Measuring context that matters: validation of the modular Tele-QoL patient-reported outcome and experience measure

Klara Greffin[1] · Holger Muehlan[1] · Neeltje van den Berg[2] · Wolfgang Hoffmann[2] · Oliver Ritter[3] · Michael Oeff[4] · Sven Speerfork[5] · Georg Schomerus[5] · Silke Schmidt[1]

Abstract

Purpose A setting-sensitive instrument for assessing Quality of Life (QoL) in Telemedicine (TM) was unavailable. To close this gap, a content-valid "add-on" measure was developed. In parallel, a brief index was derived featuring six items that summarise the main content of the multidimensional assessment. After pre- and pilot-testing, the psychometric performance of the final measures was investigated in an independent validation study.

Methods The questionnaires were applied along with other standardised instruments of similar concepts as well as associated, yet disparate concepts for validation purposes. The sample consisted of patients with depression or heart failure, with or without TM ($n = 200$). Data analyses were aimed at calculating descriptive statistics and testing the psychometric performance on item, scale, and instrument level, including different types of validity and reliability.

Results The proposed factor structure of the multidimensional Tele-QoL measure has been confirmed. Reliability coefficients for internal consistency, split-half, and test-retest reliability of the subscales and index reached sufficient values. The Tele-QoL subscales and the index demonstrated Rasch scalability. Validity of both instruments can be assumed. Evidence for discriminant construct validity was provided. Known-groups validity was indicated by respective score differences for various classes of disease severity.

Conclusion Both measures show convincing psychometric properties. The final multidimensional Tele-QoL assessment consists of six outcome scales and two impact scales assessing (un-)intended effects of TM on QoL. In addition, the Tele-QoL index provides a short alternative for outcome assessment. The Tele-QoL measures can be used as complementary modules to existing QoL instruments capturing healthcare-related aspects of QoL from the patients' perspective.

Keywords Patient-reported outcome measure · Quality of life · Digital health · Telemedicine · Telehealth · Validation

Background

Existing health-related or disease-specific quality of life (QoL) questionnaires assess the patient-reported impact of diseases or treatments on the concept [1, 2]. Any aspects related to the context of healthcare, that might influence QoL beyond treatment, were hardly considered so far [3–5]. As part of the digitalization of healthcare, medical procedures and therapeutic treatment strategies are made available within the context of telemedicine (TM; [6]). Furthermore, additional health services are provided through innovative solutions, like telemonitoring [7–11]. This digital transformation has led to a change in healthcare contexts which is widely neglected in TM evaluations [12]. In an extensive review including 293 TM studies [4, 5], results indicated that TM-sensitive instruments were used in only about 5% of the articles included. Moreover, these instruments were only available for a limited range of concepts, as the majority was solely directed to assess satisfaction [13]. Thus, TM-specific aspects of care are not sufficiently covered by existing instruments, yet. Moreover, even though QoL is frequently considered as a core patient-reported outcome [14] in TM [3, 15, 16], there is no QoL instrument available for telehealth in particular. For this reason, we emphasize that more attention should be paid to contextual factors of healthcare, their influence on patients' experiences and health outcomes [12, 17, 18].

Klara Greffin and Holger Muehlan have shared first authorship.

Extended author information available on the last page of the article

The aim of the "Tele-QoL" project was to close this gap by developing a suitable "add-on" QoL instrument to enable a setting-sensitive evaluation of TM applications [19]. As such, this modular questionnaire shall assess QoL of patients with chronic conditions or mental illnesses in the context of TM care. This validation paper aims to document the psychometric performance of the Tele-QoL measures in terms of different forms of reliability and validity.

Methods

Instrument development

Developing the Tele-QoL instrument was based on current recommendations for patient-reported outcome measures [20, 21] and to some extent inspired by a needs-based approach of QoL assessment [22]. The items of the Tele-QoL questionnaires were directly derived from qualitative interviews and focus groups, and assess various facets of the healthcare-related domain of QoL [12]. After developing the initial version, an expert workshop for external validation ($n = 6$), an online expert survey to test the instrument's content validity ($n = 15$), and a pre-testing of the initial items with a sample of patients ($n = 32$) were conducted. Subsequently, the revised version of the questionnaire was piloted. Therefore, a sample of patients with depression or heart failure with or without TM care ($n = 200$) was recruited. As a result, we identified an appropriate measurement model comprising six factors related to patient-relevant "outcomes" (Information & Education, Perceived Safety & Well-Being, Needs Orientation & Trust, Perceived Control & Monitoring, Patient Relief & Independence, Cooperation & Communication), and two factors related to the unintended "impact" of telehealth on patients (Data Processing & Surveillance, Patient Burden & Limitation) [23]. The Tele-QoL measure aims to assess healthcare-related aspects of QoL in the context of TM applications (version A) or standard care (comparison version B). It is used as an "add-on instrument" to already existing QoL questionnaires. The target group of the Tele-QoL are patients aged 18 years and older who receive TM care (version A). It is irrelevant whether the patients are being treated for chronic physical or mental illnesses. For now, the Tele-QoL instruments are available in German in form of a short index or full version.

The questionnaire opens with a short instruction on the objective and how to carry it out; this is followed by the respective items. In addition, the temporal reference of four weeks is referred to again at the beginning of each page. Patients rate their healthcare-related experiences of the last four weeks using a 4-point Likert scale with the ratings 1 = "Do not agree" to 4 = "Highly agree". The objective of this paper is to document the evaluation of performance and

psychometric properties of the modular Tele-QoL instruments, including the multidimensional Tele-QoL measure with six outcome scales and two impact scales as well as the brief Tele-QoL index with six items. The alternative short version represents the main content of the outcome subscales as closely as possible with one item per dimension, excluding the content of the impact dimensions.

Data sample

For the validation study, patients with chronic heart failure or depression ($n = 200$), with (Tele-QoL version A) or without (Tele-QoL version B) TM care were recruited. The recruitment was implemented in several hospitals of the project's consortium partners (Brandenburg an der Havel, Greifswald, Leipzig) as well as at ambulatory healthcare facilities, all located in Northeastern Germany. In addition to the disease and treatment criteria mentioned above, a minimum age of 18 was an inclusion criterion; cognitive impairment and severe cognitive comorbidities as well as non-proficient knowledge of the German language were considered as exclusion criteria [19].

Treatment providers or study nurses in the recruitment centers informed interested patients according to pre-defined criteria in person or via telephone about purpose of the study, voluntariness, dropout options, and compensations. In addition, the patients received the information in written form along with phone and e-mail contact details of the recruiting centers and the scientific research assistant. After the patients had given informed consent for the study, the questionnaires were handed out with the request to fill them in. Personal codes were generated for a pseudonymous assignment of the follow-up survey, that was scheduled four weeks later. Personal assistance during the completion of the survey was available upon request. Completed questionnaires were mailed or dropped off in a prepaid envelope. After the questionnaires have been received, data were entered into an Excel spreadsheet and stored on a secured file server. Finally, the original questionnaires were stored in lockable cabinets.

