CrossMark

ORIGINAL ARTICLE

German validation of the quality of life profile for spinal disorders (QLPSD)

Tobias L. Schulte[1] · Meinald T. Thielsch[2] · Georg Gosheger[3] · Patrick Boertz[2] ·
Jan Henrik Terheyden[3] · Mark Wetterkamp[1]

Abstract

Background The Quality of Life Profile for Spine Deformities (QLPSD) is a self-reporting questionnaire designed for studying patients with spinal deformities.

Purpose The aim of the present study was to systematically translate the QLPSD into German (G-QLPSD) and to test its reliability and validity. Special emphasis was intended to be given to patients with different Cobb angles and ages.

Methods The QLPSD was systematically translated into German and was responded to in a web-based online survey by patients with idiopathic scoliosis and by healthy control individuals to carry out a matched-pair analysis. Participants aged 14 years and older were included. All participants answered a battery of validated questionnaires (SRS 22-r, PHQ-9, PANAS, FKS, WHO-5, BFI-S, PTQ). Reliability testing included Cronbach's alpha and test–retest reliability (retest 8 weeks after initial testing). Factorial, convergent, divergent, concurrent, and discriminant validity were calculated.

Results A total of 255 scoliosis patients (age 30.0 ± 16.7 years, Cobb angle $43.5° \pm 20.9°$) and 189 matched healthy control individuals were finally included. Cronbach's alpha for the G-QLPSD total score was 0.93 and the test–retest reliability was 0.84. The G-QLPSD total score correlated with the SRS 22-r total score ($r = -0.86$). All concurrently applied scores showed strong correlations with the G-QLPSD (e.g., depression score PHQ-9: $r = 0.70$). The matched-pair analysis of 189 pairs showed strong discriminant validity (Cohen's $d = 0.78$). Patients with more severe Cobb angles ($\geq 40°$) and those ≥ 18 years of age had significantly poorer results than patients with minor curves and younger patients.

Conclusion The G-QLPSD proved to be a highly reliable and valid instrument that can be recommended for clinical use in scoliosis patients.

Keywords Quality of Life Profile for Spine Deformities · QLPSD · Idiopathic scoliosis · German · Translation · Validation · Questionnaire · Quality of life

Tobias L. Schulte and Meinald T. Thielsch contributed equally to this work.

✉ Tobias L. Schulte
  dr.tobias.schulte@web.de

[1] Department of Orthopedics and Trauma Surgery, Katholisches Klinikum Bochum, St. Josef-Hospital, Ruhr University Bochum, Gudrunstr. 56, 44791 Bochum, Germany

[2] Department of Psychology, University of Münster, Münster, Germany

[3] Department of Orthopedics and Tumor Orthopedics, University Hospital Münster, Münster, Germany

Introduction

Idiopathic scoliosis, its course and the success of treatment are commonly assessed using objective quantitative measures, the most frequently used of which is the radiologic magnitude of the curve. However, in addition to purely radiologic measurements, it is becoming increasingly important to give strong consideration to patients' subjective views of their own health and of the effects of treatment, using patient-reported outcome measures, especially in connection with quality of life.

The Quality of Life Profile for Spine Deformities (QLPSD) was introduced by Climent et al. to establish a

quality-of-life instrument for assessing adolescents with spinal deformities [1]. It contains 21 items grouped into five dimensions: psychosocial functioning (seven items), sleep disturbances (four items), back pain (three items), body image (four items), and back flexibility (three items). Optional answers are composed as a five-point Likert scale, with scores from 1 (i.e., "strongly disagree") to 5 (i.e., "strongly agree"). The resulting total score ranges from 21 (i.e., best quality of life) to 105 (i.e., poorest quality of life). The original version of the QLPSD was in Spanish [1], and English [2], French [3], and Greek versions [4] have been subsequently been established and validated. Since then, the QLPSD has been in clinical use and has been widely accepted for assessment of scoliosis and treatment results [5]. It has been used to measure quality of life following surgery or brace treatment for adolescent idiopathic scoliosis and also for Scheuermann kyphosis [2–4, 6–9].

However, a German version of the QLPSD has never previously been systematically introduced or validated. The purpose of the present study was, therefore, to translate the QLPSD into German and to assess its reliability as well as its factorial, convergent, divergent, concurrent, and discriminant validity. Reliability and validity are the two most important criteria to interpret questionnaire scores [10, 11]. Reliability indicates the exactness of measurements, often assessed in terms of internal consistency (e.g., Cronbach's alpha) [10, 12], and temporal stability measured with test–retest correlations [10]. For both criteria, values above 0.7 are desirable [12, 22, 23]. Yet, reliability is a necessary but not sufficient quality criterion, an instrument can be reliable without being valid (for further explanation and examples see Cook et al. [10] and Kimberlin and Winterstein [11]). Thus, most important for practical use is the validity of a given measure, i.e., the ability of a questionnaire to measure the intended construct. As there is no single statistical value that indicates the validity of a measure, several strategies should be applied to examine validity. For example, a common approach is to test the factorial structure of a questionnaire with factor analysis [10]; relations to established measures are often calculated using correlations (expecting high correlations for theoretically highly related measures, and accordingly low correlations to theoretically unrelated measures [22]). Furthermore, statistic procedures such as t tests can be applied to assess discriminant validity, i.e., the ability of a questionnaire to distinguish between patients and healthy individuals [23].

