clin Epidemiol. 2013;66(7):710-718.
18. **Kroenke K**, **Price RK**. Symptoms in the community. Prevalence, classification, and psychiatric comorbidity. Arch Intern Med. 1993;153(21):2474-2480.
19. **Kroenke K**, **Talib TL**, **Stump TE**, **et al**. Incorporating PROMIS symptom measures into primary care practice - a randomized clinical trial. J Gen Intern Med. 2018;33(8):1245-1252.
20. **Davis LL**, **Kroenke K**, **Monahan P**, **Kean J**, **Stump TE**. The SPADE symptom cluster in primary care patients with chronic pain. Clin J Pain. 2016;32(5):388-393.
21. **Kroenke K**, **Wu J**, **Bair MJ**, **Krebs EE**, **Damush TM**, **Tu W**. Reciprocal relationship between pain and depression: a 12-month longitudinal analysis in primary care. J Pain. 2011;12(9):964-973.
22. **Kroenke K**, **Arrington ME**, **Mangelsdorff AD**. The prevalence of symptoms in medical outpatients and the adequacy of therapy. Arch Intern Med. 1990;150(8):1685-1689.
23. **Kroenke K**, **Mangelsdorff AD**. Common symptoms in ambulatory care: incidence, evaluation, therapy, and outcome. Am J Med. 1989;86(3):262-266.
24. **Kroenke K**, **Spitzer R**, **Williams JB**, **et al**. Physical symptoms in primary care: predictors of psychiatric disorders and functional impairment. Arch Fam Med. 1994;3(9):774-779.
25. **Khan AA**, **Khan A**, **Harezlak J**, **Tu W**, **Kroenke K**. Somatic symptoms in primary care: etiology and outcome. Psychosomatics. 2003;44(6):471-478.
26. US Burden of Disease Collaborators. The state of US health, 1990-2010: burden of diseases, injuries, and risk factors. JAMA. 2013;310(6):591-608.
27. **Rooth MA**. The Prevalence and Impact of Vision and Hearing Loss in the Elderly. N C Med J. 2017;78(2):118-120.
28. **Chia E-M**, **Mitchell P**, **Rochtchina E**, **Foran S**, **Golding M**, **Wang JJ**. Association Between Vision and Hearing Impairments and Their Combined Effects on Quality of Life. Arch Ophthalmol. 2006;124(10):1465-1470.
29. **Vancampfort D**, **Koyanagi A**, **Ward PB**, **et al**. Chronic physical conditions, multimorbidity and physical activity across 46 low- and middle-income countries. Int J Behav Nutr Phys Act. 2017;14(1):6.
30. **Kroenke K**, **Monahan PO**, **Kean J**. Pragmatic characteristics of patient-reported outcome measures are important for use in clinical practice. J Clin Epidemiol. 2015;68(9):1085-1092.
31. **Nunnally JC**, **Bernstein IH**. Psychometric theory. 3rd ed. New York: McGraw-Hill; 1994.
32. **Nunes BP**, **Flores TR**, **Mielke GI**, **Thume E**, **Facchini LA**. Multimorbidity and mortality in older adults: A systematic review and meta-analysis. Arch Gerontol Geriatr. 2016;67:130-138.
33. **Fabbri E**, **Zoli M**, **Gonzalez-Freire M**, **Salive ME**, **Studenski SA**, **Ferrucci L**. Aging and Multimorbidity: New Tasks, Priorities, and Frontiers for Integrated Gerontological and Clinical Research. J Am Med Dir Assoc. 2015;16(8):640-647.
34. **Garin N**, **Koyanagi A**, **Chatterji S**, **et al**. Global Multimorbidity Patterns: A Cross-Sectional, Population-Based, Multi-Country Study. J Gerontol A Biol Sci Med Sci. 2016;71(2):205-214.
35. **Pefoyo AJ**, **Bronskill SE**, **Gruneir A**, **et al**. The increasing burden and complexity of multimorbidity. BMC Public Health. 2015;15:415.
36. **Bahler C**, **Huber CA**, **Brungger B**, **Reich O**. Multimorbidity, health care utilization and costs in an elderly community-dwelling population: a claims data based observational study. BMC Health Serv Res. 2015;15:23.
37. **Cassell A**, **Edwards D**, **Harshfield A**, **et al**. The epidemiology of multimorbidity in primary care: a retrospective cohort study. Br J Gen Pract. 2018;68(669):e245-e251.
38. **Palladino R**, **Tayu Lee J**, **Ashworth M**, **Triassi M**, **Millett C**. Associations between multimorbidity, healthcare utilisation and health status: evidence from 16 European countries. Age Ageing. 2016;45(3):431-435.
