

Ger J Exerc Sport Res 2021 · 51:474–486  
<https://doi.org/10.1007/s12662-021-00776-w>  
 Received: 15 April 2021  
 Accepted: 6 October 2021  
 Published online: 1 November 2021  
 © The Author(s) 2021



Nadja Schott · Bettina Johnen · Thomas Jürgen Klotzbier

Department of Sport and Exercise Science, University of Stuttgart, Stuttgart, Germany

# Assessing the well-being of residents in nursing facilities

## Translation and validation of the German version of the Laurens Well-being Inventory for Gerontopsychiatry (LWIG-GER)

### Supplementary Information

The online version of this article (<https://doi.org/10.1007/s12662-021-00776-w>) contains supplementary material, which is available to authorized users.

### Introduction

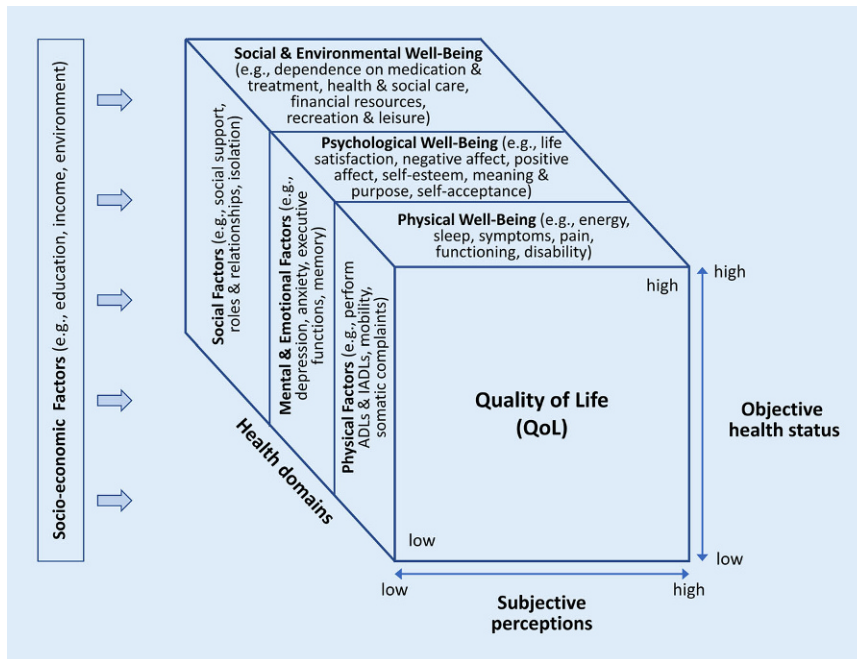
Several hundred thousand people live in German care facilities. At the end of the first half of 2020, around 731,000 people were cared for in Germany's full inpatient care facilities. According to this, the number of residents in the facilities has risen steadily over the past 10 years: from 620,249 in 2010 to 676,584 in 2015 and now around 731,000. Among 80- to 85-year-olds, one in four people already need care (Statistisches Bundesamt, 2020). If the number of older adults in the population increases, the number of people in need of care and assistance is also likely to rise. Almost two-thirds of people in Germany suffer from three or more chronic diseases associated with physical and cognitive limitations in old age (RKI, 2015). However, there is no linear relationship between such limitations and subjectively perceived quality of life (QoL) and well-being (WB); QoL and WB can be subjectively experienced as high despite physical or cognitive limitations (Megari, 2013). Therefore, the maintenance and improvement of QoL are increasingly seen as essential tasks for health care.

Numerous studies and surveys show that people's desires and expectations in

need of care and their relatives are focused on providing good care and including WB factors such as consideration of individuality and participation, and respect for human dignity (García & Ramírez, 2018). Over the years, this has been recognized, and WB and QoL have become central in caring for older people and an essential topic for research (van Leeuwen et al., 2019). While the concept of QoL finds its scientific origins in various disciplines, especially in philosophy, sociology, and medicine, the term WB originates from psychology (Schumacher, Klaiberg, & Brähler, 2003). Earlier research focused mainly on QoL's social and economic indicators, such as financial security, social equity, health care. More recently, subjective indicators of QoL, such as subjective WB and life satisfaction, have been increasingly considered (Diener & Suh, 1997).

Although research on QoL and WB has a long history, there is considerable debate in the literature about what is meant by QoL and WB. Attempts to describe and define QoL and WB in a universally valid way usually failed because they are individual and influenced by age and related circumstances (Opikofor & Mayorova, 2016). Moreover, a clear distinction between the two concepts is hardly possible; sometimes they are used synonymously, or aspects of WB are used to define QoL; although they are associated with different theoretical concepts (e.g., Ryff's Psychological Well-being, Ryff & Keyes, 1995): eudaimonic well-being emphasizing the realization of

a person's potential (Ryan & Deci, 2001); Diener's (2000) hedonic approach of Subjective Well-Being; PERMA model (Seligman, 2018); Wilson & Cleary Model of HRQoL (1995) (for an extensive discussion see Das, Jones-Harrell, & Fan, 2020; and Skevington & Böhnke, 2018). Subjective WB and subjective QoL are overarchingly described as key concepts that describe experiences, abilities, states, behaviors, appraisals, and emotional responses to circumstances. Perhaps most widely used is the concept of QoL as a multidimensional construct that encompasses physical, emotional, mental, social, and everyday functional factors (Bullinger, 2014; Koller et al., 2009). Health-related QoL is defined as how subjective QoL is influenced by physical and mental health (Karimi & Brazier, 2016). Subjective WB (SWB) is composed of general life satisfaction, satisfaction with individual life domains, the presence of positive feelings, and a low level of negative feelings (Diener, 2000). In a recent study using network analysis, Skevington and Böhnke (2018) proposed an integrated model of SWB and QoL through testing its overlapping and exclusive dimensions, which resulted in 14 facets: energy, sleep (physical domain); positive feelings, self-esteem (psychological); dependence on medication, & treatment (independence); personal relationships, social support, sex-life (social relationships); home environment, financial resources, health & social care, recreation & leisure (environment); wholeness, inner peace



**Fig. 1** ▲ Conceptual model for quality of life and well-being. QoL Quality of Life, ADL Activities of Daily Living, IADL Instrumental Activities of Daily Living. (Modified after Testa & Simonson, 1996)

(spiritual). In addition, they showed that health-related variables are closely linked to physical QoL facets on medication, activity, and mobility. Testa and Simonson (1996) proposed an interesting conceptual framework of the subjective and objective domains to unravel the different components of QoL and WB. Each domain of QoL can be viewed from two dimensions: as an objective examination of a person's health status, based on an assessment of their physical, mental, and social functioning, and as a subjective perception of that status, in terms of physical, mental, and social well-being.

Figure 1 shows a modified conceptual model of their QoL framework supplemented by socioeconomic (objective) factors and the classification into physical, psychological, and social well-being. The x-axis represents the subjective, the y-axis the objective dimensions of the three central domains of the concept of QoL. The concept of QoL spans between the two axes. The objective dimension of QoL is to be understood as a gauge of the state of health in physical, mental, and social terms, which is, however, only transformed into the concept of QoL by the person's subjective judgment. Using the example of the physical domain, this

means that the frequency and severity of limitations and their relevance from the individual's point of view are important (Fig. 1).

However, there remains ongoing debate about how QoL and WB should be conceptualized and measured, and therefore, a variety of approaches exist to measure QoL and WB in surveys. More recently, Lindert, Bain, Kubzansky, and Stein (2015) identified 60 unidimensional and multidimensional instruments to survey subjective WB alone. Some of these are generic, i.e., appropriate for the general population, and can be applied to various conditions (e.g., diabetes, cancer; Ryff's Scales of Psychological Well-being (Ryff & Keyes, 1995); WHO (Five) Well-being Index (Bech, 1996); WHOQOL-100 (WHOQOL, 1998)), while others are disease-specific and relate to a particular pathology (e.g., depression, Parkinson's disease; Beck Depression Inventory II (Beck, Steer, Ball, & Ranieri, 1996); Parkinson's disease questionnaire (Hagell & Nygren, 2007)). However, few of these instruments have been developed specifically for older "healthy" adults; instead, there are mainly questionnaires for individuals with dementia (e.g., Qualidem). Also,

no instruments are available specifically for the nursing home setting.

Previous studies on older adults in a nursing home setting (Ballmer, Wirz, & Gantschnig, 2019; van der Wolf, van Hooren, Waterink, & Lechner, 2019) often use questionnaires designed for young and middle-aged adults to evaluate WB or QoL, such as the SF-36 and its short form the SF-12 by Bullinger and Kirchberger (1998), the World Health Organization Quality of Life 100 (WHOQOL-100), or the 5-item World Health Organization Well-Being Index (WHO-5). These questionnaires cover medical and physical aspects, while social and psychological factors are only marginally addressed. However, they do not address individual priorities and needs arising from different life experiences (Meyer, Drewniak, Hovorka, & Schenk, 2019). Another problem is that the questionnaires, which are also used for community-dwelling individuals, often do not correspond to the nursing home residents' living environment. For example, the SF-36 or SF-12 asks whether there are limitations to activities such as running fast, lifting heavy objects, or engaging in strenuous sports due to current health status, to name just a few examples. One of the few questionnaires developed for nursing home residents is the Qualidem (Ettema, Dröes, de Lange, Mellenbergh, & Ribbe, 2007). In this questionnaire, residents are not directly interviewed, but their close caregivers, such as relatives or caregivers, can assess their situation. However, considering the resident's perspective is important as the third party to self-assessment can be very different (Oppikofer & Mayorova, 2016).

