


# Evaluation of the Responsiveness of the SarQoL<sup>®</sup> Questionnaire, a Patient-Reported Outcome Measure Specific to Sarcopenia

Anton Geerinck  · Olivier Bruyère · Médéa Locquet ·  
Jean-Yves Reginster · Charlotte Beaudart

Received: September 12, 2018  
© The Author(s) 2018

## ABSTRACT

**Introduction:** The Sarcopenia Quality of Life (SarQoL<sup>®</sup>) questionnaire was developed to provide a patient-reported outcome measure specific to sarcopenia. Its psychometric properties indicate that it is a valid and reliable instrument. However, until now, its ability to detect change over time has not been examined. Therefore, the objective of this study is to evaluate the responsiveness (also known as sensitivity to change) of the SarQoL<sup>®</sup> questionnaire in a prospective, longitudinal cohort of community-dwelling, older, sarcopenic subjects.

**Methods:** Sarcopenic subjects from the SarcoPhAge (Sarcopenia and Physical impairment with advancing Age) study were included. Responsiveness was evaluated with nine pre-specified hypotheses on the correlation between the evolution of the SarQoL<sup>®</sup> scores after a 2-year interval and the evolution of the scores on the Short Form-36 (SF-36) and the Euroqol 5-dimension (EQ-5D) questionnaires. This

**Enhanced digital features** To view enhanced digital features for this article go to <https://doi.org/10.6084/m9.figshare.7195778>.

technique considers responsiveness to be a form of longitudinal validity. Additionally, standardized response means were also calculated to compare the quantity of change measured by the different questionnaires.

**Results:** A total of 42 sarcopenic subjects were included. The median age of the sample was 72.9 (68.9–78.8) years, 59.5% were female, and the mean body mass index was 23.3 (20.4–25.7) kg/m<sup>2</sup>. A good responsiveness was observed, as evidenced by the confirmation of eight out of nine hypotheses, well above the 75% confirmation threshold. The standardized response mean of the Overall SarQoL<sup>®</sup> score was significantly higher than those of the SF-36 Physical Component Summary ( $p = 0.005$ ), the EQ-5D Utility Index ( $p < 0.001$ ) and the Euroqol visual analogue scale ( $p = 0.003$ ).

**Conclusion:** The first data available on the ability of the SarQoL<sup>®</sup> questionnaire to detect change over time indicates that the questionnaire has good responsiveness. This, together with the previously established psychometric properties, confirms that the SarQoL<sup>®</sup> questionnaire is a relevant instrument for the assessment of quality of life in sarcopenic populations.

A. Geerinck (✉) · O. Bruyère · M. Locquet ·  
J.-Y. Reginster · C. Beaudart  
Department of Public Health, Epidemiology and  
Health Economics, University of Liège, Liège,  
Belgium  
e-mail: [anton.geerinck@uliege.be](mailto:anton.geerinck@uliege.be)

**Keywords:** Older people; Patient-reported outcome measure; Psychometrics; Quality of life; Questionnaire; Responsiveness; Sarcopenia

---

## INTRODUCTION

### Background

Sarcopenia, defined as “a syndrome characterised by progressive and generalised loss of skeletal muscle mass and strength and with a risk of adverse outcomes such as physical disability, poor quality of life and death” by the European Working Group on Sarcopenia in Older People (EWGSOP), is a growing public health problem [1]. It has recently been recognized as a geriatric condition with an ICD-10-CM code (M62.84) [2]. Sarcopenia has been shown to be associated with negative health outcomes, such as a higher rate of mortality and functional decline, a higher rate of falls, and a higher incidence of hospitalization [3]. Other research has shown an association between sarcopenia and depression [4]. Not much is yet known about the relationship between sarcopenia and quality of life. Although several studies have incorporated quality of life outcomes in their designs, the results are difficult to compare because of the different diagnostic criteria used to establish sarcopenia. Some studies that diagnosed sarcopenia with the EWGSOP criteria have found lower health-related quality of life (HRQoL) scores for sarcopenic subjects in select domains of the Short-Form 36-item (SF-36) questionnaire, but other studies (using other diagnostic criteria) have found no difference in SF-36 scores between sarcopenic and non-sarcopenic subjects [5].