39. **Besser J**, **Stropahl M**, **Urry E**, **Launer S**. Comorbidities of hearing loss and the implications of multimorbidity for audiological care. Hear Res. 2018;369:3-14.
40. **Eckerblad J**, **Theander K**, **Ekdahl A**, **et al**. Symptom burden in community-dwelling older people with multimorbidity: a cross-sectional study. BMC Geriatr. 2015;15:1.
41. **Eckerblad J**, **Theander K**, **Ekdahl AW**, **Jaarsma T**. Symptom trajectory and symptom burden in older people with multimorbidity, secondary outcome from the RCT AGe-FIT study. J Adv Nurs. 2016;72(11):2773-2783.
42. **Gould CE**, **O'Hara R**, **Goldstein MK**, **Beaudreau SA**. Multimorbidity is associated with anxiety in older adults in the Health and Retirement Study. Int J Geriatr Psychiatry. 2016;31(10):1105-1115.
43. **Helbig AK**, **Stockl D**, **Heier M**, **et al**. Relationship between sleep disturbances and multimorbidity among community-dwelling men and women aged 65-93 years: results from the KORA Age Study. Sleep Med. 2017;33:151-159.
44. **Jones SM**, **Amtmann D**, **Gell NM**. A psychometric examination of multimorbidity and mental health in older adults. Aging Ment Health. 2016;20(3):309-317.
45. **Portz JD**, **Kutner JS**, **Blatchford PJ**, **Ritchie CS**. High Symptom Burden and Low Functional Status in the Setting of Multimorbidity. J Am Geriatr Soc. 2017;65(10):2285-2289.
46. **Quinones AR**, **Markwardt S**, **Thielke S**, **Rostant O**, **Vasquez E**, **Botoseneanu A**. Prospective Disability in Different Combinations of Somatic and Mental Multimorbidity. J Gerontol A Biol Sci Med Sci. 2018;73(2):204-210.
47. **Read JR**, **Sharpe L**, **Modini M**, **Dear BF**. Multimorbidity and depression: A systematic review and meta-analysis. J Affect Disord. 2017;221:36-46.
48. **Sharpe L**, **McDonald S**, **Correia H**, **et al**. Pain severity predicts depressive symptoms over and above individual illnesses and multimorbidity in older adults. BMC Psychiatry. 2017;17(1):166.
49. **Stubbs B**, **Vancampfort D**, **Veronese N**, **et al**. Depression and physical health multimorbidity: primary data and country-wide meta-analysis of population data from 190 593 people across 43 low- and middle-income countries. Psychol Med. 2017;47(12):2107-2117.
50. **Vancampfort D**, **Koyanagi A**, **Hallgren M**, **Probst M**, **Stubbs B**. The relationship between chronic physical conditions, multimorbidity and anxiety in the general population: A global perspective across 42 countries. Gen Hosp Psychiatry. 2017;45:1-6.
51. **Villarreal AE**, **Grajales S**, **Lopez L**, **Britton GB**, **Panama Aging Research Initiative**. Cognitive Impairment, Depression, and Cooccurrence of Both among the Elderly in Panama: Differential Associations with Multimorbidity and Functional Limitations. Biomed Res Int. 2015;2015:718701.
52. **Wang YM**, **Song M**, **Wang R**, **et al**. Insomnia and Multimorbidity in the Community Elderly in China. J Clin Sleep Med. 2017;13(4):591-597.

53. **Ryan A**, **Wallace E**, **O'Hara P**, **Smith SM**. Multimorbidity and functional decline in community-dwelling adults: a systematic review. Health Qual Life Outcomes. 2015;13:168.
54. **Smith SM**, **Wallace E**, **O'Dowd T**, **Fortin M**. Interventions for improving outcomes in patients with multimorbidity in primary care and community settings. Cochrane Database Syst Rev. 2016;3:CD006560.
55. **Wallace E**, **Salisbury C**, **Guthrie B**, **Lewis C**, **Fahey T**, **Smith SM**. Managing patients with multimorbidity in primary care. BMJ. 2015;350:h176.
56. **Kirshner B**, **Guyatt G**. A methodological framework for assessing health indices. J Chronic Dis. 1985;38(1):27-36.
57. **Herdman M**, **Gudex C**, **Lloyd A**, **et al**. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011;20(10):1727-1736.
58. **Janssen MF**, **Pickard AS**, **Golicki D**, **et al**. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. Qual Life Res. 2013;22(7):1717-1727.
59. **Kroenke K.** A practical and evidence-based approach to common symptoms: a narrative review. Ann Intern Med. 2014;16(8):579-586.
60. **Aktas A**, **Walsh D**, **Rybicki L**. Symptom clusters: myth or reality? Palliat Med. 2010;24(4):373-385.