The Laurens Well-Being Inventory for Gerontopsychiatry (LWIG; van der Wolf, van Hooren, Waterink, & Lechner, 2018) is a self-rated 30-item well-being measure (that relate to the last 7 days) for individuals residing in a nursing home driven by both theoretical and data-driven considerations. In creating their scale, van der Wolf et al. (2018) followed the WHO view that at least the physical, social, and psychological dimensions should be included (WHOQOL Group, 1995; see Fig. 1). Two models that,

when used together, cover and explain these dimensions are the Social Production Function (SPF) model (Lindenberg, 1986; for both the physical and social domains) and Ryff's model of psychological well-being (Ryff & Keyes, 1995; for the psychological domain). Based on previous matching question instruments, interviews with geriatric professionals, and focus groups, the authors created an item pool from which selections were made based on their chosen theoretical models. Further decisions about whether to retain or reject items in this process were based on an empirical study of nursing home residents, reducing the initial pool of more than 300 possible items to 30 items. The authors of the original scale (van der Wolf et al., 2018) analyzed the proposed 3-factor structure of wellbeing through confirmatory analysis (PCA), but only on the questionnaire with 53 variables (preliminary stage). Overall, the fit statistics did not show a perfect model (CFI = 0.688, TLI = 0.675, RMSEA = 0.063). In addition, they examined whether one or more factors were present within the three different dimensions. The final instrument contains six subscales in three dimensions: physical well-being' (6 items: items 1, 7, 13, 19, 24, 27), social well-being' (subscales positive social experience, 6 items: 4, 10, 16, 22, 26, 29; negative social experience, 4 items: 5, 11, 17, 23; communal living, 3 items: items 6, 12, 18), and psychological well-being' (subscales affect, 7 items: 2, 8, 14, 20, 2, 28, 30; self-worth, 4 items: 3, 9, 15, 21). The LWIG has demonstrated adequate reliability and validity (van der Wolf et al., 2018). Reliability of the dimensions and their underlying factors was assessed using McDonald  $\omega$  and Cronbach  $\alpha$ : All but one subscale "negative social experience" had values  $>0.70$ . Although self-rated and observer-rated WB is generally weakly correlated, the LWIG subscales were significantly related to most of the observer-rated Qualidem subscales and the Cantril Ladder (Ettema et al., 2007; van der Wolf et al., 2018). In addition, older adults with depressive symptoms scored significantly lower on all subscales of the LWIG compared to individuals with no depression.

Ger J Exerc Sport Res 2021 · 51:474–486 <https://doi.org/10.1007/s12662-021-00776-w>  
© The Author(s) 2021

N. Schott · B. Johnen · T. J. Klotzbier

## Assessing the well-being of residents in nursing facilities. Translation and validation of the German version of the Laurens Well-being Inventory for Gerontopsychiatry (LWIG-GER)

### Abstract

**Purpose.** Well-adapted and validated well-being (WB) instruments for the nursing home population are scarce. To our knowledge, the Laurens Well-Being Inventory for Gerontopsychiatry (LWIG) is a practical and reliable well-being assessment tool that has never been validated for German nursing home populations. Therefore, the purpose of our study was to (1) translate and cross-culturally adapt the LWIG to a German context and (2) test the reliability and validity of the German LWIG in a group of older nursing home residents using the Rasch model.

**Methods.** This study has a cross-sectional, descriptive study design. Cross-cultural adaption of the LWIG-GER from English to German was performed according to a standardized method. The data obtained from 104 long-term nursing home residents (57 women, 47 men) aged 60–99 years (mean 79.5, standard deviation  $\pm 9.11$ ) were analyzed for psychometric testing (exploratory and confirmatory factor analysis,

item fit, McDonald's  $\omega$ , convergent validity, and known-group validity, Rasch).

**Results.** The final LWIG-GER consists of 19 items with three subscales, including "psychological WB", "social WB", and "physical WB". The LWIG-GER showed good overall reliability with McDonald's  $\omega$  of 0.83; the LWIG-GER dimensions' scores were significantly correlated with depression, functional performance, activities, fear of falling, and education.

**Conclusions.** Our findings suggest that the German language version of the LWIG is a reliable and valid tool for measuring WB in nursing home residents. Furthermore, we propose that the LWIG-GER questionnaire can broaden and deepen our understanding of residents' perception of quality of care and their environment.

### Keywords

Well-being · Instrument validation · Nursing home residents · Rasch analysis · Factor analysis

**Purpose of this study.** To our knowledge, the LWIG is a practical and reliable WB assessment tool that has never been validated or implemented in German nursing home populations. Therefore, the purpose of our study was to (1) translate and cross-culturally adapt the LWIG to a German context and (2) test the reliability and validity of the German LWIG in a group of older nursing home residents using the Rasch model.

Following the Standards for Educational and Psychological Testing (Frey, 2018), the present study addressed evidence related to dimensionality, reliability, and construct validity (Boateng, Neilands, Frongillo, Melgar-Quinonez, & Young, 2018). Dimensionality, in general, refers to the structure of a specific phenomenon and examines the extent to which the internal components of a scale match the defined constructs and is concerned with item homogeneity. We examined dimensionality using both ex-

ploratory and confirmatory factor analytical procedures as well as Rasch analysis. Reliability refers to the consistency and relative freedom from error of an instrument. This study assessed internal consistency using the reliability coefficients Cronbach's alpha ( $\alpha$ ), composite reliability (CR), and McDonald's omega ( $\omega$ ) coefficient. Finally, construct validity refers to how well a scale measures the construct it is intended to measure and is based on, among other things, the relationships of the constructs to other variables. There are two subsets of construct validity: convergent construct validity and discriminant construct validity. Convergent construct validity tests the relationship between the construct and a similar measure; it shows that constructs that are supposed to be related are related. Discriminant construct validity tests the relationships between the construct and an unrelated measure; this indicates that constructs are not associated with some-

thing unexpected. Consistent with previous studies, factors such as physical and cognitive activities as well as physical function are found to be positively related to well-being in nursing home residents (Brett, Traynor, & Stapley, 2016; Grönstedt et al., 2011; Saadeh, Welmer, Dekhtyar, Fratiglioni, & Calderón-Larrañaga, 2020) and negatively related to impaired cognition, fear of falling, and depression (Morsch, Shenk, & Bos, 2015; Smalbrugge et al., 2006). Therefore, these constructs were selected to assess convergent construct validity using correlational analyses.

## Methods

### Study design

This cross-sectional study was designed to examine the psychometric properties (dimension, reliability, validity) of the German version of the LWIG (LWIG-GER) in nursing home residents.

### Transcultural translation and adaptation of the instrument

The translation was performed according to translation guidelines of the World Health Organization (WHO, 2010). Four different phases were followed: (1) the initial translation from English to German by two independent bilingual translators; (2) the synthesis of the first two translations to provide the first version of the translated questionnaire; (3) the backward translation by two different bilingual blinded to the original English version and having German as their first language; (4) an expert committee review to compare the backward translations with the original questionnaire and consent on a second version of the translated questionnaire; (v) the pretest of the second version of the LWIG-GER to ensure good comprehension of each question of the questionnaire and conclude with the third version, the final version of the German LWIG.

Two geriatric nursing experts were convened for the consensus process, which identified and addressed discrepancies between different language versions to achieve conceptual, idiomatic,

and semantic equivalences until they agreed on the preliminary German LWIG version. Last, a pretest of 16 older adults within the selected nursing homes was performed using face-to-face cognitive interviews to evaluate the preliminary version's comprehension and clarity. No changes were made to the applied version because they were judged as straightforward and easy to understand. The questionnaire has three different answer scales with four answering options: (1) "not, sometimes, often, always", (2) "not, seldom, sometimes, often", (3) "completely disagree, mostly disagree, mostly agree, completely agree". The total scale range of the original scale with 30 items is 30–120, with a higher score indicating a higher level of well-being. Eventually, the final version was obtained, which is reported in the Appendix.

### Participants

Between March and September 2019, a convenience sample of 104 long-term nursing home residents was recruited from five different nursing homes in Baden-Württemberg, Germany. Inclusion criteria for the study were (i) the consent to participate, (ii)  $\geq 60$  years, (iii) the ability to comprehend and carry out simple instructions, and (iv) no communication deficits (due to chronic conditions such as sensory loss, dementia, and stroke or inability to speak German). Subjects were recruited in cooperation with the nursing staff based on the mentioned criteria. In the first stage, participants completed the LWIG-GER and the Montreal Cognitive Assessment (MoCA); one week after completing the questionnaire and test, all respondents were approached with an invitation to participate in further tests. As a result, 64 respondents agreed to participate in this second stage. There were no significant differences in age, gender, and MoCA between participants and nonparticipants in stage 2. All nursing home residents or their legal caregivers provided written informed consent before participating in the study. The experimental procedure was explained in detail to participants. The study was

carried out according to the institute's ethical standards and the 1964 Helsinki declaration and its later amendments (WMA, 2015).