Until recently, researchers only had generic questionnaires, such as the SF-36, available to assess quality of life in sarcopenic patients. These questionnaires are designed for use in broad populations and may thus not be sensitive enough to accurately measure quality of life in sarcopenic populations [6]. To address this problem, Beudart et al. developed the Sarcopenia Quality of Life (SarQoL<sup>®</sup>) questionnaire in 2015 [7].

Until now, no study has evaluated the responsiveness, defined as “the ability of an instrument to detect change over time in the construct to be measured”, of the SarQoL<sup>®</sup> questionnaire [8]. When an instrument is used for

evaluative purposes, i.e. when the aim is to detect and measure longitudinal change in subjects or populations, responsiveness is a key psychometric property [9, 10]. This situation is often present in clinical studies aimed at testing the effect of an intervention, where an accurate assessment of HRQoL before and after the intervention is an important outcome. Researchers need to have valid data on the responsiveness of the instrument they wish to use to be certain of the results they obtain.

The psychometric properties of the SarQoL<sup>®</sup> questionnaire have been evaluated in several cross-sectional studies, but until now, its ability to detect change over time (responsiveness) had not yet been examined [11–14]. This study aimed to evaluate the responsiveness of the SarQoL<sup>®</sup> questionnaire in a sample of older, community-dwelling, sarcopenic subjects from the SarcoPhAge (Sarcopenia and Physical impairment with advancing Age) cohort.

## METHODS

### Design

The current article describes an instrument validation study that examined data collected at the 2nd and 4th annual visit of the SarcoPhAge study, an ongoing 5-year prospective, longitudinal, observational cohort study being carried out in Liège (Belgium) [15, 16]. Participants in the SarcoPhAge study all provided written informed consent. The research protocol and its amendments were approved by the Ethics Committee of the University Teaching Hospital of Liège (no. 2012-277).

### Participants

Participants from the SarcoPhAge study with valid data from the 2nd (T1) and 4th (T3) study visit (a 2-year interval) who were diagnosed as sarcopenic according to the EWGSOP criteria were included [1]. This 2-year interval was chosen because it covers the first and last available administrations of the questionnaire and because the SarcoPhAge study is an

observational study; therefore, we relied on the natural progression of sarcopenia to cause a change in health status between the two measurements. The details of this study have been reported previously [11, 15, 17, 18].

Sarcopenia was diagnosed according to the EWGSOP algorithm, which demands the presence of low muscle mass in combination with low muscle strength and/or low physical performance [1]. Muscle mass was measured by dual-energy X-ray absorptiometry (DXA) (Hologic Discovery A, USA), which was calibrated daily by scanning a spine phantom. Male subjects with a skeletal muscle mass index (SMI = appendicular lean mass/height<sup>2</sup>) below 7.26 kg/m<sup>2</sup> and women with an SMI below 5.5 kg/m<sup>2</sup> were considered to have low muscle mass. Muscle strength was measured with a hydraulic hand dynamometer (Saehan Corporation, Korea), calibrated at the beginning of the study for 10, 40 and 90 kg. Men with a maximal handgrip strength below 30 kg and women below 20 kg were considered to have low muscle strength. Physical performance was examined with the help of the Short Physical Performance Battery (SPPB), with a value of 8 or less being considered low [15].

Participants were included in the current analysis when diagnosed as sarcopenic at T1 and/or T3 and when both SarQoL<sup>®</sup> questionnaires (T1 and T3) had less than 20% missing data for the calculation of the Overall score.

## Measures

### *The SarQoL<sup>®</sup> Questionnaire*

The SarQoL<sup>®</sup> questionnaire is a patient-reported outcome measure (PROM) specific to sarcopenia. The SarQoL<sup>®</sup> questionnaire consists of 22 questions incorporating 55 items, which fall into seven domains of HRQoL. These domains are “Physical and Mental Health”, “Locomotion”, “Body Composition”, “Functionality”, “Activities of Daily Living”, “Leisure activities” and “Fears”. Each domain is scored from 0 to 100, and an Overall score is calculated. The questionnaire is auto-administered and takes 10 min to complete [7]. The questionnaire is available in 16 languages and can be found on its webpage [19].