Materials and methods

A professional medical translator systematically translated the original QLPSD questionnaire [1] into German. Afterwards, another professional medical translator

translated the German QLPSD back into English. In a consensus meeting of the authors and translators, the two English versions—the original and the re-translation—were then compared with each other. No relevant differences were found. Minor discrepancies were clarified in a consensus. The final German version of the QLPSD and the scoring instructions are presented in Online Appendix 1.

Patients with idiopathic scoliosis (the scoliosis group) were recruited at the Department of Orthopedics at Münster University Hospital, Germany, and from the self-help group for scoliosis patients in Germany (*Bundesverband Skoliose-Selbsthilfe e.V.*). A healthy control group was established using an online panel named PsyWeb (available via http://psyweb.uni-muenster.de/), organized by the Universities of Münster, Leipzig, and Munich, and the University of Applied Sciences in Osnabrück, Germany, with a total of 12,000 members (as of March 2016).

Only participants with a minimum age of 14 years were invited to voluntarily take part in an online questionnaire. The study intentionally also included adults to be able to identify age effects. The participants did not receive any compensation. Data transfer was encrypted, and the answers given were stored solely in an anonymized form. The local ethics committee approved the study (ref. no. 2014-660-f-S).

Participants were asked for their age, gender, height, weight (body mass index was assessed), academic level, average level of back pain during the previous 6 months on the visual analog scale (VAS), current degree of scoliosis (Cobb angle of the most severe curve), and history of scoliosis treatment as well as current treatment. Data acquisition took place in a self-reported form.

In addition, all of the participants responded to several questionnaires already available and validated in German: the Scoliosis Research Society 22-r (SRS 22-r) Questionnaire [13]; Patient Health Questionnaire (PHQ-9) [14]; the Positive and Negative Affect Schedule (PANAS; only the negative scale applied in this study) [15]; the Questionnaire on Body Dysmorphic Symptoms *Fragebogen körperdysmorpher Symptome* (FKS) [16]; the neuroticism subscale from the GSOEP Big Five Inventory (BFI-S) [17]; and the Perseverative Thinking Questionnaire (PTQ) [18]. Three additional questions created by the authors were also used: (1) Do you think your back's shape will lead to less success in your professional career (job-related worries)? (2) Do you think your back's shape will lead to less satisfaction in your private life (social life-related worries)? Possible answers to these two questions were: definitely not (1)—rather not (2)—maybe (3)—probably yes (4)—definitely yes (5). (3) All in all, how stressed are you by the look of your back (overall stress)? Possible answers: not at all (1)—a little bit (2)—

moderately (3)—very (4)—extremely (5). Due to time restrictions, the WHO-5 Well-Being Index (WHO-5) [19] was only added during the retest.

The web-based questionnaire was available between March 2015 and March 2016. The data were partly used during validation of the G-BIDQ-S [20], but have never before been analyzed in relation to the G-QLPSD.

Statistical analysis was carried out using IBM SPSS Statistics, version 23.

Results

A total of 677 scoliosis patients took part in the study. In total, 149 participants dropped out during answering of the questionnaire; 181 participants reported types of spinal deformity other than idiopathic scoliosis; among those with idiopathic scoliosis, 87 reported a Cobb angle below 10°; and five did not provide consent for their data to be analyzed and thus also had to be excluded. In total, data for 255 patients (37.67%) were included in the current analyses. In addition, 626 individuals were tested as controls (a further 347 participants started the study, but did not complete it), leading to a total of 189 perfectly matched pairs in relation to age (full years) and gender (i.e., 74.12% of analyzed patients could be matched).

The basic data, demographics, and results of the G-QLPSD and the other questionnaires for all of the participants included are shown in Table 1.

Reliability

Reliability was assessed in two ways. Table 2 illustrates internal consistency (i.e., Cronbach's alpha) and test–retest reliability (stability over time). Cronbach's alpha was 0.85 or higher for each subscale and 0.93 for the total score. A retest to check reproducibility was performed on average about 8 weeks after the primary test (on average 55.44 ± 26.32 days). The participants received the G-QLPSD once again, plus—at both measurement points—a few additional measures not pertinent to the current study. There were no significant differences in the means for the G-QLPSD total and subscale scores, except for a small difference in the pain subscale (T_1: 2.82 ± 1.19, T_2: 2.67 ± 1.19; $T = 2.47$, $df = 132$, $P = 0.02$). The retest reliability was $r = 0.84$ ($P < 0.01$) for the G-QLPSD total score and $r > 0.8$ for three of the five subscales (Table 2), indicating good long-term stability with the measure. Altogether, these results indicate that the G-QLPSD is a very reliable instrument.