Applied measures

Whereas all measures were applied to the first wave of the validation study, only some of them were also used within the second wave after four weeks to detect test-retest reliability, stability over time, and sensitivity to change. All instruments included in the validation study are described in the study protocol [19]. Therefore, we will provide a short overview, only.

Sociodemographic characteristics were assessed based on recommended demographic standards in Germany [24]. The "Goldman Specific Activity Scale" (original version: [25]) was

used to assess severity of heart failure. Participants were asked to rate whether they are able to perform specific daily activities, and were classified according to four Specific Activity Scale Functional Classes (Class I = least burdened; Class IV = most burdened) based on their answers. It was complemented by the "New York Heart Association Classification" (NYHA; original version: [26]; German version [27]:). Based on their answers, the participants were assigned to one of four classes (NYHA 1 = least burdened; NYHA 4 = most burdened). Depressive symptoms were assessed using the "Patient Health Questionnaire 9 (PHQ-9)" [28]. The "SeCu-20" Questionnaire (German original version: [29]) was used to assess perceived security in experiences with TM care. Patient satisfaction was assessed by the "Questionnaire to evaluate patient satisfaction (ZUF-8)" (original version: [30]; German version: [31]). Additionally, the general item of the "Youth Health Care Measure (YHC-SUN)" [32] was used to assess the general satisfaction with treatment. Patient activation was assessed using the "Patient Activation Measure (PAM13-D)" (original version: [33]; German version: [34]). To assess body-related self-consciousness, the subscale "private" of the "Body-related Self-Consciousness (KSA)" questionnaire (German original version: [35]) was used. From the "Body-related Locus of Control (KLC)" questionnaire (German original version: [36, 37]) for assessing the body-related locus of control, the subscale "health" was used. The "European Health Literacy Survey (HLS-EU-Q6)" (original version in multiple languages: [38]) was used to assess health literacy. In addition, we used a newly adapted version of HLS-EU-Q6 for digital healthcare, referred to as D-HLS-EU-Q6. The"WHO-Five Well-Being Index (WHO-5)" [39, 40] was used to assess QoL of participants with mental health issues. The "Minnesota Living with Heart Failure questionnaire (MLHFQ)" was used to assess QoL of patients with heart failure [41–43]. The "Veterans RAND 12 Item Health Survey (VR-12)" (original version: [44]; German version: [45]) was used to assess the subjective health status of the participants. Health-related QoL was assessed with the "European Quality of Life 5 Dimensions (EQ-5D)" (original version: [46, 47]). The short form of the "World Health Organization Quality of Life (WHOQOL-BREF)" (original version in multiple languages: [48]) was used to assess the general QoL in four different life domains (physical, psychological, social, environmental).

Data analyses

Factorial validity was investigated by conducting a confirmatory factor analysis with maximum likelihood estimations to test the multidimensional measurement model. Amongst the fit indices [49] we inspected the comparative fit index (CFI), Tucker-Lewis index (TLI), and the root mean square error of approximation (RMSEA). *Discriminant validity* was investigated by calculating Pearson correlation coefficients for associations between Tele-QoL scores and various indicators of general, health-related, and disease-specific QoL as well as measures related to the assessment of satisfaction with care, patient activation, and health literacy. All concepts were assumed to be low or moderately associated with the Tele-QoL scores.

Concerning *convergent validity*, we assumed high associations with the subscales of a setting-sensitive measure for patient experiences in TM. Finally, we tested for correlations with further associated constructs, including self-monitoring and locus of control.

To examine *known-groups validity* with respect to different clinical variables known for differences in QoL, standardized effect sizes for differences of two independent means were estimated using Cohen's d [50]. We expected that patients with stronger disease severity show lower Tele-QoL outcome and higher impact scores.

Rasch analysis was used to detect possible misfit on item level. The partial credit model was applied to the data, using Q index statistics and threshold ordering estimation for detecting item misfit [51].

For *reliability* testing, homogeneity of the subscales was investigated by computing Cronbach's alpha coefficient α. Split-half reliability was determined by the correlation between both test-halves. Intraclass correlation coefficient (ICC) was used to estimate test-retest reliability of the Tele-QoL scores.

Statistical software

Descriptive statistics and item-scaling analysis were performed using the IBM SPSS Version 28.0 [52]. Confirmatory factor analysis was processed using IBM AMOS Version 28 [53]. For Rasch analysis, the WINMIRA software package was used [54].

Results

Sample characteristics

In total, $n = 200$ patients aged 19 to 88 years participated in the Tele-QoL validation study (Table 1). Of these, 51.5% ($n = 103$) reported being male, 48.0% ($n = 96$) female, and 0.5% ($n = 1$) non-binary/gender nonconforming, respectively. Patients included were being treated for chronic heart failure (52.0%, $n = 104$) or depression (48.0%, $n = 96$). Sociodemographic characteristics for each patient group are provided in Table 1. Eighty patients participated in the second wave.

There have been patients who did not want to participate in this study, mostly due to a lack of interest or to feeling overwhelmed by the amount of study participation requests they got from the hospitals. A third group was already stressed due

Table 1 Sociodemographic characteristics of the Tele-QoL validation study sample ($n = 200$)

Characteristics*	Patients with heart failure		Patients with depression	
	Standard care	Telemedical care	Standard care	Telemedical care
Age group				
< 35 years	–	–	22 (48.0%)	7 (13.7%)
36—50 years	1 (1.8%)	1 (2.0%)	12 (26.7%)	11 (21.6%)
51—65 years	20 (36.4%)	9 (18.4%)	9 (20.0%)	29 (56.9%)
66—80 years	23 (41.8%)	26 (53.1%)	2 (4.4%)	4 (7.8%)
> 80 years	8 (14.5%)	9 (18.4%)	–	–
Gender				
Female	18 (32.7%)	16 (32.7%)	27 (60.0%)	35 (68.6%)
Male	37 (67.3%)	33 (67.3%)	17 (37.8%)	16 (31.4%)
Non-binary / gender nonconforming	–	–	1 (2.2%)	–
Education (Highest Degree)				
Primary school (8th/9th class)	3 (5.5%)	10 (20.4%)	3 (6.7%)	2 (3.9%)
Secondary School (10th class)	27 (49.1%)	25 (51.0%)	14 (31.1%)	29 (56.9%)
High School (12th/13th class)	15 (27.3%)	9 (18.4%)	26 (57.8%)	15 (29.4%)
Other Degree	7 (12.7%)	2 (4.1%)	2 (4.4%)	2 (3.9%)
No Formal Degree	2 (3.6%)	1 (2.0%)	–	–

*Data referring to frequencies and percentages. Absolute frequencies vary as a function of the amount of missing data for each variable

**Sum of percent value may vary resulting from rounding of single percent rates

to covid-19 and related circumstances, which was also a factor for non-participation. Unfortunately, the number of these individuals was not systematically recorded in the clinics.