## Instruments

### Timed-up-and-go test and walking speed

The TUG (timed-up-and-go test; Podsiadlo & Richardson, 1991) is a widely used assessment to examine balance, gait speed, and a predictor for the risk of falling (Allison, Painter, Emory, Whitehurst, & Raby, 2013; Herman, Giladi, & Hausdorff, 2011; Shumway-Cook, Brauer, & Woollacott, 2000). The TUG measures the time it takes a person to stand up from a chair, walk 3 m, turn around, walk back, and sit down again. Subjects independently selected the walking speed that was appropriate for them. If individuals require less than 10 s, they are freely mobile. Times between 10–20 s can be considered as independent mobile. If the task is completed in 20–29 s, the individual has variable mobility, and if it takes more than 29 s, the individual is mobility impaired (Podsiadlo & Richardson, 1991).

In addition to the TUG, walking speed was measured over a distance of 5 m (Cromwell & Newton, 2004). The number of steps was counted, and the time was measured in seconds. Each subject performed two trials with the task of completing the distance as fast as possible. The better value was included in the analysis. Walking speed is used to assess lower extremity mobility and strength. It is considered a reliable measure of physical functioning in older adults and performs the so-called ADLs (Activities of Daily Living) (Buchner, Larson, Wagner, Koepsell, & de Lateur, 1996).

### Depression in Aging Scale

The DIA-S (Depression in Aging Scale; Heidenblut & Zank, 2010) is a test designed as a self-assessment procedure for screening for depressive disorders. The scale consists of ten short statements about depression to be evaluated as true or false with a simple yes/no answer format. The items were constructed to be brief, easy to use, and easy to interpret.

Care was taken to select items to ensure that they were context-free so that the instrument could be used in various health settings (Heidenblut & Zank, 2014). The scale is based on the current WHO definition of depressive disorders and addresses everyday clinical practice demands in geriatric institutions. In addition, the implementation and evaluation should require little time and not overburden both the examiner and the respondent. Regarding the interpretation, a total score of 0–3 points is considered normal, whereas 4 points and more indicate depressive disorders (Heidenblut & Zank, 2010).

### Montreal Cognitive Assessment

Global cognitive status was assessed using the MoCA (Montreal Cognitive Assessment; Nasreddine, Phillips, Bédirian, & Charbonneau, 2005). It is a brief screening tool to detect mild cognitive impairment. The test consists of 30 items categorized into the following cognitive domains: executive functions, visuospatial abilities, language; short-term memory; attention, concentration, working memory, and temporal and spatial orientation. MoCA scores range between 0 and 30, with a score of 26 or above considered no cognitive impairment, a score of 14 or less indicating cognitive impairment (Nasreddine et al., 2005).

### Demographic data, activities, and fear of falling

In addition to sociodemographic data (age, gender, height, weight, and education), weekly activities (e.g., church service, game sessions, gymnastics, singing and making music, and memory training) were collected. Additionally, the subjects were asked if they were afraid of falling in general (yes/no) to determine their fear of falling (Cameron et al., 2000).

### Statistical analysis

Descriptive statistics, correlation, and independent t-tests were performed using IBM SPSS Statistics ver. 27.0 (IBM Corp., Armonk, NY, USA). Exploratory factor analysis and calculating omega were performed using JASP (v0.14.1); AMOS 27.0 was used for the confirma-

tory factor analysis. All Rasch analyses were performed using Winstep ver. 4.8.0 (Winsteps, Beaverton, OR, USA).

There are several approaches for assessing a scale's psychometric characteristics and reducing items: classical test theory (CTT) and modern test theory (item response theory, Rasch Measurement; Ellis & Mead, 2002). For the CTT evaluation of the 30-items LWIG-GER, we examined the data quality and scaling evaluation, the scaling assumptions, and the reliability. For the modern test theory, we used the Rasch Measurement approach. The deletion of the items is discussed in terms of both Rasch Measurement properties and impact on the content of the final instrument, taking into account the importance of the items. The most meaningful and psychometrically sound solution was retained to create the final version of the LWIG-GER questionnaire.

**Classical test theory (CTT).** The 30 items of the original LWIG were subject to item analysis using standard statistical procedures. The distributional properties of each item were examined by inspecting the skewness and kurtosis of the item's distribution and the pattern of response frequency. Excess kurtosis (SPSS-specific) greater than 4 indicates substantial deviance from the normal distribution; an absolute skew value > 2 represents a substantial departure from normality (West, Finch, & Curran, 1995).

The initially hypothesized 6-factor did not fit the data, and thus we followed the initial confirmatory factor analysis (CFA) with exploratory factor analysis (EFA). The EFA was conducted with JASP, using weighted least-squares estimation as the factor extraction method and geomin rotation as the factor rotation method to examine the construct validity of the LWIG-GER. Only factor loadings above 0.40 were considered indicative of item loading and were considered for CFA. The CFA was used to confirm the factor structure determined by EFA. The mean- and variance-adjusted weighted least squares (WLSMV) estimator was employed for CFA, as it is superior to other estimation methods for Likert-type rating scales (Li, 2016).

Model fit indices included the  $X^2$  test, the comparative fit index (CFI) and the Tucker–Lewis index (TLI)  $\geq 0.90$ , and the root mean square error of approximation (RMSEA)  $\leq 0.06$  (Hu & Bentler, 1999). Finally, following the basics assumptions of CTT (Ellis & Mead, 2002), a summary score of the reduced questionnaire was obtained by summing and averaging the scores of their component dimensions.

*Internal consistency* was evaluated to determine the reliability of the LWIG-GER. Internal consistency, which indicates the degree to which all items in the instrument refer to the same construct, can be assessed by several coefficients, each with its strengths and limitations (Deng & Chan, 2017). Cronbach's  $\alpha$ , composite reliability (CR), and McDonald's  $\omega$  coefficient were calculated for this study.

*Convergent and divergent validity* was explored via correlations with other scales. The LWIG-GER scores (total with 19 items [after EFA, CFA] and subdimensions) were correlated with age, body mass index (BMI), the DIA, the MoCA score, 5 m gait speed, and the TUG. It was hypothesized that the LWIG-GER would present higher correlations with motor performance. Lower correlations were expected with the DIA and the MoCA. Known-group validity was studied by examining the LWIG-GER score values against DIA and gait speed after transforming them into categorical variables.

**Rasch measurement.** To provide additional information on the psychometric properties of the LWIG-GER items, Rasch analysis (Boone, Staver, & Yale, 2014) was conducted on the original item set of 30 questions, INFIT and OUTFIT statistics, reliability, and separation indices were calculated. First, validity was represented through fit statistics, reported as log odd units (logits) in an unstandardized (mean square [MnSq]). We have applied an iterative procedure to identify the persons causing an item to misfit and start off with the worst person and give them a PWEIGHT of 0. We repeated this process until the item fit. Next, person and item reliability and separation were calculated. The

**Table 1** Sample characteristics

	<b>N</b>	<b>M</b>	<b>SD</b>	<b>Min</b>	<b>Max</b>
Age (years)	104	79.47	9.11	60.0	99.9
Sex	104	57 (54.8%) female; 47 (45.2%) male			
Weight (kg)	55	74.4	14.0	37.0	110
Height (cm)	62	167	10.8	110	187
BMI (kg/m <sup>2</sup> )	55	26.8	4.98	16.6	41.5
Medication (n)	63	0.98	0.13	0	1
MoCA (score)	104	19.03	5.24	6.00	30.0
Education (years)	64	9.28	2.74	7.00	17.0
DIA-S (score)	63	1.89	2.36	0.00	9.00
Activities (n)	64	2.23	1.63	0	6
TUG (s)	63	18.3	12.4	4.10	66.0
5 m Walking duration (s)	63	9.67	5.65	2.19	32.6
Fear of falls (% yes)	64	24 (37.5%)			

*BMI* Body-Mass-Index, *MoCA* Montreal Cognitive Assessment, *DIA-S* Depression in Aging Scale, *TUG* Timed-up-and-go Test

item–person map, often called a Wright map, displays both persons (in terms of their well-being) and items (in terms of their difficulty to agree with) along a common vertical axis (▣ Fig. 3). Thus, it can be shown whether there are redundant items, too many items of the same difficulty, and gaps between items.

## Results

### Participants

A total of 104 long-term nursing home residents (57 women, 47 men) aged 60–99 years ( $M = 79.5 \pm 9.1$ , participant characteristics shown in ▣ Table 1) completed the translated LWIG-GER. While the mean age is comparable to the population living in nursing homes in Germany, as presented in epidemiological studies, the percentage of men in our sample is higher than in similar studies (von Renteln-Kruse & Ebert, 2003). The MoCA total score across all residents was 19.03 ( $\pm 5.24$ ) points, which is just above the cut-off value (19 points) for differentiating mild cognitive impairment and Alzheimer's disease (Roalf et al., 2013). In addition to the sample's sociodemographic information, ▣ Table 1 shows the descriptive results on mobility, depression, the number of regular activities, fear of falling, and cognitive function. However, these data were collected from only a subsample ( $n = 64$ ).

### Psychometric properties of the LWIG-GER

#### Item analysis

Nonresponse, item difficulty, skewness, and kurtosis were computed to give descriptive information and distribution properties (Supplementary material: Table S2). Items 5, 22, and 23 showed unacceptable values for skewness and kurtosis and a low item difficulty. Therefore, these three variables are excluded from the factor analyses (EFA, CFA).