Factorial validity

An exploratory factor analysis (EFA) was used to test the proposed structure of the G-QLPSD and the homogeneity of its scales. The EFA was intended to allow investigation of the independence of the various subscales within the G-QLPSD. With a value of 0.91, the Kaiser–Meyer–Olkin (KMO) test indicated the high suitability of the data for factor analysis [21]. As the original QLPSD contained five subscales, the EFA was set to extract five factors. These explained 64.81% of the variance, and this factor solution reflected exactly the proposed back pain, body image, and trunk flexibility scales. However, one item from the psychosocial function scale did not show any substantial loading >0.3 (item number 6). Two items (numbers 4 and 8) showed substantial multiple loadings, and one of the two had no loading on the intended scale. The other G-QLPSD items showed loadings between 0.47 and 0.98 on the respective factors (Online Appendix 2).

The G-QLPSD subscales correlated to some extent (Online Appendix 3); the correlation ranged between $r = 0.36$ (back flexibility with body image) and $r = 0.70$ (back pain with sleep disturbances). As these correlations were within an acceptable range, all further analyses were performed with the item arrangement on the five scales as proposed in the original QLPSD.

Convergent validity

Convergent validity is the extent of agreement among theoretically highly related measures [22]. The G-QLPSD and its subscales showed significant moderately to highly negative correlations with each domain in the SRS 22-r (Table 3). Thus, a higher (poorer) G-QLPSD score is associated with a lower (poorer) SRS 22-r score. In addition, high correlations were found for the G-QLPSD total score and subscales with job-related and social life-related worries, with overall stress, and with the VAS pain scale. In particular, the G-QLPDS pain subscale correlated strongly with the SRS 22-r pain scale ($r = -0.76$) and the VAS (pain) ($r = 0.79$); the G-QLPSD body image subscale correlated well with the SRS 22-r self-image subscale ($r = -0.73$); and the G-QLPSD function subscale correlated well with the SRS 22-r function scale ($r = -0.59$). To a somewhat lesser extent, a higher Cobb angle also corresponded with higher G-QLPSD scores, and this was further analyzed in the subgroup differentiation (see below).

Divergent validity

Divergent validity refers to the degree of disagreement between theoretically unrelated (or less related) constructs

Table 1 Basic data, demographics, and results of questionnaires in the scoliosis group and control group

Parameter	Scoliosis group ($n = 255$)	Scoliosis subgroup for matched-pair analysis ($n = 189$)	Controls for matched-pair analysis ($n = 189$)
Age (years)	30.0 ± 16.7	33.6 ± 17.0	33.6 ± 17.0
Gender			
Male	38	17	17
Female	217	172	172
Weight (kg)	63.1 ± 11.6	64.1 ± 11.9	67.7 ± 14.4[b]
Height (cm)	169.6 ± 9.5	168.7 ± 9.0	169.1 ± 7.1[c]
BMI (kg/m^2)	22.0 ± 4.0	22.6 ± 4.2	23.6 ± 4.7[d]
Cobb angle (°)	43.5 ± 20.9	47.8 ± 20.3	–
Academic level			
Secondary school	13	11	2
Junior high	65	47	18
Technical college entry qualification	25	21	10
High School	83	54	67
University degree	69	56	92
Scoliosis treatment[a]			
Physiotherapy	222	160	0
Brace	182	127	0
Surgery	84	73	0
G-QLPSD total score (range 1, i.e., best, to 5, i.e., poorest), mean (sum score)	2.15 ± 0.81 (43.54 ± 16.45)	2.27 ± 0.83	1.70 ± 0.61[e]
G-QLPSD psychosocial functioning, mean (sum score)	1.64 ± 0.78 (11.48 ± 5.47)	1.72 ± 0.81	1.40 ± 0.70[e]
G-QLPSD sleep disturbances, mean (sum score)	2.13 ± 1.02 (8.54 ± 4.09)	2.26 ± 1.07	1.89 ± 0.91[e]
G-QLPSD back pain, mean (sum score)	2.63 ± 1.20 (7.88 ± 3.59)	2.78 ± 1.21	1.95 ± 1.03[e]
G-QLPSD body image, mean (sum score)	2.52 ± 1.20 (10.08 ± 4.80)	2.64 ± 1.19	1.99 ± 0.97[e]
G-QLPSD back flexibility, mean (sum score)	1.85 ± 1.08 (5.56 ± 3.24)	1.99 ± 1.10	1.29 ± 0.65[e]
SRS 22-r score (range 5, i.e., best, to 1, i.e., poorest)			
Overall mean	3.76 ± 0.61		
Function	3.89 ± 0.65		
Pain	3.90 ± 0.90		
Self-image	3.46 ± 0.75		
Mental health	3.77 ± 0.78		
Satisfaction	3.77 ± 0.98		
Job-related worries (range 1, i.e., best, to 5, i.e., poorest)	2.42 ± 1.24		
Social life-related worries (range 1, i.e., best, to 5, i.e., poorest)	2.75 ± 1.29		
Overall stress (range 1, i.e., best, to 5, i.e., poorest)	2.37 ± 1.12		
VAS (pain) (range 0, i.e., best, to 10, i.e., poorest)	4.20 ± 2.61		
PANAS (mood) (range 10, i.e., best, to 50, i.e., poorest)	13.94 ± 5.41		
PHQ-9 (range 0, i.e., best, to 27, i.e., poorest)	4.95 ± 4.63		
FKS (body dysmorphic disorder) (range 0, i.e., best, to 64, i.e., poorest)	12.17 ± 9.57		
BFI-S (neuroticism) (range 1. i.e., best to 7, i.e., poorest)	3.71 ± 1.52		