Several psychometric properties of the SarQoL<sup>®</sup> questionnaire have been examined previously. The questionnaire has demonstrated its ability to distinguish between sarcopenic and non-sarcopenic subjects (discriminative power). It has good internal consistency and construct validity, and its test–retest reliability is excellent. Furthermore, it has been demonstrated that there are no floor or ceiling effects for the Overall score [11–14].

### *The Short-Form 36-Item (SF-36) Questionnaire*

The SF-36 is a multi-item generic health survey that uses 36 questions to measure functional health and wellbeing from the patient’s perspective. It measures eight domains: “Physical Functioning”, “Role limitation due to physical problems”, “Bodily Pain”, “General Health Perceptions”, “Vitality”, “Social Functioning”, “Role limitations due to emotional problems” and “Mental Health”, each of which provides a score between 0 and 100. Additionally, two composite scores can be calculated: the Physical Component Summary (PCS) and the Mental Component Summary (MCS) [20–22].

### *The EuroQol 5-Dimension 3-Level (EQ-5D-3L)*

The EQ-5D-3L is a standardized measure of health status developed by the EuroQol Group in 1990. The instrument consists of two pages: the EQ-5D descriptive system, which is composed of five questions encompassing five dimensions of health (mobility, self-care, usual activities, pain/discomfort and anxiety/depression); and the Visual Analogue Scale (EQ-VAS), which records the respondent’s self-rated health on a vertical scale going from best (100) to worst imaginable health (0). The EQ-5D descriptive system is used to calculate an index score, which represents the utility value for current health [23, 24].

## Physical Parameters

Parameters related to muscle mass, muscle strength and physical performance were collected. Apart from the SMI, we also determined appendicular lean mass (ALM) and ALM divided

by body mass index (ALM/BMI) by DXA. As mentioned previously, muscle strength was determined with a hydraulic hand dynamometer. For physical performance, the patients performed the SPPB test, which also includes the usual gait speed on a 4-m track. The subjects also performed the timed-up-and-go (TUG) test, which uses the time that a subject takes to rise from a chair, walk three metres, turn around, walk back to the chair, and sit down to determine a subject's mobility. Lastly, the chair stand test (CST) was administered as part of the SPPB. In this test, the subjects are asked to stand up from a chair and sit back down five times as fast as they can.

## Methodological Approach

### *Hypotheses Testing*

It is recommended to treat responsiveness as the longitudinal form of construct validity and to evaluate it in much the same way as the construct validity of a questionnaire [25]. Thus, we formulated hypotheses between the changes in the scores of the SarQoL<sup>®</sup> questionnaire and the changes observed for the SF-36 and the EQ-5D. AG, CB and OB were responsible for the formulation of the hypotheses, on the basis of similarity in the construct of the different domains, and previously found results for the construct validity of the questionnaire. The data used in this analysis were collected before the formulation of the hypotheses, but no statistical manipulations in relation to the evaluation of responsiveness were carried out before the final set of hypotheses was agreed upon.

The hypotheses used for the evaluation of the responsiveness, the expected strength of the correlations and the rationale for their formulation are detailed in Table 1.

We employed the criteria formulated by De Boer et al. to evaluate the results of the hypotheses testing. These state that a questionnaire has high responsiveness when less than 25% of hypotheses are refuted, moderate responsiveness when 25–50% are refuted and poor responsiveness when more than 50% are refuted [26].

### *Standardized Response Means (SRMs)*

We also calculated SRMs for the different questionnaires, by dividing the mean difference between T1 and T3 by the standard deviation of the differences between the paired measurements [27]. The SRM reflects the magnitude of the change measured by the different questionnaires. Consequently, when greater SRMs are obtained, this is an indication of better responsiveness. To allow the use of the thresholds for responsiveness formulated by Cohen et al., which are designed for use with the effect size and which categorize an observed change, we applied the correction developed by Middel and Van Sonderen [28, 29]. After correcting the SRMs with the formula  $[(SRM/\sqrt{2})/\sqrt{(1-r)}]$ ; with  $r$  = correlation between baseline and follow-up score], we categorized them as trivial when  $SRM < 0.20$ , small when  $0.20 \leq SRM < 0.49$ , moderate when  $0.50 \leq SRM < 0.79$  and large when  $SRM \geq 0.80$  [29].