#### Exploratory factor analysis

The Kaiser–Meyer–Olkin (KMO) test and Bartlett's test of sphericity were conducted to assess the dataset's suitability for factor analysis. In this study, the KMO value was 0.76, and Bartlett's test of sphericity was significant ( $\chi^2 = 1130$ ,  $p < 0.001$ ), which means that the sample was appropriate for factor analysis. The number of factors was explored based on parallel analysis, a scree plot, and Eigenvalues. A three-factor model was determined as the final factor structure of the LWIG-GER (Eigenvalues 3.30, 3.34, 2.01). In the EFA, all items except nine (Items 2, 3, 8, 16, 17, 18, 21, 24, 29) were loaded ( $> 0.40$ ) in at least one factor (▣ Table 1). Two items (16, 17) loaded on two factors and did not maximize McDonald's  $\omega$ ; item 29 led to an improved McDonald's  $\omega$ . The short LWIG-GER total scale comprises 19 items. The three-factor solution 'Psychological well-

being' (7 items; affect and self-worth), 'Social well-being' (8 items; positive and negative social experience, communal living), and 'Physical well-being' (4 items; absence of physiological needs, feeling fit, being cared for physically) accounted for 39.4% of the variance. McDonald's  $\omega$  were 0.827, 0.795, and 0.685, respectively.

#### Rasch analysis

Participants who had unexpected values for a given item were weighted with 0 using the PWEIGHT function to get the fit statistics (INFIT and OUTFIT MnSq) within the intended range (0.5–1.5). Thus, weights of 0 were applied to four individuals. This brought the MnSq into the intended range, and all items fit the Rasch model (▣ Table 2). The acceptable MnSq value of an item ranges from +0.5 to +1.5 considered a high yield for measurement (Boone et al., 2014). All items fall within the designated range. The mean MnSq INFIT value for all items is +1.04 (P.S.D = 0.21). The mean MnSq OUTFIT value is +1.01 (P.S.D = 0.22).

Item measures show that item 12 (1.20 logits) is the most difficult item to agree with, and items 5, 22, and 23 are most easily agreed with  $-1.03$  to  $-1.97$  logits (▣ Table 2; ▣ Fig. 2; the items are shown in red). The combined analysis of the 30 items showed a good separation index for both the participants studied (2.27; reliability 0.84) and the 30 items (3.89; reliability 0.94).

**Table 2** Means, internal consistency reliability, factor loadings, and item fit statistics (N=104)

Item content	Mean (SD)	Item total correlation	McDonald's $\omega$ if dropped	Factor loadings			INFIT MnSq	OUTFIT MnSq	Measure Logits
				F1	F2	F3			
<i>Factor 1 Psychological well-being (0–21), Mean (SD) 15.9 (4.23); Cronbach's <math>\alpha = 0.835</math>, <math>\omega = 0.827</math></i>									
2. How often did you feel relaxed?	1.93 (0.98)	–	–	0.175	0.254	0.272	1.01	1.01	0.70
8. How often did you feel anxious or tense?	2.64 (0.62)	–	–	0.196	0.138	0.285	0.97	0.90	–0.67
9. I am satisfied with how my life has turned out so far	2.27 (0.81)	0.512	0.815	<b>0.600</b>	0.070	–0.006	0.94	0.89	0.14
14. How often did you feel sad or depressed?	2.18 (0.76)	0.544	0.812	<b>0.498</b>	–0.099	0.279	0.75	0.91	0.28
15. I think life is meaningful	2.36 (0.82)	0.739	0.777	<b>0.843</b>	0.023	–0.055	0.73	0.68	–0.04
18. How often did you enjoy the communal mealtimes?	2.25 (0.98)	–	–	<b>0.373</b>	0.137	–0.010	1.26	1.36	0.14
20. How often did you feel empty or flat?	2.33 (0.75)	0.583	0.806	<b>0.592</b>	–0.033	0.251	0.66	0.67	–0.03
21. I have accomplished what I wanted to in life	2.19 (0.85)	–	–	<b>0.339</b>	0.198	–0.012	0.90	0.95	0.28
25. How often did you feel bored?	2.44 (0.82)	0.608	0.797	<b>0.667</b>	–0.007	0.002	0.94	0.88	–0.18
28. How often did you feel lonely?	2.38 (0.80)	0.625	0.798	<b>0.555</b>	0.114	0.223	0.74	0.62	–0.04
30. How often did you worry about the purpose of your life?	1.98 (1.05)	0.546	0.821	<b>0.560</b>	0.034	0.124	1.13	1.24	0.61
<i>Factor 2 Social well-being (0–24), Mean (SD) 17.5 (4.48); Cronbach's <math>\alpha = 0.795</math>, <math>\omega = 0.795</math></i>									
3. I think I am worth the effort	2.42 (0.75)	–	–	0.280	<b>0.337</b>	–0.076	0.90	0.86	–0.18
4. How often did you receive sufficient respect?	2.33 (0.89)	0.598	0.748	–0.014	<b>0.645</b>	0.236	0.89	0.80	0.03
6. How often did you experience a sociable atmosphere when with the other residents?	1.79 (0.99)	0.522	0.736	0.019	<b>0.622</b>	–0.093	1.01	1.01	0.91
10. How often did you receive sufficient attention?	2.30 (0.81)	0.600	0.747	0.072	<b>0.608</b>	0.078	0.79	0.81	0.10
11. How often did you feel ignored?	2.65 (0.65)	0.353	0.779	–0.205	<b>0.464</b>	0.310	1.16	1.04	–0.70
12. How often did you feel you fit in with the other residents?	1.57 (1.06)	0.600	0.727	0.012	<b>0.667</b>	–0.051	0.96	0.94	1.20
16. There are fellow residents with whom I have good contact	2.25 (1.01)	–	–	0.328	<b>0.385</b>	–0.197	1.24	1.29	0.16
17. How often did you feel that others saw you as a burden?	2.61 (0.73)	–	–	–0.010	<b>0.376</b>	0.323	1.20	1.12	–0.61
19. How often were you cared for in a pleasant way?	2.38 (0.91)	0.437	0.765	–0.029	<b>0.495</b>	0.033	1.30	1.40	–0.02
26. I see the people that are important to me as often as I would like to	2.08 (0.98)	0.428	0.779	0.073	<b>0.450</b>	0.066	1.05	0.96	0.46
29. There are people with whom I can feel completely at ease	2.51 (1.04)	0.383	0.778	0.160	<b>0.380</b>	–0.128	1.31	1.14	–0.34

**Table 2** (Continued)

Item content	Mean (SD)	Item total correlation	McDonald's $\omega$ if dropped	Factor loadings			INFIT MnSq	OUTFIT MnSq	Measure Logits
				F1	F2	F3			
<i>Factor 3 Physical well-being (0–12), Mean (SD) 8.11 (2.77); Cronbach's <math>\alpha = 0.693</math>, <math>\omega = 0.685</math></i>									
1. How often did you feel fit?	1.87 (1.00)	0.457	0.643	0.078	0.107	<b>0.498</b>	1.07	1.09	0.74
7. How often did you have severe pain?	2.14 (1.03)	0.505	0.606	–0.084	0.028	<b>0.575</b>	1.50	1.40	0.31
13. How often did you not feel well?	2.12 (0.82)	0.434	0.645	0.166	–0.003	<b>0.462</b>	0.82	1.43	0.31
24. How often did you enjoy your food and drinks?	2.27 (0.89)	–	–	0.266	0.170	0.163	1.04	0.36	0.16
27. How often did you suffer from physical limitations or difficulties in daily life?	1.98 (0.97)	0.551	0.567	0.306	–0.064	<b>0.535</b>	0.94	–0.46	0.57
<i>Mean</i>							1.04	0.1	0.00
<i>PSD</i>							0.21	1.3	0.65
<i>TOTAL LWIG_GER (0–57), Mean (SD) 41.6 (8.68); (Cronbach's <math>\alpha = 0.841</math>, <math>\omega = 0.831</math>)</i>									

*PSD* Population standard deviation (standard deviation assuming that this sample of the statistic is the entire population)

*Mean* the average value of the statistic

*MnSq* mean-square statistic with expectation 1 (The acceptable MnSq value of an item ranges from +0.5 to +1.5, which is considered to be productive for measurement, values substantially below 1 indicate dependency in your data; values substantially above 1 indicate noise or the presence of unexpected outliers)

*Measure* Item measures expressed in logit units (A logit is defined as a natural log of an odds ratio (Ludlow & Haley, 1995). A higher item measure indicates an item that was harder to agree with, whereas a lower item measure denotes an item that was easier to agree with)

Rasch person reliability and Rasch item reliability values range from 0–1 and can be interpreted similarly to Cronbach's  $\alpha$ . Separation indices can range from 0 to infinity, with higher values indicating better separation. The separation statistics for the individuals indicated that two different groups with different ability levels could be reliably distinguished based on the scale (in terms of person separation, an index of 1.50 is acceptable, 2.00 is good, and 3.00 is excellent (Duncan, Bode, Lai, & Perera, 2003)). In addition, items separation statistics show that three item groups that differ in item difficulty can be distinguished from each other (Linacre (2012) suggests that item separation indices of 3 or greater are desirable).