Table 1 continued

Parameter	Scoliosis group ($n = 255$)	Scoliosis subgroup for matched-pair analysis ($n = 189$)	Controls for matched-pair analysis ($n = 189$)
PTQ (negative thinking) (range 0, i.e., best, to 48, i.e., poorest)	18.99 ± 12.43		
WHO-5 (well-being) (range 25, i.e., best, to 0, i.e., poorest)	13.50 ± 5.75 ($n = 133$)[f]		

BFI-S neuroticism subscale of the Big Five Inventory-Short, *BMI* body mass index, *FKS* Fragebogen körperdysmorpher Symptome (Questionnaire on Body Dysmorphic Symptoms), *G-QLPSD* German Quality of Life Profile for Spine Deformities, *PANAS* Positive And Negative Affect Schedule, *PHQ-9* Patient Health Questionnaire, *PTQ* Perseverative Thinking Questionnaire, *SRS* Scoliosis Research Society, *VAS* Visual Analog Scale, *WHO-5* the WHO-5 Well-Being Index

[a] Binomials included

[b] Difference in weight is significant ($T = -2.63$, $df = 376$, $P < 0.01$), effect size is rather small ($d = -0.27$)

[c] Difference in height is not significant ($T = -0.50$, $df = 374$, $P = 0.62$)

[d] Difference in BMI is significant ($T = -2.24$, $df = 374$, $P = 0.03$), effect size is rather small ($d = -0.23$)

[e] Differences between scoliosis subgroup and controls significant at $P < 0.01$; see subsection on discriminant validity

[f] Only given during the secondary test

Table 2 Reliability of the G-QLPSD

	G-QLPSD total score	G-QLPSD psychosocial functioning	G-QLPSD sleep disturbances	G-QLPSD back pain	G-QLPSD body image	G-QLPSD back flexibility	n
Cronbach's alpha	0.93	0.86	0.85	0.87	0.88	0.89	255
Test–retest reliability	0.84**	0.63**	0.84**	0.83**	0.73**	0.81**	133

** $P < 0.01$

[22]. With regard to this aspect of validity, the G-QLPSD correlated at a low level and only partly significantly with the BMI; as expected, this connection was weak (Table 3).

Concurrent validity

Concurrent validity refers to the ability of a measure to predict a concurrently assessed criterion [23]. All concurrently evaluated criteria (PANAS, PHQ-9, FKS, WHO-5 PTQ, and BFI-S) (Table 3) showed strong correlations with the G-QLPSD scores. Particularly notable are the sometimes very high correlations with the depression score (G-QLPSD total score and PHQ-9, $r = 0.70$), high correlation between G-QLPSD body image and FKS ($r = 0.71$), and strong negative correlations with the Well-Being Index (WHO-5): e.g., G-QLPSD total score and WHO-5 $r = -0.65$.

Discriminant validity

Discriminant validity refers to the ability of the G-QLPSD and its subscales to distinguish between patients with scoliosis and individuals in a healthy control group. This comparison between the scoliosis and the control group in

the matched-pair analysis (Table 1) showed a clear difference in the G-QLPSD total score ($F = 14.88$, $df = 1$, 376, $P < 0.01$) and all five subscales ($13.30 \leq F \leq 56.66$, $df = 1$, 376, $P < 0.01$). With a Cohen's $d = 0.78$ for the G-QLPSD total score, the effect size of this group difference can be considered to be large; subscale differences are medium to large ($0.42 \leq d \leq 0.77$)[1] [24].

Subgroup analysis: Cobb angle and age

Table 4 comprises data from a subgroup analysis of patients with Cobb angles of less than 40° and those with ≥40°, as well as the correlation of the G-QLPSD with age and mean differences between adolescent and adult patients. G-QLPSD values are increased in patients with Cobb angles ≥40° and adults in general; age correlations showed mostly medium effect sizes.

---

[1] According to the guidelines provided by Cohen [24], standardized mean differences of 0.2, 0.5, and 0.8 and more are considered to represent small, medium, and large effects, respectively.