Data

Factorial validity was explored by applying confirmatory factor analysis (CFA). We used Maximum-likelihood parameter estimation for testing the model. Despite impaired normal distribution of items, this method can be applied as it is assumed to be robust even if the data violates the assumption of normal distribution. The model did fit the data well ($\chi^2(df = 436) = 696.53$, $p < 0.001$, *CFI = 0.94*, *TLI* = 0.93, *RMSEA* = 0.056 [0.048; 0.064]).

The six "outcome" subscale scores of the multidimensional Tele-QoL instrument correlate with each other with $r = 0.39\text{-}0.81$, the two "impact" subscales with $r = 0.44$ (see Table 2). The Tele-QoL index score correlates moderately to highly with all outcome scales of the multidimensional Tele-QoL ($r = 0.59 - 0.83$), but slightly negatively with both impact scales ($r = -0.12$ and $r = -0.16$).

Rasch analysis (Partial Credit Model) with emphasis on the operational characteristics of the items showed that none of the items in any of the Tele-QoL subscales or the Tele-QoL index displays infit, indicating no substantial deviation from the model. The range of item locations for most of the scales is moderate (< 2 logits), but the effective range carried by threshold distributions along the latent traits varies between > 4 and < 11 logits. Ordering of thresholds is in

accordance with the model assumptions for all items in any of the subscales and the index as well (Table 3).

For reliability testing, the *internal consistency* was calculated using Cronbach's alpha (α) coefficient for all subscales and the index score. For the Tele-QoL index, a value of $\alpha = 0.90$, and for the Tele-QoL subscales values between $\alpha = 0.84$ and 0.95 were obtained. Thus, the internal consistencies for all scales of the Tele-QoL instruments can be judged as very good. All subscales of the Tele-QoL measure as well as the Tele-QoL index also yielded very good values for the *split-half-reliability*, which varied between 0.81 and 0.91. *Test-retest reliability* was determined over a period of approximately four weeks, controlling for the course of the disease. The corresponding intraclass correlation coefficients vary between 0.64 and 0.74 and are thus sufficient to good. T-tests indicate no significant differences between test and retest scores, except for the subscale "Patient Burden & Limitation" ($p = 0.01$). All reliability coefficients for the subscales of the Tele-QoL and the Tele-QoL index are also depicted in Table 3.

Evidence for *known-groups validity* of the Tele-QoL measure is displayed by expected group differences ($d = 0.01 < 0.44$) in the Tele-QoL scores for patients with different disease severity (Table 4). With regard to *discriminant construct validity related to QoL* results show low to moderate correlations with different indices of general QoL (WHOQOL-BREF), health-related QoL (EQ-5D, VR-12), disease-specific QoL, and well-being (MLHFQ, WHO-5). These findings indicate a sufficient *divergent validity* of the Tele-QoL instruments, since they capture different aspects of QoL than other instruments

Table 2 Intercorrelations of the Tele-QoL subscale scores (n = 200)

| Tele-QoL instruments | Number of items | Tel-QoL outcome scales | | | | | | Tel-QoL impact scales | | Tele-QoL index |
Multidimensional Tele-QoL measure		Needs Orientation & Trust	Patient Relief & Autonomy	Information & Education	Cooperation & Communication	Perceived Control & Monitoring	Perceived Safety & Well-Being	Data Processing & Surveillance	Patient Burden & Limitation	Total score
Needs Orientation & Trust	4	–	0.44	0.50	0.40	0.52	0.51	−0.24	−0.25	0.59
Patient Relief & Autonomy	4	0.44	–	0.69	0.63	0.47	0.39	−0.07	−0.07	0.63
Information & Education	4	0.50	0.69	–	0.80	0.63	0.49	−0.12	−0.13	0.76
Cooperation & Communication	4	0.40	0.63	0.80	–	0.60	0.53	−0.06	−0.14	0.73
Perceived Control & Monitoring	4	0.52	0.47	0.63	0.60	–	0.81	−0.12	−0.28	0.83
Perceived Safety & Well-Being	4	0.51	0.39	0.49	0.53	0.81	–	−0.20	−0.26	0.74
Data Processing & Surveillance	4	−0.24	−0.07	−0.12	−0.06	−0.12	−0.20	–	0.44	−0.12
Patient Burden & Limitation	4	−0.25	−0.07	−0.13	−0.14	−0.28	−0.26	0.44	–	−0.16
Tele-QoL index	6	0.59	0.63	0.76	0.73	0.83	0.74	−0.12	−0.16	–

Interpretation of correlation coefficients: r < 0.50: low; r = 0.50 - 0.70: moderate; r > 0.70: high. In bold print: r > 0.30

assessing already established concepts of QoL (Table 5). Most coefficients for correlations between the six Tele-QoL outcome subscales with scores from other QoL measures are notably higher for those domains related to mental issues (WHOQOL-BREF: Mental/Psychological Domain, VR-12 Mental Health Status) than for domains related to physical issues (WHOQOL-BREF: Physical Domain, VR-12 Physical Health Status). Also, domains related to social or environmental issues show higher correlations than domains related to physical issues, but not as high as the "mental" domains. Correlation coefficients with physical domains of QoL are generally weak or low ($r = −0.17$ to 0.27). Both impact scales of the multidimensional Tele-QoL measure show low negative correlations with almost all QoL scores ($r = −0.25$ to 0.02).

Discriminant construct validity related to patients' experiences with healthcare provision was also investigated using other measures of related concepts assessing satisfaction with healthcare (YHC-SUN), patient satisfaction (ZUF-8) as well as patient activation (PAM13-D). For almost all correlations between Tele-QoL outcome subscale scores and index score, coefficients indicate low to moderate associations ($r = 0.22–0.61$). In addition, discriminant construct validity was also investigated with selected patient experiences covered by the Tele-QoL scales.

Considering "Information & Education", health-literacy (HLS-EU-Q6) as well as digital health literacy (D-HLS-EU-Q6) were assessed. Correlation coefficients indicate very low associations ($r < 0.10$). With respect to "Perceived Control & Monitoring", we applied instruments assessing related concepts such as private body-related self-monitoring (KSA) as well as internal and external health-related locus of control (KLC). Again, correlations coefficients also indicate very low associations ($r < 0.10$).