Figure 3 shows how the individuals with their different ability expressions and the difficulty of the 30 items can be represented in a common metric. Here, one can see some psychometrically redundant items and assess the same level of difficulty on the construct (e.g., WB10\_SW, WB16\_SW, WB18\_SW, WB24\_PHW, WB9\_PSW or WB15\_PSW, WB19\_PHW, WB20R\_PS, WB28R\_PS, WB4\_SW\_P). Moreover,

gaps between items can be observed, for example, between WB12\_SW and WB6\_SW\_6, especially between WB22\_SW and WB23R\_SW. If, as with the former gap, individuals fall into this gap, a researcher would not be able to differentiate these individuals.

### Confirmatory factor analysis

Subsequently, the three-factor structure obtained with the EFA was verified with the CFA, excluding redundant items from the Rasch analysis or items with low factor loadings. As modification indices suggested that model fit would improve provided correlated error terms were included, we added one theoretically plausible correlated error term (item 6: How often did you experience a sociable atmosphere when with the other residents? And item 12: How often did you feel you fit in with the other residents?). Results indicated that the three-factor model was appropriate with an acceptable fit ( $X^2(147) = 190$ ; CFI = 0.937, TLI = 0.919, and RMSEA = 0.045).

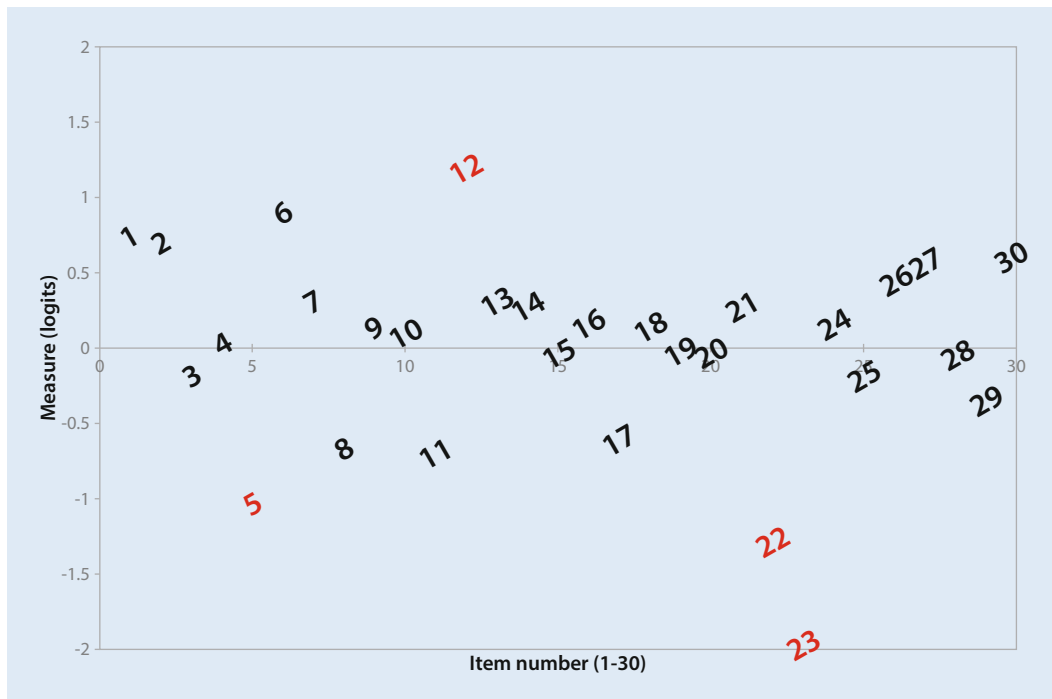
### Convergent, divergent, and known group validity

Convergent, divergent, and known group validity data are presented in Table 3. A higher score on the DIA scale was associated with lower psychological and physical well-being. Better functional status in mobility and balance was associated with significantly better social and physical well-being scores. Higher education was also reflected in better social and physical well-being. No significant associations were present for the LWIG-GER and age, sex, BMI, number of medications, and MoCA. As far as the known-group validity was concerned, higher gait speed and lower DIA scores were significantly associated with higher well-being scores in almost all subdimensions and the total score (based on 19 items).

### Discussion

This study translated the LWIG—a questionnaire specially designed for nursing home residents reflecting their environment—into German (i.e., LWIG-GER) and examined its psychometric properties based on both factor analysis (EFA, CFA) and Rasch modeling. The CFA re-





**Fig. 2** ◀ Item measures of the LWIG-GER (If the measure is positive, the ability is higher than difficulty, and the probability of solution is higher than 50%. The most difficult and the easiest items to agree with are shown in red)

sults supported the validity of the three-factor model of the original LWIG (psychological well-being, social well-being, physical well-being) but not the 6-factor model. Based on classic test theory, the analysis produced a shorter version of 19 items of the LWIG-GER with acceptable reliability, validity, and moderate to strong correlations with other depression and functional performance measures. Rasch modeling suggested that the LWIG-GER was unidimensional, with acceptable INFIT and OUTFIT values, good person separation, and high person reliability. It also generated a reduced version of the LWIG-GER, excluding similar items as the classic test theory procedure.

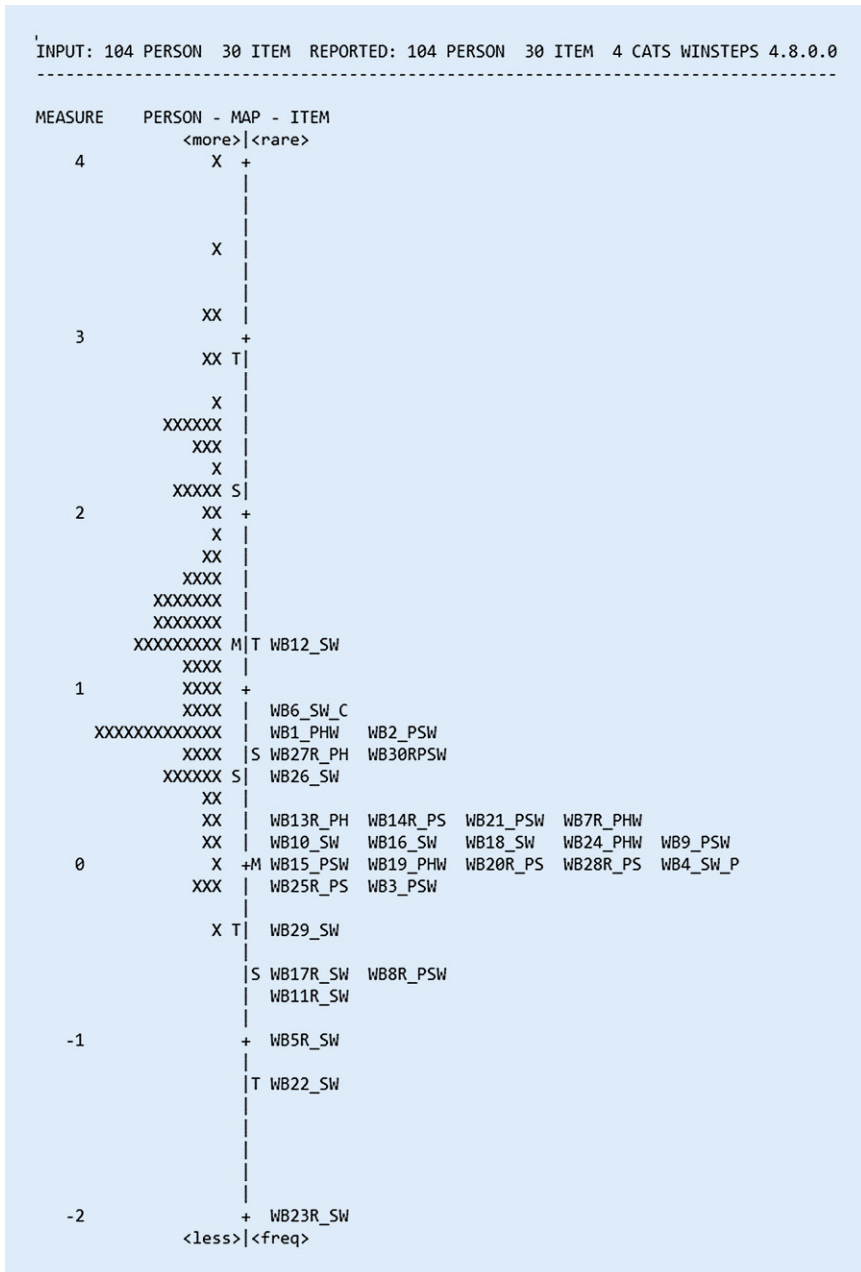
In this convenience sample from different nursing homes, the answers of almost all items of the LWIG are somewhat on the side of high well-being, leading to left-skewed distributions of the scores. This should not necessarily be regarded as unfavorable; the values of other instruments measuring QoL or WB show a similar pattern (Lindert et al., 2015). However, classical test theory as well as item response theory analysis (using Rasch) clearly show that items 5, 22, and 23 have unacceptable values for skewness and kurtosis (compared to factor analysis, a normal distribution is not

a requirement for Rasch analysis) and a low item difficulty (item analysis: 0.92, 0.93, 0.94; Rasch measure: -1.03, -1.28, -1.97). Therefore, we omitted these three items, which addressed the questions “be bothered by other residents”, “relationship with nurses”, and “be bullied”.