Table 3 Correlations for convergent, divergent, and concurrent validity

	Correlation with G-QLPSD total score	Correlation with G-QLPSD psychosocial functioning	Correlation with G-QLPSD sleep disturbances	Correlation with G-QLPSD back pain	Correlation with G-QLPSD body image	Correlation with G-QLPSD back flexibility	n
SRS 22-r							
Overall mean	−0.86**	−0.68**	−0.72**	−0.73**	−0.62**	−0.58**	255
Function	−0.72**	−0.59**	−0.56**	−0.60**	−0.43**	−0.61**	255
Pain	−0.71**	−0.42**	−0.70**	−0.76**	−0.33**	−0.051**	255
Self-image	−0.72**	−0.61**	−0.51**	−0.50**	−0.73**	−0.45**	255
Mental health	−0.64**	−0.62**	−0.58**	−0.50**	−0.49**	−0.34**	255
Satisfaction	−0.40**	−0.31**	−0.29**	−0.29**	−0.39**	−0.25**	255
Job-related worries	0.56**	0.46**	0.42**	0.44**	0.45*	0.41**	255
Social life-related worries	0.59**	0.50**	0.38**	0.45**	0.62**	0.32**	255
Overall stress	0.59**	0.47**	0.39**	0.37**	0.68**	0.32**	255
VAS (pain)	0.73**	0.46**	0.67**	0.79**	0.38**	0.48**	255
Cobb angle	0.28**	0.13*	0.13*	0.21**	0.24**	0.32**	255
BMI	0.28**	0.08	0.23**	0.26**	0.26**	0.19**	253
PANAS (mood)	0.59**	0.54**	0.48**	0.41**	0.50**	0.36**	255
PHQ-9 (depression)	0.70**	0.66**	0.63**	0.52**	0.49**	0.43**	255
FKS (body dysmorphic disorder)	0.57**	0.54**	0.38**	0.33**	0.71**	0.23**	255
WHO-5 (well-being)	−0.65**	−0.63**	−0.57**	−0.49**	−0.50**	−0.41**	133
PTQ (negative thinking)	0.44**	0.51**	0.33**	0.30**	0.42**	0.20**	255
BFI-S (neuroticism)	0.45**	0.47**	0.38**	0.35**	0.39**	0.18**	255

Convergent validation = SRS-22r, BIDQ-S, job-related worries, social life-related worries, overall stress, VAS, Cobb angle. Divergent validation = BMI. Concurrent validation = PANAS, PHQ-9, FKS, WHO-5, PTQ, BFI-S

*$P < 0.05$; **$P < 0.01$

Table 4 Subgroup and age analysis

	G-QLPSD total score	G-QLPSD psychosocial functioning	G-QLPSD sleep disturbances	G-QLPSD back pain	G-QLPSD body image	G-QLPSD back flexibility
Cobb angle 10°–39° (n = 133)	1.92 ± 0.69	1.51 ± 0.71	1.96 ± 0.91	2.39 ± 1.13	2.24 ± 1.05	1.49 ± 0.80
Cobb angle ≥40° (n = 122)	2.41 ± 0.87	1.79 ± 0.83	2.32 ± 1.10	2.89 ± 1.21	2.82 ± 1.29	2.25 ± 1.21
G-QLPSD correlations with age (n = 255)	0.42**	0.20**	0.33**	0.37**	0.29**	0.40**
Age 14 to <18 years (n = 59)	1.76 ± 0.66	1.47 ± 0.73	1.72 ± 0.85	2.06 ± 1.01	2.07 ± 1.00	1.48 ± 0.82
Age ≥18 years (n = 196)	2.27 ± 0.82	2.26 ± 1.04	2.26 ± 1.04	2.80 ± 1.20	2.65 ± 1.22	1.97 ± 1.12

Unless otherwise indicated, means and standard deviations of the G-QLPSD are presented. All differences between Cobb angle and age subgroups are significant with $P < 0.01$; for correlations shown: *$P < 0.05$, **$P < 0.01$

Discussion

In summary, the German version of the QLPSD (G-QLPSD) proved to be a reliable and valid instrument and can therefore be recommended for everyday use in treating scoliosis patients. Including 255 patients, the present study is the largest available on the use of the QLPSD in scoliosis patients. In doing so, we not only found mostly equal or even higher values for reliability and validity compared with the original QLPSD [1] but also additional prove for validity of the G-QLPSD that goes beyond the analysis performed on the original questionnaire.

To allow better comparison of the subscales with one another, QLPSD mean scores were calculated: G-QLPSD total score 2.15, psychosocial functioning 1.64, sleep disturbances 2.13, back pain 2.63, body image 2.52, back flexibility 1.85. These results correspond to the following sum scores in relation to the available literature: G-QLPSD total score 43.54 (36.61–44.57 [1]; 42.8–53.6 [6]; 32.2–48 [3]); psychosocial functioning 11.48 (10.44–13.37 [1]; 11.2–15.9 [6]); sleep disturbances 8.54 (5.71–6.93 [1]; 9.1–9.8 [6]); back pain 7.88 (6.16–7.16 [1]; 6.3–7.3 [6]), body image 10.08 (9.06–9.48 [1]; 9.3–10.1 [6]; 11.42 [25]); and back flexibility 5.56 (5.22–7.61 [1]; 5.6–11.4 [6]; 4.7–6.7 [8]).