All six outcome subscales scores of the Tele-QoL instrument and the index score correlate with the three subscales of the SeCu-instrument assessing patient experiences in TM ($r = 0.36–0.90$). This supports the assumption of *convergent validity*. Missing substantial correlations ($r = −0.06 < 0.07$) with the SeCu subscale assessing negative experiences in TM ("Technology Anxiety") indicate *divergent validity*. Correspondingly, the subscale "Data Processing & Surveillance" of the Tele-QoL instrument shows a low correlation coefficients with "Technology Anxiety" ($r = 0.31$), but both impact subscales correlate slightly negative with the three "positive" SeCu subscales ($r = −0.30 < −0.08$).

Discussion

Main results

With the Tele-QoL measures, we provide a modular instrument that assesses the impact of the TM healthcare context

Table 3 Rasch analysis and reliabilities of the multidimensional Tele-QoL sub-scales and the Tele-QoL index (n = 200)

Tele-QoL instruments	Number of items	Range of item locations	Range of threshold parameters	Non-ordered thresholds	Item fit (Q index)	Internal consistency	Spilt-half reliability	Test–retest reliability ICC (CI)	p-value (t-test)
Multidimensional Tele-QoL measure		($n_{min-max}$ = 178–185)	($n_{min-max}$ = 178–185)	($n_{min-max}$ = 178–185)	($n_{min-max}$ = 178–185)	($n_{min-max}$ = 178–185)	($n_{min-max}$ = 178–185)	($n_{min-max}$ = 75–78)	($n_{min-max}$ = 75–78)
Needs Orientation & Trust	4	−0.84 to 0.75	−3.90 to 3.10	–	0.025 to 0.048	0.90	0.89	0.64 (0.48-0.76)	0.18
Patient Relief & Autonomy	4	−0.13 to 0.23	−2.39 to 2.92	–	0.034 to 0.069	0.87	0.83	0.74 (0.62-0.83)	0.23
Information & Education	4	−0.43 to 0.03	−2.34 to 2.47	–	0.030 to 0.092	0.83	0.83	0.73 (0.61-0.82)	0.48
Cooperation & Communication	4	−1.16 to 0.69	−4.35 to 3.49	–	0.028 to 0.082	0.90	0.84	0.66 (0.51-0.77)	0.36
Perceived Control & Monitoring	4	−0.82 to 0.80	−2.28 to 2.24	–	0.038 to 0.071	0.84	0.91	0.72 (0.59-0.81)	0.46
Perceived Safety & Well-Being	4	−1.17 to 0.96	−6.27 to 4.46	–		0.87	0.84	0.70 (0.57-0.80)	0.07
Data Processing & Surveillance	4	−0.95 to 0.88	−3.20 to 3.00	–	0.019 to 0.050	0.93	0.81	72 (0.59-0.81)	0.13
Patient Burden & Limitation	4	−0.84 to 0.84	−2.60 to 5.63	–	0.021 to 0.038	0.95	0.91	0.72 (0.59-0.81)	0.01*
Tele-QoL index	6	−0.26 to 0.74	−2.71 to 3.05	–	0.046 to 0.075	0.90	0.84	0.69 (0.55-0.79)	0.26

Table 4 Known-groups validity of the Tele-QoL subscales scores and Tele-QoL index score (n = 200)

Tele-QoL instruments	Patients with heart failure (NYHA = 1 vs. NYHA > 1)		Patients with depression (PHQ < 15 vs. PHQ > 14)	
Multidimensional Tele-QoL measure	Mean difference*	Effect size Cohen's d	Mean difference*	Effect size Cohen's d
Needs Orientation & Trust	0.93	**0.28**	0.42	0.16
Patient Relief & Autonomy	0.33	0.12	1.09	**0.41**
Information & Education	0.52	0.17	1.09	**0.38**
Cooperation & Communication	1.21	**0.33**	1.29	**0.41**
Perceived Control & Monitoring	1.51	**0.44**	0.02	0.01
Perceived Safety & Well-Being	0.63	0.19	0.78	**0.33**
Data Processing & Surveillance	-0.79	**−0.27**	−0.46	−0.18
Patient Burden & Limitation	-0.93	**−0.37**	−0.80	**−0.34**
Tele-QoL index	0.84	0.17	0.94	**0.40**

*Tele-QoL index and subscales scores are linear transformed raw means (range: 0–100). In bold print: d > 0.20

on QoL of patients, beyond the effects of the disease and the treatment [12]. This study tested and determined the psychometric properties of the Tele-QoL measures when applied to a sample of German patients with depression or heart failure.

Summarizing the results of this study, the Tele-QoL measures show a promising psychometric performance. Our results confirm the factorial structure of the multidimensional measure. The multidimensional measure suggests

that the Tele-QoL captures distinct domains of healthcare-related QoL in a reliable and valid manner. The reliabilities of all subscales and of the index are satisfying, with alpha coefficients and split-half reliabilities being very good and test retest-reliabilities with sufficient to good values. This indicates that the items within each subscale are highly consistent in measuring the intended construct. Considering that each subscale consists of only four items, these results are

Table 5 Intercorrelations of the Tele-QoL scores with subscale scores of other measures for convergent and discriminant validation (n = 200)

Tele-QoL module (subscales) Items		Needs Orientation & Trust	Patient Relief & Autonomy	Information & Education	Cooperation & Communication	Perceived Control & Monitoring	Perceived Safety & Well-Being	Data Processing & Surveillance	Patient Burden & Limitation	Tele-QoL index
		4	4	4	4	4	4	4	4	6
WHOQOL-BREF	Physical domain	0.10	0.18	0.21	0.27	0.10	0.05	−0.14	−0.17	0.10
	Mental domain	0.15	**0.40**	**0.36**	0.28	0.17	0.09	−0.18	−0.07	0.27
	Social domain	0.17	**0.34**	0.27	0.21	0.23	0.19	−0.16	−0.04	**0.30**
	Environmental domain	0.17	0.29	0.27	0.25	0.26	0.13	−0.25	−0.22	0.29
EQ-5D	Health-related QoL	0.04	0.06	0.09	0.19	0.01	0.04	−0.21	−0.11	0.01
VR-12	Physical health status	0.04	−0.01	0.01	0.12	−0.01	0.05	−0.13	−0.11	−0.02
	Mental health status	0.18	**0.43**	**0.37**	**0.31**	0.20	0.10	−0.15	−0.12	0.26
MLHFQ *	Disease-specific QoL (Heart failure)	0.20	0.26	**0.33**	**0.34**	**0.34**	0.27	−0.14	−0.15	**0.30**
WHO-5**	Disease-specific QoL (depression)	0.14	**0.37**	0.29	**0.36**	0.17	0.15	0.02	−0.04	0.24
YHC-SUN-1	Satisfaction with healthcare	**0.39**	**0.41**	**0.46**	**0.47**	**0.47**	**0.48**	−0.10	−0.12	**0.52**
ZUF-8	Patient satisfaction	**0.46**	**0.48**	**0.56**	**0.58**	**0.59**	**0.56**	−0.19	−0.17	**0.61**
PAM-13-D	Patient activation	0.22	0.28	**0.33**	**0.34**	0.27	0.22	−0.14	−0.01	**0.33**
HLS-EU-Q6	Health literacy	0.16	0.07	0.07	0.21	0.17	0.28	−0.19	−0.19	0.15
D-HLS-EU-Q6	Digital health literacy	0.13	0.05	0.04	0.18	0.10	0.21	−0.18	−0.12	0.08
KSA	Private body-related self-monitoring	0.10	−0.03	−0.01	−0.01	−0.05	0.03	0.11	0.10	0.05
KLC	Health-related locus of control (internal)	0.17	0.01	0.05	0.02	0.06	0.10	0.01	−0.05	0.04
KLC	Health-related locus of control (external)	−0.08	0.01	0.05	−0.03	−0.01	−0.09	0.03	−0.02	−0.05
TBS	Technology acceptance	0.12	0.12	0.13	0.16	−0.03	−0.00	−0.12	−0.05	0.07
	Technology competence	0.12	−0.09	−0.11	0.01	−0.02	0.06	−0.19	−0.18	−0.04
	Technology control	0.02	−0.05	0.04	0.15	0.02	0.02	−0.08	−0.09	−0.00
	Technology anxiety	−0.06	−0.02	−0.04	0.05	0.04	0.07	0.31	0.20	0.07
SeCu***	Perceived security	**0.58**	**0.57**	**0.81**	**0.74**	**0.80**	**0.80**	−0.22	**−0.30**	**0.84**
	Perceived autonomy	**0.38**	**0.63**	**0.78**	**0.90**	**0.62**	**0.66**	−0.08	−0.08	**0.82**
	Physician–patient-relationship	**0.36**	**0.59**	**0.70**	**0.66**	**0.72**	**0.71**	−0.08	−0.17	**0.81**