Factor analysis demonstrated the appropriateness of a three-factor structure for the LWIG. The structure for all 30 items was identified for the first time because factor analysis on all items was not conducted on all items simultaneously in the original LWIG development study in the Netherlands (van der Wolf et al., 2018). Subscales were labeled to reflect the theory guiding the original LWIG item pool. In the EFA, the factor loadings of nine items on the scale are below 0.40 or loaded significantly on more than one dimension. For example, according to van der Wolf et al. (2018), item 3 (“I think I am worth the effort”) belongs to the psychological dimension, but our results show that this item loads with a factor of 0.337 on the social WB dimension. In contrast, for item 18 (“How often did you enjoy the communal mealtimes?”; on the social dimension), we see a factor loading of 0.373 on the psychological dimension. The different assignment of the items to

a specific dimension can be attributed to a different interpretation of the questions. Undeniably, short questionnaires improve assessment by saving response time and effort, increasing the response rate, minimizing burden, and decreasing fatigue. Therefore, based on the previous results for subsampling, items 2, 3, 8, 16, 17, 18, 21, and 24 were excluded. Concerning the results based on CFA, we found the LWIG-GER with 19 items had good performance on all psychometric indicators.

Item fit analysis for each item was performed to ensure that all items matched expected responses based on the Rasch model. INFIT and OUTFIT MnSq statistics for each item were within the acceptable range, indicating a satisfactory fit to the underlying global trait, i.e., the residents’ well-being. Nevertheless, our results also indicate that the LWIG-GER had some redundant or poorly constructed items. In this context, six items can be considered redundant (e.g., 8, 16, 28, 24). The Wright map shows that most of the items that are too easy to agree with can be assigned to the social WB dimension (e.g., items 5, 22, and 23), which corresponds with the item analysis (skewness and kurtosis of the item’s



**Fig. 3** ▲ Wright item–person map of the LWIG-GER (Items that are harder to agree with are plotted toward the top of the Wright map. Items that are easier to agree with are plotted toward the base of the Wright map)

distribution and the pattern of response frequency).

A widely used way to measure a questionnaire’s reliability is to assess internal consistency. We preferred the use of McDonald’s  $\omega$  rather than the widely used Cronbach’s  $\alpha$  since Cronbach’s  $\alpha$  is based on the assumption of equal factor loadings (however, for comparison, we also reported Cronbach’s  $\alpha$ ). Therefore, Cronbach’s  $\alpha$ , McDonald’s  $\omega$  values

>0.7, can be interpreted as good internal reliability. For the LWIG-GER, we obtained McDonald’s  $\omega$  of 0.685, 0.795, and 0.827 indicating a moderate to good internal consistency for the three dimensions of the LWIG-GER. Both reliability measures were comparable to the Dutch instrument (van der Wolf et al., 2018).

In terms of convergent, divergent, and known group validity, the LWIG-GER was significantly correlated with educa-

tion, depression, functional performance (TUG, gait speed), and fear of falling (measures as a one-item question). Like van der Wolf et al. (2018, 2019), we observe a negative correlation between the LWIG-GER total score and the depression scale (DIA-S). Our results are also consistent with those of Meeks and Murrell (2001), who showed that higher education levels are related to lower levels of negative affect in older adults. Also, Subaşı and Hayran (2005) demonstrated that educational level is a statistically significant independent predictor of life satisfaction among nursing home residents. We also see an association between fear of falling and the LWIG-GER total score and physical WB and an association between mobility (TUG and 5 m gait speed) and the LWIG-GER total score and in almost all of its dimensions. This study conducted a known-group comparison to examine whether the LWIG-GER and its subdimensions scores could discriminate between three groups according to depression and gait speed. Our results confirm this relationship with respect to DIA-S ( $p < 0.001$  to  $p = 0.003$ ) and 5 m gait speed ( $p < 0.001$  to  $p = 0.121$ ). In this context, the study by Painter et al. (2012) aimed to examine the relationship of fear of falling, depression, anxiety, activity level, and activity limitation. They were able to show strong correlations for most of these constructs. This illustrates that these are multifaceted constructs that have a reciprocal relationship along with the construct of WB. Through these analyses, convergent validity and known group validity were confirmed to be satisfactory.

The present study has several limitations that should be noted. One issue is achieving a sufficient sample size to conduct an EFA or CFA since getting a sample in a clinical setting is not easy. Although our sample size was smaller than that recommended for traditional factor analysis (minimum subject to item ratio of at least 5:1, Hatcher, 1994; or more recently 3:1, Bujang, Ghani, Soelar, & Zulkifli, 2012), the Rasch rating scale model functions with a minimum of 10 observations per category (Boone et al., 2014). As a Likert-type scale of four categories was used for the present study, the minimum

**Table 3** Convergent, divergent, and known-group validity of the LWIG-GER scale with DIA and gait speed

Spearman correlations	LWIG-GER total (19 items)	Psychological well-being (7 items)	Social well-being (8 items)	Physical well-being (4 items)
Age	0.127	0.101	0.097	0.127
Sex	-0.144	-0.108	-0.092	-0.144
BMI	-0.121	-0.023	-0.237	-0.121
Education	-0.249*	0.177	0.356**	0.249*
Medication	-0.197	0.021	0.049	-0.197
MoCA	0.031	0.025	-0.133	0.031
Activities	0.151	0.018	0.309*	0.151
DIA-S	-0.505**	-0.695**	-0.245	-0.505**
5 m gait speed	-0.629**	-0.257*	-0.362**	-0.629**
TUG	-0.466**	-0.180	-0.306*	-0.466
Fear of falling	-0.447**	-0.067	0.075	-0.477**
<b>Known groups</b>				
<i>DIA-S</i>				
Inconspicuous atmosphere	44.8 (7.27)	18.0 (2.62)	18.0 (5.10)	8.74 (2.88)
Suspected depression	38.4 (6.23)	14.0 (2.12)	18.2 (3.03)	6.20 (2.17)
Depression of disease value is likely	28.9 (5.48)	10.8 (4.88)	12.6 (3.15)	5.58 (2.35)
<i>p</i> -value	< 0.001	0.003	0.001	< 0.001
<i>Gait speed</i>				
≤ 0.6 m/s	38.0 (9.07)	15.5 (4.97)	16.4 (6.40)	6.23 (2.95)
0.61–1.0 m/s	42.0 (8.59)	16.5 (3.14)	15.9 (5.17)	9.68 (1.84)
≥ 1.0 m/s	49.0 (5.64)	18.6 (3.20)	21.1 (4.38)	9.30 (2.50)
<i>p</i> -value	0.003	0.121	0.014	< 0.001

MoCA Montreal Cognitive Assessment, *DIA-S* Depression in Aging Scale, *TUG* Timed-up-and-go Test  
 \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$

sample size for Rasch model measurement analysis for the measurement instrument in this study was  $n = 40$ . These recommendations suggest a valid sample size of  $n = 104$  in the present study to analyze the LWIG-GER. However, more significant numbers of cases are needed to replicate our findings and draw reliable conclusions.

We are confronted with the same attendance problem mentioned by van der Wolf et al., (2018). Accordingly, it can be assumed that residents who did not attend are either more physically (e.g., gait speed, TUG) or mentally (e.g., MoCa) impaired, which should negatively affect WB. In particular, a more comprehensive range of facilities should be considered in terms of socioeconomic status.

Regarding reliability, an internally consistent questionnaire can be assumed based on correlations, Cronbach's  $\alpha$ , McDonald's  $\omega$ , and calculated Rasch

reliability. However, no conclusions can be made about test-retest reliability and interrater reliability. In future studies, testing these reliability measures is an exciting aspect. Moreover, even though we did not include items 5, 22, and 23 in the factor analysis, and most of the items that are too easy to agree with belong to the social dimension, we would not recommend removing all of these items. This is because the present data were collected in only five different facilities, and thus it is not certain that the full range of facilities was covered. Therefore, it is important to replicate our results not only with larger samples in future studies to revisit the questionnaire's undoubtedly reasonable shortening; but in addition, other formulations regarding items that are too "easy" should be tested with the residents to avoid falling into the trap of social desirability in response behavior.

## Conclusion

Taken together, the German version of the short LWIG-GER with 19 items provides a valid and reliable measurement tool to evaluate well-being in nursing home residents. Although there is no consistent definition of WB, it seems to be a multidimensional construct (Lindert et al., 2015; Su et al., 2014). Based on the CFA and comparable to the results of van der Wolf et al. (2018), we can verify a three-factor structure (psychological, social, and physical well-being) with an acceptable fit.

We see great potential in this questionnaire because, on the one hand, there have been few specific questionnaires on WB in healthy nursing home residents. On the other hand, the questionnaire is meaningful, cost-effective, time-efficient, and easy to use. However, significantly more data are needed to determine the sensitivity, specificity, and cut-off values of the LWIG-GER. Nevertheless, this provides an instrument for future studies that measure WB validly and reliably.