Reliability

In the original paper, Climent et al. described an internal consistency measured with Cronbach's alpha of 0.88 (in the present study: 0.93) for the QLPSD total score, 0.81 (0.86) for psychosocial functioning, 0.84 (0.85) for sleep disturbances, 0.75 (0.87) for back pain, 0.70 (0.88) for body image, 0.70 (0.89) for back flexibility, and a test–retest correlation (intraclass correlation coefficient) of 0.91 (test–retest reliability in the present data: 0.84) for the total score, 0.89 (0.63) for psychosocial functioning, 0.78 (0.84) for sleep disturbances, 0.91 (0.83) for back pain, 0.66 (0.73) for body image, and 0.67 (0.81) for back flexibility [1]. Feise et al. reported similar data from their trial: Cronbach's alpha: 0.91 QLPSD total score, 0.82 psychosocial functioning, 0.86 sleep disturbances, 0.84 back pain, 0.86 body image, 0.69 back flexibility, intraclass correlation coefficients: 0.91 QLPSD total score; 0.59 psychosocial functioning, 0.76 sleep disturbances, 0.88 back pain, 0.87 body image, 0.87 back flexibility [26]. Matamalas et al. described a Cronbach's alpha of 0.80 for the body image subscale of the QLPSD in their scoliosis cohort [25].

These data prove that the G-QLPSD offers strong reliability, corresponding well with the original version and the available literature.

Validity

Exploratory factorial analysis showed that with only very few exceptions, nearly all of the items presented a strong factor loading on the intended subscale. This can serve as an indicator of good factorial validity of the G-QLPSD. However, it needs to be borne in mind that the G-QLPSD subscales are not fully independent and showed medium to large intercorrelations. Comparable subscale correlations in the QLPSD were reported by Feise et al. [26]. In the present data, the strong correlation of G-QLPSD back pain and G-QLPSD sleep disturbances in particular, at $r = 0.70$, is noteworthy.

With regard to convergent validity, the G-QLPSD and its subscales showed significant moderate to highly negative correlations with the SRS 22-r. The G-QLPSD total score correlated significantly with the SRS 22-r overall score ($r = -0.86$). In particular, the G-QLPSD pain subscale correlated well with the SRS 22-r pain subscale ($r = -0.76$), the G-QLPSD body image subscale correlated well with the SRS 22-r self-image subscale ($r = -0.73$), and the function subscales correlated with an $r = -0.59$. This corresponds well with the findings reported by Climent et al., who noted a significant correlation coefficient between the QLPSD total score and the SRS-22 total score of 0.84 [27]. In their study, the correlation coefficients of common dimensions (pain, function, image) between the two scores were 0.85, 0.52 and 0.62. Corresponding well with these data, Matamalas et al. found a correlation of $r = -0.76$ between the QLPSD body image subscale and the SRS 22 self-image subscale [25].

In addition, in the present analysis, high correlations for the G-QLPSD total score and subscales with job-related and social life-related worries, and with overall stress and pain (VAS), also indicate good convergent validity.

The expected weak relation of the BMI with the G-QLPSD was verified in terms of divergent validity.

With regard to concurrent validity, the concurrently evaluated criteria (PANAS, PHQ-9, FKS, WHO-5, PTQ, BFI-S) showed strong correlations with the G-QLPSD. In particular, the high correlation of the G-QLPSD total score with the PHQ-9 depression score ($r = 0.70$) and the WHO-5 Well-Being score ($r = -0.65$), as well as the high correlation of the G-QLPSD body image subscale with the FKS (body dysmorphic disorder), at $r = 0.71$, confirmed strong concurrent validity. These findings might also serve as a starting-point for future research on quality of life in scoliosis patients and possible psychological strains.

The discriminant validity testing comparing matched pairs of patients and controls, including a very large group of 189 pairs, clearly showed that the G-QLPSD is able to differentiate between patients and controls. This is crucial for clinical practice.

Subgroup analysis: Cobb angle and age

The mean Cobb angle of $43.5° \pm 20.9°$ in the scoliosis group represents a wide range of severities of scoliosis. In the present study, Cobb angles were grouped below 40° versus 40° or more, to compare patients who are candidates for conservative treatment with those who are likely to have surgery. Comparing patients with Cobb angles of less than 40° with those of 40° or more clearly showed that the G-QLPSD (total score and subscales) differed significantly between these groups, with poorer scores in patients with more severe deformity. Significant correlations were found between the G-QLPSD total score, as well as all subscales, and the Cobb angle. Climent et al. did not find any statistically significant correlations between the QLPSD scores and the size of the Cobb angle in their cohort of patients, with a mean Cobb angle of 21° [1]. However, they also found significantly higher—i.e., poorer—QLPSD scores in patients with structural curves in comparison with those with postural curves (except for body image). Matamalas et al. reported a significant correlation between the Cobb angle and the QLPSD body image subscale ($r = 0.36$), finding a sum score of 10.4 for the QLPSD body image subscale in patients with Cobb angles <45°, in comparison with 12.4 in patients with ≥45° ($P = 0.05$) [25].

One difference between the present study and comparable investigations of the QLPSD is the fact that it included adolescents as well as adult participants. This was done on purpose to validate the questionnaire for all age groups. When constructing the original QLPSD, Climent et al. only included adolescents [1]. Our correlation analysis of the G-QLPSD total score and subscales with age clearly indicated a significant effect of age on the G-QLPSD results—i.e., with poorer results in older patients. Matamalas et al. differentiated patients <18 years from those ≥18 years, and found a sum score for the QLPSD body image subscale of 10.8 versus 12.11. However, the correlation with the Cobb angle was significant only for the older group ($r = 0.47$), not for the younger group ($r = 0.2$) [25]. The present results correspond well with these data, showing better QLPSD results in patients under the age of 18 in comparison with those aged 18 or over.