*For patients with chronic heart failure only; **for patients with depression only, ***for patients with telemedical care only. Interpretation of correlation coefficients: r < 0.50: low; r = 0.50–0.70: moderate; r > 0.70: high. In bold print: r ≥ 0.30. *WHOQOL-BREF* World Health Organization Quality of Life measure (short version); *EQ-5D* EuoQol Quality of Life measure; *VR-12* Veterans RAND 12 Item Health Survey (short version); *MLHFQ* Minnesota Living with Heart Failure Questionnaire; *WHO-5* World Health Organization Well-Being Scale; *YHC-SUN-1* Generic single-item measure of satisfaction with healthcare from the Satisfaction, Utilization & Needs Questionnaire (Youth version); *ZUF-8* Patient Satisfaction Questionnaire (8 item version); *PAM-13-D* Patient Activation Measure; *HLS-EU-Q6* Health Literacy Scale HLS-EU (6 item version); *D-HLS-EU-Q6* (Adopted) Digital version of the Health Literacy Scale HLS-EU (6 item version); *KSA* Body-related Self-Awareness Scale; *KLC* Body-related Locus of Control Scale; *TBS* Technology Commitment Scale; *SeCu* Perceived Security in Telemedicine Scale

quite convincing. Also, operational characteristics of the items for all scales were in line with the model assumptions implied by the Rasch model. Furthermore, this validation study suggests that the Tele-QoL measures possess both discriminant and convergent validity. There is reasonable evidence that the concept of healthcare-related QoL and the domains representing this construct in the measurement model are not identical with related constructs and are sufficiently distinguished from each other in terms of the discriminant validity. This implies that these measures effectively distinguish between various aspects of QoL. In addition, data provides evidence bolstering the incremental validity, indicating that the Tele-QoL measures offer unique insights beyond what is already captured by other QoL questionnaires. Moreover, a high level of content validity can be assumed, as the Tele-QoL questionnaires underwent a rigorous development process, drawing from extensive qualitative material, collected in previous studies of the Tele-QoL project [12]. Noteworthy, the observed correlations among the Tele-QoL outcome subscales imply that there are substantial interrelationships among various aspects of healthcare-related QoL in TM contexts, which is in line with our previous qualitative study [12]. Collectively, these findings affirm the robustness and utility of the Tele-QoL measures in assessing healthcare-related QoL.

How can the Tele-QoL measures benefit the evaluation of TM applications?

According to a modern understanding, the majority of patients are considered active protagonists who no longer want to be treated passively, but also want to make their own contribution to their health [55, 56]. With a long-lasting illness, however, the needs and challenges in everyday life that a patient is confronted with also increase [57]. For this reason, it is the purpose of TM care for long-term illnesses to support patients in the management of their illness and the needs associated with it [58]. In order to assess whether and to what extent TM applications are able to provide this support, appropriate assessments are needed that reflect the patient's perspective [59]. Therefore, the development and implementation of setting-sensitive questionnaires like the Tele-QoL measures are crucial as they allow for a more valid assessment in TM studies. In this way, the healthcare context is included in the evaluation of care components, in addition to the effects of the disease and respective treatment. As a result, the demand for a valid and quantitative summative evaluation of the medical benefit can now be better met [59].

In general, patients using TM will have the opportunity to better represent the impact of TM on their QoL via the Tele-QoL questionnaire. The extended conceptualization of QoL in TM settings may also lead to potential improvements

of TM applications and individualized TM care for patients with chronic diseases and mental illnesses.

Strengths and limitations

The Tele-QoL is developed based on an extensive mixed-methods approach, which is a strength in terms of content validity [20, 21]. Moreover, patients were included in all stages of the development and validation process. Another advantage is the sample composition for validation, consisting of respondents with complementary diseases and forms of treatment. Thus, half of the sample consisted of patients with TM or standard treatment, half of whom were chronically physically or mentally ill. Amongst patients with TM care, half of them were treated with an active TM approach (regular phone calls), the other half were treated with a passive TM application (remote vital monitoring). The aim was to represent all potential user groups and to test whether the questionnaire can be used independently of the disease and the treatment.

However, our validation study also has limitations. First of all, in planning the project, a compromise had to be made between an adequate sample size and the feasibility of data collection. A sample of $n = 200$ is considered fair [60] and is therefore sufficient, but can be expanded. Future evaluation of the psychometric properties should be based on larger samples, including more disease groups and other TM settings. Moreover, other important properties of the measures need to be investigated, such as readability or responsiveness.

To assess test-retest reliability, patients were asked to complete a second questionnaire four weeks after the initial survey. The date for the second questionnaire was written on the instrument. In addition, after completing the second questionnaire, patients were asked to write the current date under the questionnaire's items. Unfortunately, not all patients did so. Therefore, we cannot be sure in every case that the questionnaires were filled out exactly four weeks later.