## Corresponding address



**Nadja Schott**  
 Department of Sport and Exercise Science, University of Stuttgart  
 Allmandring 28,  
 70569 Stuttgart, Germany  
 nadja.schott@inspo.uni-stuttgart.de

**Acknowledgements.** The authors thank the volunteers who participated in the study.

**Author Contribution.** We confirm that all authors were fully involved in studying and preparing the manuscript and the material within. In addition, all authors fulfill the ICMJE (International Committee of Medical Journal Editors) recommendations on authorship. Credit taxonomy: N. Schott: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Data Curation, Writing—Original draft, Visualization, Supervision, Project administration; B. Johnen: Methodology, Validation, Formal analysis, Investigation, Writing—Original Draft, Visualization; T. J. Klotzbier: Methodology, Validation, Formal analysis, Writing—Original Draft, Visualization

**Availability of data and material.** Data can be obtained from the corresponding author upon reasonable request.

**Funding.** Open Access funding enabled and organized by Projekt DEAL.

## Declarations

**Conflict of interest.** N. Schott, B. Johnen and T. J. Klotzbier declare that they have no competing interests.

All assessments were conducted in accordance with ethical rules for research in human subjects following the Declaration of Helsinki (WMA, 2013). The local ethics committee approved the study. All nursing home residents or their legal guardians received written and verbal information about the study and signed informed consent before their participation.

**Open Access.** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Allison, L. K., Painter, J. A., Emory, A., Whitehurst, P., & Raby, A. (2013). Participation restriction, not fear of falling, predicts actual balance and mobility abilities in rural community-dwelling older adults. *Journal of Geriatric Physical Therapy*, 36(1), 13–23. <https://doi.org/10.1519/jpt.0b013e3182493d20>.
- Ballmer, T., Wirz, F., & Gantschnig, B. E. (2019). Assessing the quality of life and well-being of older adults with physical and cognitive impairments in a German-speaking setting: A systematic review of validity and utility of assessments / Die Erfassung von Lebensqualität und Wohlbefinden älterer Menschen mit psychischen und kognitiven Einschränkungen: ein systematisches Literaturreview zur Validität und Praktikabilität deutschsprachiger Assessments. *International Journal of Health Professions*, 6(1), 124–143. <https://doi.org/10.2478/ijhp-2019-0014>.
- Bech, P. (1996). Health-related quality of life rating scales. In P. Bech (Ed.), *Rating scales for psychopathology, health status, and quality of life: a compendium on documentation in accordance with the DSM-III-R and WHO systems* (pp. 395–424). New York: Springer.
- Beck, A. T., Steer, R. A., Ball, R., & Ranieri, W. (1996). Comparison of Beck Depression Inventories –IA and –II in psychiatric outpatients. *Journal of Personality Assessment*, 67, 588–597.
- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Frontiers in Public Health*, 6, 149. <https://doi.org/10.3389/fpubh.2018.00149>.
- Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch analysis in the human sciences*. New York: Springer.
- Brett, L., Traynor, V., & Stapley, P. (2016). Effects of physical exercise on health and well-being of individuals living with a dementia in nursing homes: a systematic review. *Journal of the American Medical Directors Association*, 17(2), 104–116. <https://doi.org/10.1016/j.jamda.2015.08.016>.
- Buchner, D. M., Larson, E. B., Wagner, E. H., Koepsell, T. D., & de Lateur, B. J. (1996). Evidence for a non-linear relationship between leg strength and gait speed. *Age and Ageing*, 25(5), 386–391.
- Bujang, M. A., Ghani, P. A., Soelar, S. A., & Zulkifli, N. A. (2012). *Sample size guideline for exploratory factor analysis when using small sample: taking into considerations of different measurement scales*. 2012 International Conference on Statistics in Science, Business and Engineering (ICSSBE). (pp. 1–5). <https://doi.org/10.1109/ICSSBE.2012.6396605>.
- Bullinger, M. (2014). Das Konzept der Lebensqualität in der Medizin – Entwicklung und heutiger Stellenwert. *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen*, 108(2–3), 97–103. <https://doi.org/10.1016/j.zefq.2014.02.006>.
- Bullinger, M., & Kirchberger, I. (1998). *SF-36 Fragebogen zum Gesundheitszustand. Handanweisung*. Göttingen: Hogrefe.
- Cameron, I. E., Stafford, B., Cumming, R. G., Birks, C., Kurlle, S. E., Lockwood, K., et al. (2000). Hip protectors improve falls self-efficacy. *Age and Ageing*, 29, 57–62.
- Cromwell, R. L., & Newton, R. A. (2004). Relationship between balance and gait stability in healthy old adults. *Journal of Aging and Physical Activity*, 11, 90–100.
- Das, K. V., Jones-Harrell, C., Fan, Y., et al. (2020). Understanding subjective well-being: perspectives from psychology and public health. *Public Health Review*, 41, 25. <https://doi.org/10.1186/s40985-020-00142-5>.
- Deng, L., & Chan, W. (2017). Testing the difference between reliability coefficients Alpha and Omega. *Educational and Psychological Measurement*, 77(2), 185–203. <https://doi.org/10.1177/0013164416658325>.
- Diener, E. (2000). Subjective well-being: the science of happiness and a proposal for a national index. *American Psychologist*, 55(1), 34–43. <https://doi.org/10.1037/0003-066x.55.1.34>.
- Diener, E., & Suh, E. (1997). Measuring quality of life: economic, social, and subjective indicators. *Social Indicators Research*, 40(1–2), 189–216. <https://doi.org/10.1023/A:1006859511756>.
- Duncan, P. W., Bode, R. K., Lai, S. M., & Perera, S. (2003). Rasch analysis of a new stroke-specific outcome scale: the Stroke Impact Scale. *Archives of Physical Medicine and Rehabilitation*, 84(7), 950–963.
- Ellis, B. B., & Mead, A. D. (2002). Item analysis: theory and practice using classical and modern test theory. In S. G. Rogelberg (Ed.), *Handbook of research methods in industrial and organizational psychology* (pp. 324–343). Blackwell.
- Ettema, T. P., Dröes, R. M., de Lange, J., Mellenbergh, G. J., & Ribbe, M. W. (2007). QUALIDEM: development and evaluation of a dementia specific quality of life instrument–validation. *International Journal of Geriatric Psychiatry*, 22(5), 424–430. <https://doi.org/10.1002/gps.1692>.
- Frey, B. (2018). *The SAGE encyclopedia of educational research, measurement, and evaluation*. Vol. 1–4. Thousand Oaks: SAGE. <https://doi.org/10.4135/9781506326139>.
- García, L. M. R., & Navarro, R. J. M. (2018). The impact of quality of life on the health of older people from a multidimensional perspective. *Journal of Aging Research*. <https://doi.org/10.1155/2018/4086294>.
- Grönstedt, H., Hellström, K., Bergland, A., Helbostad, J. L., Puggaard, L., Andresen, M., Granbo, R., & Frändin, K. (2011). Functional level, physical activity and wellbeing in nursing home residents in three Nordic countries. *Aging Clinical and Experimental Research*, 23(5–6), 413–420. <https://doi.org/10.1007/BF03337766>.
- GROUP, T. H. E. W. H. O. Q. O. L. (1998). Development of the world health organization WHOQOL-BREF quality of life assessment. *Psychological Medicine*, 28(3), 551–558. <https://doi.org/10.1017/s0033291798006667>.
- Hagell, P., & Nygren, C. (2007). The 39 item Parkinson's disease questionnaire (PDQ-39) revisited: implications for evidence based medicine. *Journal of neurology, neurosurgery, and psychiatry*, 78(11), 1191–1198. <https://doi.org/10.1136/jnnp.2006.111161>.
- Hatcher, L. (1994). *A step-by-step approach to using the SAS system for factor analysis and structural equation modeling*. SAS Institute.
- Heidenblut, S., & Zank, S. (2010). Entwicklung eines neuen Depressionsscreenings für den Einsatz in der Geriatrie. *Zeitschrift für Gerontologie und Geriatrie*, 43(3), 170–176. <https://doi.org/10.1007/s00391-009-0067-z>.
- Heidenblut, S., & Zank, S. (2014). Screening for depression with the depression in old age scale (DIA-S) and the geriatric depression scale (GDS15). *GeroPsych*, 27(1), 41–49. <https://doi.org/10.1024/1662-9647/a000101>.
- Herman, T., Giladi, N., & Hausdorff, J. M. (2011). Properties of the 'timed up and go' test: more than meets the eye. *Gerontology*, 57(3), 203–210. <https://doi.org/10.1159/000314963>.
- Hu, L. T., & 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>.
- Karimi, M., & Brazier, J. (2016). Health, health-related quality of life, and quality of life: what is the difference? *Pharmacoeconomics*, 34(7), 645–649. <https://doi.org/10.1007/s40273-016-0389-9>.
- Koller, M., Neugebauer, E. A. M., Augustin, M., Büssing, A., Farin, E., Klinkhammer-Schalke, M., Lorenz, W., Münch, K., Petersen-Ewert, C., Steinbüchel, N. V., & Wieseler, B. (2009). Die Erfassung von Lebensqualität in der Versorgungsforschung – konzeptuelle, methodische und strukturelle Voraussetzungen. *Das Gesundheitswesen*, 71(12), 864–872. <https://doi.org/10.1055/s-0029-1239516>.
- van Leeuwen, K. M., van Loon, M. S., van Nes, F. A., Bosmans, J. E., de Vet, H. C. W., Ket, J. C. F., Widdershoven, G. A. M., & Ostelo, R. W. J. G. (2019). What does quality of life mean to older adults? A thematic synthesis. *PLOS ONE*, 14(3), e213263. <https://doi.org/10.1371/journal.pone.0213263>.
- Li, C.-H. (2016). Confirmatory factor analysis with ordinal data: comparing robust maximum