Another difference between the findings reported in most of the available literature on QLPSD and the present study is the fact that instead of using sum scores for the G-QLPSD and its subscales, we used mean values to make the data comparable between the scales. This technique has also been applied before for the QLPSD by Zeh et al. [28]. Using mean scores has the additional effect that all subscales were given the same weight in the total score.

From a practical point of view one might ask at which situation one should use the G-QLPSD or a comparable

quality of life measure for scoliosis patients such as the SRS-22. In general, it is important to monitor the experienced quality of life of scoliosis patients during and after treatment to provide optimal medical and, if needed, psychological support. In doing so, the German versions of QLPSD and SRS-22 are comparable in terms of length and reliability, yet more systematic validation data in German are available for the G-QLPSD. A practical guide could be to consider which assessed aspects are most important for the treatment of a specific patient: Both questionnaires ask for patients' evaluations of pain, body image and psychosocial functioning. Yet, the SRS-22 follows a more global screening approach, additionally assessing mental health and satisfaction, while the QLPSD is a little more specific, additionally asking for sleep disturbances and back flexibility. Thus, if a screening is needed the SRS-22 might better fit due to it's more global approach. If quality of life should be assessed in more detail with a highly valid instrument the G-QLPSD is to prefer. Especially in the treatment of patients the additional aspects asked in the QLPSD (such as sleep quality) could be of high interest. Finally, one has to keep in mind that the original QLPSD was constructed and validated for adolescents [1]. Yet, we successfully expanded the age range and validated the German version for adults as well.

One of the limitations of this study is the fact that no radiographic data for the patients were analyzed, due to the self-reporting online format of the study. The web-based self-reporting character of the study involves a risk of incorrect or misleading answers being given, in addition to misunderstandings. Nevertheless, this study format makes it possible to include large groups of participants. The results of the subgroup analyses showed that samples as large as this are required to identify the effects of variables such as age. In addition, a sample size >250 is recommended when seeking stable estimates of correlations [29]. With regard to factor analysis, most sample size requirements for producing a reliable factor solution were met in the present study, although definitive identification of a multifactorial model might require larger sample sizes [30, 31] and the five-factor solution found for the G-QLPSD should be tested in the future in a confirmatory factor analysis.

In summary, the G-QLPSD (Online Appendix 1) is a simple and feasible, valid and reliable questionnaire that provides valuable information for physicians and patients in the management of scoliosis.

Conclusion

- This is the first publication of a systematically translated German version of the QLPSD (G-QLPSD),

including a sophisticated analysis of reliability and validity.

- The G-QLPSD proved to be a highly reliable instrument.
- The G-QLPSD showed strong factorial, convergent, divergent, concurrent, and discriminant validity.
- Older patients and patients with Cobb angles of $\geq 40°$ showed poorer G-QLPSD results.
- The G-QLPSD can be recommended for clinical use to evaluate quality of life in patients with spinal deformities.

References

1. Climent JM, Reig A, Sánchez J, Roda C (1995) Construction and validation of a specific quality of life instrument for adolescents with spine deformities. Spine (Phil Pa 1976) 20(18):2006–2011

2. Howard A, Donaldson S, Hedden D, Stephens D, Alman B, Wright J (2007) Improvement in quality of life following surgery for adolescent idiopathic scoliosis. Spine (Phil Pa 1976) 32(24):2715–2718

3. Pham VM, Houlliez A, Carpentier A, Herbaux B, Schill A, Thevenon A (2008) Determination of the influence of the Chêneau brace on quality of life for adolescent with idiopathic scoliosis. Ann Réadaptation Med Phys 51:3–8

4. Korovessis P, Zacharatos S, Koureas G, Megas P (2007) Comparative multifactorial analysis of the effects of idiopathic adolescent scoliosis and Scheuermann kyphosis on the self-perceived health status of adolescents treated with brace. Eur Spine J 16:537–546

5. Bagó J, Climent JM, Pérez-Grueso FJS, Pellisé F (2013) Outcome instruments to assess scoliosis surgery. Eur Spine J 22(Suppl 2):S195–S202

6. Climent JM, Sánchez J, Group for the Study of QoL in Spine Deformities (1999) Impact of the type of brace on the quality of life of adolescents with spine deformities. Spine (Phil Pa 1976) 24(18):1903–1908

7. Davies E, Norvell D, Hermsmeyer J (2011) Efficacy of bracing versus observation in the treatment of idiopathic scoliosis. Evid Based Spine Care J 2(2):25–34

8. Sanchez-Raya J, Bago J, Pellise F, Cuxart A, Villanueva C (2012) Does the lower instrumented vertebra have an effect on lumbar mobility, subjective perception of trunk flexibility, and quality of life in patients with idiopathic scoliosis treated by spinal fusion? J Spinal Disord Tech 25(8):437–442