The severity of the respective disease, which was used for calculating the known-groups validity, was based on patients' self-reports, assessed via patient-reported outcome measures. The data may be biased, for example, by how someone feels on a particular day. In addition, the validation was conducted as a questionnaire study in which patients were asked to fill out different questionnaires one after the other. We arranged the order of the questionnaires in such a way that the questions on general health run towards specific health questions in order to cause as little priming as possible. Nevertheless, answering one questionnaire may have an impact on answering subsequent questionnaires.

It remains unclear what effect the SARS-CoV-2 pandemic outbreak had on our sample. The recruiting institutions

had the impression that more severely burdened patients were less willing to participate in the study than before the pandemic, but this circumstance was not systematically recorded. Nevertheless, it should be reported that in this context there may have been a selection and nonresponse bias in our sample regarding the severity levels included. Besides, TM was suddenly used as a substitute, not as a complement.

In summary, this instrument development demonstrates that the psychometric properties of the Tele-QoL measures are convincing. However, it only remains the first step towards a fully validated questionnaire [61].

Conclusion & outlook

The modular Tele-QoL instrument represents a methodologically sound measure to assess QoL in TM settings. It can be used as complementary module to existing QoL instruments to assess healthcare-related aspects of QoL from the patients' perspective in telehealth contexts. It is an important and necessary contribution to developing, implementing, and evaluating digital health applications. In the future, the Tele-QoL approach will be further adapted so that it can also be used for children and adolescents (new development of a Tele-QoL Kids) as well as in other countries (cultural adaptation and translation) facing similar healthcare challenges.

References

1. Haraldstad, K., Wahl, A., Andenæs, R., Andersen, J. R., Andersen, M. H., Beisland, E., Borge, C. R., Engebretsen, E., Eisemann, M., Halvorsrud, L., Hanssen, T. A., Haugstvedt, A., Haugland, T., Johansen, V. A., Larsen, M. H., Løvereide, L., Løyland, B., Kvarme, L. G., & P. Moons… S. Helseth,. (2019). A systematic review of quality of life research in medicine and health sciences. *Quality of Life Research, 28*(10), 2641–2650. https://doi.org/10.1007/s11136-019-02214-9

2. Patrick, D. L., & Deyo, R. A. (1989). Generic and disease-specific measures in assessing health status and quality of life. *Medical Care, 27*(3 Suppl), S217–S232. https://doi.org/10.1097/00005650-198903001-00018

3. Knapp, A., Harst, L., Hager, S., Schmitt, J., & Scheibe, M. (2021). Use of patient-reported outcome measures and patient-reported experience measures within evaluation studies of telemedicine applications: Systematic review. *Journal of Medical Internet Research, 23*(11), e30042. https://doi.org/10.2196/30042

4. Greffin, K., Muehlan, H., Rosenkranz, E., van den Berg, N., Hoffmann, W., Ritter, O., Oeff, M., Schomerus, G. & Schmidt, S, "Telemedicine and patient-reported outcomes in chronic conditions: concordance and discrepancy of purpose, concepts, and methods of measurement—A systematic literature review (submitted)".

5. Greffin, K., Muehlan, H., Rosenkranz, E., van den Berg, N., Hoffmann, W., Ritter, O., Oeff, M., Schomerus, G., & Schmidt, S., "Telemedicine and patient-reported outcomes in mental illnesses: concordance and discrepancy of purpose, concepts, and methods of measurement—a systematic literature review (submitted)".

6. Muehlan, H., & Schmidt, S. (2013). Versorgungsbezogene E-Health-Anwendungen und Lebensqualität: Empirische Befunde und konzeptuelle Perspektiven. *Psychiatrische Praxis, 40*(6), 307–312. https://doi.org/10.1055/s-0033-1349489

7. Schmidt, S. (2007). Telemedizin und Lebensqualität. *Deutsche Medizinische Wochenschrift, 132*(09), 442–447. https://doi.org/10.1055/S-2007-970355

8. Oeff, M., Müller, A., Neuzner, J., Sack, S., Schwab, J. O., Pfeiffer, D., & Zugck, C. (2008). EKG-telemonitoring. *Herzschrittmachertherapie und Elektrophysiologie, 19*(3), 137–145. https://doi.org/10.1007/s00399-008-0012-7

9. van den Berg, N., Schmidt, S., Stentzel, U., Mühlan, H., & Hoffmann, W. (2015). Telemedizinische Versorgungskonzepte in der regionalen Versorgung ländlicher Gebiete. *Bundesgesundheitsblatt-Gesundheitsforschung-Gesundheitsschutz, 58*(4–5), 367–373. https://doi.org/10.1007/s00103-015-2134-5

10. Johnson, S. E., Fedor, M. A., & Hoban, M. B. (2008). Telemonitoring: A positive impact on patient outcomes. *Caring, 27*(7), 20–25.

11. Ritter, O., & Bauer, W. R. (2006). Use of 'IEGM Online' in ICD patients—early detection of inappropriate classified ventricular tachycardia via home monitoring. *Clinical Research in Cardiology, 95*(7), 368–372. https://doi.org/10.1007/s00392-006-0390-y

12. Greffin, K., Schmidt, S., van den Berg, N., Hoffmann, W., Ritter, O., Oeff, M., Schomerus, G., & Muehlan, H. (2021). Same same—but different: Using qualitative studies to inform concept elicitation for quality of life assessment in telemedical care—A request for an extended working model. *Health and Quality of Life Outcomes.* https://doi.org/10.1186/s12955-021-01807-8

13. Hajesmaeel-Gohari, S., & Bahaadinbeigy, K. (2021). The most used questionnaires for evaluating telemedicine services. *BMC Medical Informatics and Decision Making, 21*(1), 36. https://doi.org/10.1186/s12911-021-01407-y

14. Bullinger, M., & Quitmann, J. (2014). Quality of life as patient-reported outcomes: Principles of assessment. *Dialogues in Clinical Neuroscience, 16*(2), 137–145. https://doi.org/10.31887/DCNS.2014.16.2/mbullinger

15. Riva, S., Mazzocco, K., & Pravettoni, G. (2015). Better outcomes for cancer patients using telemedicine: Health-related quality of life (HrQoL) and empowerment. In G. Gatti, G. Pravettoni, & F. Capello (Eds.), *Tele-oncology* (pp. 23–37). Springer.