- likelihood and diagonally weighted least squares. *Behavioral Research Methods*, 48(3), 936–949.
- Linacre, J.M. (2012). A user guide to winsteps ministep Rasch model computer programs: program manual 3.75.0. <http://www.winsteps.com/winstepsmanual.pdf>. Accessed: 18.02.2021.
- Lindenberg, S. (1986). The paradox of Privatization in consumption. In A. Diekmann & P. Mitter (Eds.), *Paradoxical effects of social behavior: essays in honor of anatol rapoport* (pp. 297–310). Heidelberg: Physica.
- Lindert, J., Bain, P.A., Kubzansky, L.D., & Stein, C. (2015). Well-being measurement and the WHO health policy Health 2010: systematic review of measurement scales. *The European Journal of Public Health*, 25(4), 731–740. <https://doi.org/10.1093/eurpub/cku193>.
- Ludlow, L.H., & Haley, S.M. (1995). Rasch model logits: Interpretation, use, and transformation. *Educational and Psychological Measurement*, 55(6), 967–975.
- Meeks, S., & Murrell, S.A. (2001). Contribution of education to health and life satisfaction in older adults mediated by negative affect. *Journal of Aging and Health*, 13(1), 92–119.
- Megari, K. (2013). Quality of life in chronic disease patients. *Health psychology research*, 1(3), e27. <https://doi.org/10.4081/hpr.2013.e27>.
- Meyer, R., Drewniak, D., Hovorka, T., & Schenk, L. (2019). Questioning the questionnaire: methodological challenges in measuring subjective quality of life in nursing homes using cognitive interviewing techniques. *Qualitative Health Research*, 29(7), 972–986.
- Morsch, P., Shenk, D., & Bos, A.J. (2015). The relationship between falls and psychological well-being in a Brazilian community sample. *Journal of Cross-Cultural Gerontology*, 30(1), 119–127. <https://doi.org/10.1007/s10823-014-9249-2>.
- Nasreddine, Z.S., Phillips, N.A., Bedirian, V., & Charbonneau, S. (2005). The Montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4), 695–699.
- Oppikofer, S., & Mayorova, E. (2016). Lebensqualität im hohen Alter – theoretische Ansätze, Messmethoden und empirische Befunde. *Pflege und Gesellschaft*, 21(2), 101–113.
- Painter, J.A., Allison, L., Dhingra, P., Daughtery, J., Cogdill, K., & Trujillo, L.G. (2012). Fear of falling and its relationship with anxiety, depression, and activity engagement among community-dwelling older adults. *American Journal of Occupational Therapy*, 66, 169–176. <https://doi.org/10.5014/ajot.2012.002535>.
- Podsiadlo, D., & Richardson, S. (1991). The timed "Up & Go": a test of basic functional mobility for frail elderly persons. *Journal of the American Geriatrics Society*, 39(2), 142–148.
- von Renteln-Kruse, W., & Ebert, D. (2003). Merkmale hospitalisierter geriatrischer Patienten – Zwei Kohorten im Vergleich unter Verwendung des Screenings der Arbeitsgemeinschaft Geriatrisches Basisassessment (AGAST). *Zeitschrift für Gerontologie & Geriatrie*, 36, 223–232. <https://doi.org/10.1007/s00391-003-0098-9>.
- Roalf, D.R., Moberg, P.J., Xie, S.X., Wolk, D.A., Moelter, S.T., & Arnold, S.E. (2013). Comparative accuracies of two common screening instruments for classification of Alzheimer's disease, mild cognitive impairment, and healthy aging. *Alzheimer's & Dementia*, 9(5), 529–537. <https://doi.org/10.1016/j.jalz.2012.10.001>.
- Robert Koch-Institut (Ed.). (2015). *Gesundheit in Deutschland. Gesundheitsberichterstattung des Bundes*. Berlin: RKI, Destatis.
- Ryan, R.M., & Deci, E.L. (2001). On happiness and human potentials: a review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology*, 52(1), 141–166. <https://doi.org/10.1146/annurev.psych.52.1.141>.
- Ryff, C.D., & Keyes, C.L. (1995). The structure of psychological well-being revisited. *Journal of Personality & Social Psychology*, 69, 719–727.
- Saadeh, M., Welmer, A.K., Dekhtyar, S., Fratiglioni, L., & Calderón-Larrañaga, A. (2020). The Role of psychological and social well-being on physical function trajectories in older adults. *Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, 75(8), 1579–1585. <https://doi.org/10.1093/gerona/glaa114>.
- Schumacher, J., Klaiberg, A., & Brähler, E. (2003). *Diagnostische Verfahren zu Lebensqualität und Wohlbefinden (Diagnostik für Klinik und Praxis)* (1st edn.). Hogrefe.
- Seligman, M. (2018). PERMA and the building blocks of well-being. *The Journal of Positive Psychology*, 13(4), 333–335. <https://doi.org/10.1080/17439760.2018.1437466>.
- Shumway-Cook, A., Brauer, S., & Woollacott, M. (2000). Predicting the probability for falls in community-dwelling older adults using the Timed Up & Go Test. *Physical Therapy*, 80(9), 896–903. <https://doi.org/10.1093/ptj/80.9.896>.
- Skevington, S.M., & Böhnke, J.R. (2018). How is subjective well-being related to quality of life? do we need two concepts and both measures? *Social Science & Medicine*, 206, 22–30. <https://doi.org/10.1016/j.socscimed.2018.04.005>.
- Smalbrugge, M., Pot, A.M., Jongenelis, L., Gundy, C.M., Beekman, A.T., & Eefsting, J.A. (2006). The impact of depression and anxiety on well being, disability and use of health care services in nursing home patients. *International Journal of Geriatric Psychiatry*, 21(4), 325–332. <https://doi.org/10.1002/gps.1466>.
- Statistisches Bundesamt. (2020). Pflegestatistik 2019. Verfügbar unter [https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Gesundheit/Pflege/\\_inhalt.html](https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Gesundheit/Pflege/_inhalt.html) (20.3.2021)
- Su, R., Tay, L., & Diener, E. (2014). The Development and Validation of the Comprehensive Inventory of Thriving (CIT) and the Brief Inventory of Thriving (BIT). *Applied Psychology: Health and Well-Being*, 6(3), 251–279. <https://doi.org/10.1111/aphw.12027>
- Subaşı, F., & Hayran, O. (2005). Evaluation of life satisfaction index of the elderly people living in nursing homes. *Archives of Gerontology and Geriatrics*, 41(1), 23–29.
- Testa, M.A., & Simonson, D.C. (1996). Assessment of quality-of-life outcomes. *New England Journal of Medicine*, 334(13), 835–840. <https://doi.org/10.1056/NEJM199603283341306>.
- THE WHOQOL GROUP. (1998) The World Health Organization quality of life assessment (WHOQOL): development and general psychometric properties. *Social Science and Medicine*, 46(12), 1569–1585. [https://doi.org/10.1016/s0277-9536\(98\)00009-4](https://doi.org/10.1016/s0277-9536(98)00009-4)
- The World Health Organization Quality of Life assessment (WHOQOL) (1995). Position paper from the World Health Organization. *Soc Sci Med*;41(10):1403-9. [https://doi.org/10.1016/0277-9536\(95\)00112-k](https://doi.org/10.1016/0277-9536(95)00112-k).
- West, S.G., Finch, J.F., & Curran, P.J. (1995). Structural equation models with nonnormal variables: problems and remedies. In R.H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues and applications* (pp. 56–75). Newbery Park: SAGE.
- Wilson, I.B., & Cleary, P.D. (1995). Linking clinical variables with health-related quality of life: a conceptual model of patient outcomes. *JAMA*, 273(1), 59–65.
- WMA, 2013 World Medical Association. (2013). World Medical Association Declaration of Helsinki ethical principles for medical research involving human subjects. *JAMA: Journal of the American Medical Association*, 310(20), 2191–2194
- van der Wolf, E., van Hooren, S.A.H., Waterink, W., & Lechner, L. (2018). Measurement of well-being in gerontopsychiatric nursing home residents: development of the Laurens well-being inventory for gerontopsychiatry. *Journal of Geriatric Psychiatry and Neurology*, 31(3), 136–148. <https://doi.org/10.1177/0891988718781031>.
- van der Wolf, E., van Hooren, S.A.H., Waterink, W., & Lechner, L. (2019). Well-being in elderly long-term care residents with chronic mental disorder: a systematic review. *Aging & Mental Health*, 23(3), 287–296. <https://doi.org/10.1080/13607863.2017.1408773>.
- World Health Organization (WHO) (2010). Process of translation and adaptation of instruments. WHO. [https://www.who.int/substance\\_abuse/research\\_tools/translation/en/](https://www.who.int/substance_abuse/research_tools/translation/en/) (Created 5 Nov 2010). Accessed 17 Jan 2019.
- World Medical Association (2015). *Declaration of Helsinki. Ethical principles for medical research involving human subjects*. 64th WMA General Assembly, Fortaleza.