9. Smith PL, Donaldson S, Hedden D, Alman B, Howard A, Stephens D, Wright JG (2006) Parents' and patients' perception of postoperative appearance in adolescent idiopathic scoliosis. Spine (Phil Pa 1976) 31(20):2367–2374

10. Cook DA, Beckman TJ (2006) Current concepts in validity and reliability for psychometric instruments: theory and application. Am J Med 119(2):166-e7–16

11. Kimberlin CL, Winterstein AG (2008) Validity and reliability of measurement instruments used in research. Am J Health Syst Pharm 65(23):2276–2284. doi:10.2146/ajhp070364

12. Tavakol M, Dennick R (2011) Making sense of Cronbach's alpha. Int J Med Educ 27(2):53–55. doi:10.5116/ijme.4dfb.8dfd

13. Niemeyer T, Schubert C, Halm HF, Herberts T, Leichtle C, Gesicki M (2009) Validity and reliability of an adapted German Version of Scoliosis Research Society-22 Questionnaire. Spine (Phila Pa 1976) 34(8):818–821. doi:10.1097/BRS.0b013e31819b33be

14. Kroenke K, Spitzer RL, Williams JBW (2001) The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 16(9):606–613

15. Watson D, Clark LA, Tellegen A (1988) Development validation of brief measures of positive and negative affect. The PANAS scales. J Personal Soc Psychol 54(6):1063–1070

16. Buhlmann U, Wilhelm S, Glaesmer H, Braehler E, Rief W (2009) Fragebogen körperdysmorpher symptome (FKS): Ein screening-instrument. Verhaltenstherapie 19:237–242. doi:10.1159/000246278

17. Gerlitz JY, Schupp J (2005) Zur Erhebung der Big-Five-basierten Persoenlichkeitsmerkmale im SOEP. Deutsches Institut für Wirtschaftsforschung (DIW Research Notes, 4), Berlin. Available at: http://www.diw.de/documents/publikationen/73/diw_01.c.43490.de/rn4.pdf. Accessed 18 Sept 2016

18. Ehring T, Zetsche U, Weidacker K, Wahl K, Schoenfeld S, Ehlers A (2011) The Perseverative Thinking Questionnaire (PTQ): validation of a content-independent measure of repetitive negative thinking. J Behav Ther Exp Psychiatry 42(2):225–232

19. Topp CW, Østergaard SD, Søndergaard S, Bech P (2005) The WHO-5 Well-Being Index: a systematic review of the literature. Psychother Psychosom 84(3):167–176. doi:10.1159/000376585

20. Wetterkamp M, Thielsch MT, Gosheger G, Boertz P, Terheyden JH, Schulte TL (2017) German validation of the BIDQ-S Questionnaire on body image disturbance in idiopathic scoliosis. Eur Spine J 26(2):309–315. doi:10.1007/s00586-016-4895-4

21. Dziuban CD, Shirkey EC (1974) When is a correlation matrix appropriate for factor analysis? Some decision rules. Psychol Bull 81:358–361

22. Nunnally JC, Bernstein IH (1994) Psychometric theory, 3rd edn. McGraw-Hill, New York

23. Anastasi A, Urbina S (1997) Psychological testing, 7th edn. Prentice-Hall, Upper Saddle River

24. Cohen J (1988) Statistical power analysis for the behavioral sciences, 2nd edn. Erlbaum, Hillsdale

25. Matamalas A, Bagó J, D'Agata E, Pellisé F (2014) Body image in idiopathic scoliosis: a comparison study of psychometric properties between four patient-reported outcome instruments. Health Qual Life Outcomes 12:81

26. Feise RJ, Donaldson S, Crowther ER, Menke JM, Wright JG (2005) Construction and validation of the Scoliosis Quality of Life Index in adolescent idiopathic scoliosis. Spine (Phil Pa 1976) 30(11):1310–1315

27. Climent JM, Bago J, Ey A, Perez-Grueso FJ, Izquierdo E (2005) Validity of the Spanish Version of the Scoliosis Research Society-22 (SRS-22) Patient Questionnaire. Spine (Phil Pa 1976) 30(6):705–709

28. Zeh A, Planert M, Klima S, Hein W, Wohlrab D (2008) The flexible Triac™-Brace for conservative treatment of idiopathic scoliosis. An alternative treatment option? Acta Orthop Belg 74:512–521

29. Schoenbrodt FD, Perugini M (2013) At what sample size do correlations stabilize? J Res Personal 47:609–612. doi:10.1016/j.jrp.2013.05.009

30. Beavers AS, Lounsbury JW, Richards JK, Huck SW, Skolits GJ, Esquivel SL (2013) Practical considerations for using exploratory factor analysis in educational research. Pract Assess Res Eval 18(6):1–13. Available at: http://pareonline.net/getvn.asp?v=18&n=6. Accessed 14 Mar 2016

31. Hirschfeld G, von Brachel R, Thielsch MT (2014) Selecting items for Big Five Questionnaires: at what sample size do factor loadings stabilize? J Res Personal 53:54–63. doi:10.1016/j.jrp.2014.08.003