16. Schmidt, S. (2007). Telemedicine and quality of life. *Deutsche Medizinische Wochenschrift, 132*(9), 442–447. https://doi.org/10.1055/s-2007-970355

17. Deng, G., Weber, W., Sood, A., & Kemper, K. J. (2010). Research on integrative healthcare: Context and priorities. *Explore (NY), 6*(3), 143–158. https://doi.org/10.1016/j.explore.2010.03.007

18. Ng, J. H. Y., & Luk, B. H. K. (2019). Patient satisfaction: Concept analysis in the healthcare context. *Patient Education and Counseling, 102*(4), 790–796. https://doi.org/10.1016/j.pec.2018.11.013

19. Greffin, K., Muehlan, H., van den Berg, N., Hoffmann, W., Ritter, O., Oeff, M., Schomerus, G., & Schmidt, S. (2021). Setting-sensitive conceptualization and assessment of quality of life in telemedical care—study protocol of the Tele-QoL project. *International Journal of Environmental Research and Public Health.* https://doi.org/10.3390/ijerph181910454

20. Cheng, K. K. F., & Clark, A. M. (2017). Qualitative methods and patient-reported outcomes: Measures development and adaptation. *International Journal of Qualitative Methods, 16*(1), 1–3. https://doi.org/10.1177/1609406917702983

21. Rothrock, N. E., Kaiser, K. A., & Cella, D. (2011). Developing a valid patient-reported outcome measure. *Clinical Pharmacology and Therapeutics, 90*(5), 737–742. https://doi.org/10.1038/clpt.2011.195

22. McKenna, S. P., & Doward, L. C. (2004). The needs-based approach to quality of life assessment. *Value Health, 7*(SUPPL. 1), S1–S3. https://doi.org/10.1111/j.1524-4733.2004.7s101.x

23. Muehlan, H., Greffin, K., van den Berg, N., Hoffmann, W., Ritter, O., Oeff, M, & G. Schomerus, and S. Schmidt, "Towards adjunct setting-related quality of life assessment in telemedicine - cognitive debriefing, expert rating and pilot testing of the Tele-QoL instrument (submitted)".

24. Beckmann, K., Glemser, A., Heckel, C., & Al, E. (2016). *Demographische Standards. Eine gemeinsame Empfehlung des ADM, Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V., der Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e.V. (ASI) und des Statistischen Bundesamtes*, 6. Aufl. Wiesbaden.

25. Goldman, L., Hashimoto, B., Cook, E. F., & Loscalzo, A. (1981). Comparative reproducibility and validity of systems for assessing cardiovascular functional class: Advantages of a new specific activity scale. *Circulation, 64*(6), 1227–1234. https://doi.org/10.1161/01.CIR.64.6.1227

26. Dolgin, M., New York Heart Association, & Criteria Committee. (1994). *Nomenclature and Criteria for Diagnosis of Diseases of the Heart and Great Vessels*, 9th ed. Little Brown & Co.

27. Hoppe, U. C., Böhm, M., Dietz, R., Hanrath, P., Kroemer, H. K., Osterspey, A., Schmaltz, A. A., & Erdmann, E. (2005). Leitlinien zur Therapie der chronischen herzinsuffizienz. *Zeitschrift für Kardiologie, 94*(8), 488–509. https://doi.org/10.1007/s00392-005-0268-4

28. Kroenke, K., Spitzer, R. L., Williams, J. B. W., & Löwe, B. (2010). The patient health questionnaire somatic, anxiety, and depressive symptom scales: A systematic review. *General Hospital Psychiatry, 32*(4), 345–359. https://doi.org/10.1016/j.genhosppsych.2010.03.006

29. Karg, O., Weber, M., Bubulj, C., Esche, B., Weber, N., Geiseler, J., Bachl, C., & Schellhorn, H. (2012). Acceptance of a telemonitoring device in patients with chronic obstructive pulmonary disease. *Deutsche Medizinische Wochenschrift, 137*(12), 574–579.

30. Attkisson, C. C., & Zwick, R. (1982). The client satisfaction questionnaire. *Evaluation and Program Planning, 5*(3), 233–237. https://doi.org/10.1016/0149-7189(82)90074-X

31. Schmidt, J., Lamprecht, F., & Wittmann, W. W. (1989). Zufriedenheit mit der stationären Versorgung. Entwicklung eines Fragebogens und erste Validitätsuntersuchungen. *Psychotherapie Psychosomatik Medizinische Psychologie, 39*, 248–255.

32. Schmidt, S., Thyen, U., Herrmann-Garitz, C., Bomba, F., & Muehlan, H. (2016). The youth health care measure-satisfaction, utilization, and needs (YHC-SUN)-development of a self-report version of the child health care (CHC-SUN) proxy-measure. *BMC Health Services Research, 16*(1), 1–12. https://doi.org/10.1186/s12913-016-1419-1

33. Hibbard, J. H., Mahoney, E. R., Stockard, J., & Tusler, M. (2005). Development and testing of a short form of the patient activation measure. *Health Services Research, 40*(6I), 1918–1930. https://doi.org/10.1111/j.1475-6773.2005.00438.x

34. Brenk-Franz, K., Hibbard, J., Herrmann, W., Freund, T., Szecsenyi, J., Djalali, S., Steurer-Stey, C., Sönnichsen, A., Tiesler, F., Storch, M., Schneider, N., & Gensichen, J. (2013). Validation of

the German version of the patient activation measure 13 (PAM13-D) in an international multicentre study of primary care patients. *PLoS ONE, 8*(9), 6–11. https://doi.org/10.1371/journal.pone.0074786

35. Bohner, G., Harlacher, U., Rudolf, C., Sieger, H., & Schwarz, N. (1982). Selbstaufmerksamkeit und Körperaufmerksamkeit. In G. Lüer (Ed.), *Bericht über den 33. Kongreß der Deutschen Gesellschaft für Psychologie in Mainz 1982* (pp. 649–652). Hogrefe.

36. Mrazek, J. (1989). Die Erfassung körperbezogener Kontrollüberzeugungen. In G. Krampen (Ed.), *Diagnostik von Attributionen und Kontrollüberzeugungen* (pp. 112–118). Hogrefe.

37. Albani, C., Blaser, G., Geyer, M., Schmutzer, G., Hinz, A., Bailer, H., Grulke, N., & Brähler, E. (2007). Psychometrische Überprüfung und Normierung des Fragebogens Körperbezogener Locus of Control (KLC) an einer repräsentativen deutschen Bevölkerungsstichprobe. *Zeitschrift für Medizinische Psychologie, 16*(1–2), 83–91.

38. Pelikan, J. M., & Ganahl, K. (2017). Measuring health literacy in general populations: Primary findings from the HLS-EU consortium's health literacy assessment effort. *Stud Health Technol Inform, 240*, 34–59.

39. Brähler, E., Mühlan, H., Albani, C., & Schmidt, S. (2007). Teststatistische Prüfung und Normierung der Deutschen Versionen des EUROHIS-QOL Lebensqualität-index und des WHO-5 Wohlbefindens-index. *Diagnostica, 53*(2), 83–96. https://doi.org/10.1026/0012-1924.53.2.83

40. Topp, C. W., Østergaard, S. D., Søndergaard, S., & Bech, P. (2015). The WHO-5 well-being index: A systematic review of the literature. *Psychotherapy and Psychosomatics, 84*(3), 167–176. https://doi.org/10.1159/000376585

41. Quittan, M., Wiesinger, G. F., Crevenna, R., Nuhr, M. J., Posch, M., Hülsmann, M., Müller, D., Pacher, R., & Fialka-Moser, V. (2001). Cross-cultural adaptation of the Minnesota living with heart failure questionnaire for German-speaking patients. *Journal of Rehabilitation Medicine, 33*(4), 182–186. https://doi.org/10.1080/165019701750300654

42. Rector, T. S., & Cohn, J. N. (1992). Assessment of patient outcome with the Minnesota living with heart failure questionnaire: Reliability and validity during a randomized, double-blind, placebo-controlled trial of pimobendan. *American Heart Journal, 124*(4), 1017–1025. https://doi.org/10.1016/0002-8703(92)90986-6

43. Garin, O., Ferrer, M., Pont, A., Wiklund, I., Ganse, E., Vilagut, G., Almansa, J., Ribera, A., & Alonso, J. (2013). Evidence on the global measurement model of the Minnesota living with heart failure questionnaire. *Quality of Life Research, 22*, 2675–2684.

44. Jones, D., Kazis, L., Lee, A., Rogers, W., Skinner, K., Cassar, L., Wilson, N., & Hendricks, A. (2001). Health status assessments using the veterans SF-12 and SF-36: Methods for evaluating outcomes in the veterans health administration. *Journal of Ambulatory Care Management, 24*(3), 68–86. https://doi.org/10.1097/00004479-200107000-00011

45. Buchholz, I., Feng, Y.-S., Buchholz, M., Kazis, L. E., & Kohlmann, T. (2020). The German version of the veterans rand—36/12 item health survey: Translation and adaptation of a generic health assessment instrument in a sample of inpatient rehabilitation patients. *Health Quality of Life Outcomes.* https://doi.org/10.21203/rs.3.rs-98587/v1

46. Devlin, N. J., & Brooks, R. (2017). EQ-5D and the EuroQol group: Past, present and future. *Applied Health Economics and Health Policy, 15*(2), 127–137. https://doi.org/10.1007/s40258-017-0310-5

47. Herdman, M., Gudex, C., Lloyd, A., Janssen, M., Kind, P., Parkin, D., Bonsel, G., & Badia, X. (2011). Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research, 20*(10), 1727–1736. https://doi.org/10.1007/s11136-011-9903-x

48. Skevington, S. M., Sartorius, N., Amir, M., The WHOQOL-Group. (2004). Developing methods for assessing quality of life in different cultural settings—The history of the WHOQOL instruments. *Social Psychiatry and Psychiatric Epidemiology, 39*(1), 1–8. https://doi.org/10.1007/s00127-004-0700-5

49. Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

50. Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

51. Rost, J., & von Davier, M. (1994). A Conditional item-fit index for Rasch models. *Applied Psychological Measurement, 18*(2), 171–182. https://doi.org/10.1177/014662169401800206

52. IBM Corp. (2021). "IBM SPSS Statistics for Windows." IBM Corp, Armonk.

53. Arbuckle, J. L. (2021). *Amos 28.0 user's guide*. IMB SPSS.

54. von Davier, M. (2001). WINMIRA users manual. A program system for analyses with the Rasch model, with latent class analysis and with mixed rasch model [Computer software]. Assessment Systems Corporation.

55. Haslbeck, J., Klein, M., Bischofberger, I., & Sottas, B. (2015). Leben mit chronischer Krankheit. *Die Perspektive von Patientinnen, Patienten und Angehörigen, 46*, 150.

56. Kennedy, A., Rogers, A., & Bower, P. (2007). Support for self care for patients with chronic disease. *British Medical Journal, 335*(7627), 968–970. https://doi.org/10.1136/bmj.39372.540903.94

57. McGilton, K. S., Vellani, S., Yeung, L., Chishtie, J., Commisso, E., Ploeg, J., Andrew, M. K., Ayala, A. P., Gray, M., Morgan, D., Froehlich Chow, A., Parrott, E., Stephens, D., Hale, L., Keatings, M., Walker, J., Wodchis, W. P., Dubé, V., McElhaney, J., & Puts, M. (2018). Identifying and understanding the health and social care needs of older adults with multiple chronic conditions and their caregivers: A scoping review. *BMC Geriatrics, 18*, 1–13.

58. Huygens, M. W. J., Vermeulen, J., Swinkels, I. C. S., Friele, R. D., Van Schayck, O. C. P., & De Witte, L. P. (2016). Expectations and needs of patients with a chronic disease toward self-management and eHealth for self-management purposes. *BMC Health Services Research, 16*(1), 1–11. https://doi.org/10.1186/s12913-016-1484-5

59. Federal Institute for Drugs and Medical Devices, "The fast-track process for digital health applications (DiGA) according to Section 139e SGB V." pp. 1–124, 2019, [Online]. Available: https://www.bfarm.de/EN/MedicalDevices/DiGA/_node.html

60. Tsang, S., Royse, C. F., & Terkawi, A. S. (2017). Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. *Saudi Journal of Anaesthesia, 11*(5), S80–S89. https://doi.org/10.4103/sja.SJA_203_17

61. Frost, M. H., Reeve, B. B., Liepa, A. M., Stauffer, J. W., Hays, R. D., & Sloan, J. A. (2007). What is sufficient evidence for the reliability and validity of patient-reported outcome measures? *Value in Health, 10*(SUPPL. 2), 94–105. https://doi.org/10.1111/j.1524-4733.2007.00272.x

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Klara Greffin[1] · Holger Muehlan[1] · Neeltje van den Berg[2] · Wolfgang Hoffmann[2] · Oliver Ritter[3] · Michael Oeff[4] · Sven Speerfork[5] · Georg Schomerus[5] · Silke Schmidt[1]

✉ Klara Greffin
klara.greffin@uni-greifswald.de

Holger Muehlan
holger.muehlan@uni-greifswald.de

Neeltje van den Berg
neeltje.vandenberg@uni-greifswald.de

Wolfgang Hoffmann
wolfgang.hoffmann@uni-greifswald.de

Oliver Ritter
o.ritter@klinikum-brandenburg.de

Michael Oeff
michael@oeff.eu

Sven Speerfork
sven.speerfork@medizin.uni-leipzig.de

Georg Schomerus
georg.schomerus@medizin.uni-leipzig.de

Silke Schmidt
silke.schmidt@uni-greifswald.de

[1] Department of Psychology, Chair of Health and Prevention, University of Greifswald, Greifswald, Germany

[2] Section Epidemiology of Health Care and Community Health, Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany

[3] Department of Cardiology, Nephrology and Pulmonology, Campus Clinic Brandenburg, Brandenburg Medical School Theodor Fontane, Brandenburg an der Havel, Germany

[4] Brandenburg City Hospital, Brandenburg an der Havel, Germany

[5] Clinic of Psychiatry and Psychotherapy, University of Leipzig Medical Center, Leipzig, Germany