A selection of SRMs were compared in pairs to evaluate whether they were significantly different. This was carried out using the modified jack-knife method, which uses linear regression to determine whether a significant difference exists between two SRMs [30]. For this measure, an individual SRM is first calculated for each subject by dividing their change score by the standard deviation of the change scores in the whole sample. Next, a "centred" SRM is calculated for each subject by subtracting the mean SRM score of the sample from the individual SRMs. With these variables, a linear regression is carried out with the individual SRMs of the two quality-of-life scores of interest as dependent variables and the "centred" SRM of one of the quality-of-life scores (either one will work) as the independent variable. A significant difference is demonstrated when the  $p$  value of the intercept is at most 0.05 [30, 31].

### *Correlations Between Physical Parameters and QoL*

We investigated the relationship between the evolution of physical parameters linked to sarcopenia and the changes observed by the different questionnaires with the help of correlations. We selected the five summary/total scores available (SarQoL<sup>®</sup> Overall score, SF-

**Table 1** Hypotheses for the evaluation of responsiveness

Hypotheses	Expected strength of correlation	Rationale
1. $\Delta$ SarQoL Overall score and $\Delta$ SF-36 General Health domain are correlated	$r > 0.4$	The SarQoL Overall score and the SF-36 General Health score have been shown to be correlated in the French ( $r = 0.67$ ) and the English ( $r = 0.49$ ) validations. These domains are similar in that they both measure a subject's general view of either their HRQoL or health. Because of the strong interaction between general health status and HRQoL, we expect a correlation of at least 0.4, despite the difference in underlying construct
2. $\Delta$ SarQoL Overall score and $\Delta$ SF-36 Vitality domain are correlated	$r > 0.3$	Here also, two different constructs are measured, but they have been shown to be correlated (FR: $r = 0.72$ ; ENG: $r = 0.74$ ). Since the underlying constructs are less similar than in hypothesis 1, and we expect the influence of a change in vitality to be less impactful than one in General Health, the expected correlation was set to at least 0.3
3. $\Delta$ SarQoL Overall score and $\Delta$ SF-36 Physical Functioning domain are correlated	$r > 0.5$	The domain Physical Functioning covers a significant portion of the content used to calculate the Overall score of the SarQoL <sup>®</sup> , although the Overall score also takes into account other aspects of HRQoL. The English validation confirmed this similarity with a correlation of 0.82, although the French validation found a smaller correlation of 0.49. Nevertheless, we expect changes on both measures to be correlated at a strength of at least 0.5
4. $\Delta$ SarQoL Overall score and $\Delta$ EQ-VAS are correlated	$r > 0.4$	The Overall score and the EQ-VAS both give a general view of the subjects' current health or HRQoL, and should thus, in theory, be correlated. We expect the difference in health as measured by the EQ-VAS to be reflected in changes in HRQoL (as evidenced by a cross-sectional correlation of $r = 0.597$ ) but, since they measure two different but related constructs, it was decided to fix the expected strength of this association to at least 0.4

**Table 1** continued

Hypotheses	Expected strength of correlation	Rationale
5. $\Delta$ SarQoL domain 1 (Physical and Mental Health) and $\Delta$ SF-36 General Health domain are correlated	$r > 0.3$	Domain 1 of the SarQoL <sup>®</sup> questionnaire carries significant weight in the calculation of the Overall score. Since we know a correlation exists between the Overall score and the General Health domain for the construct validity (see hypothesis 1), we theorized that this same correlation should exist between Physical and Mental Health and General Health. We did expect this correlation to be weaker, although the cross-sectional correlation was $r = 0.655$ , since some aspects covered in the Overall score are not represented in Physical and Mental Health. It was decided to expect a correlation of at least 0.3
6. $\Delta$ SarQoL domain 1 (Physical and Mental Health) and $\Delta$ EQ-VAS are correlated	$r > 0.3$	In the same vein as hypothesis 5, we expected changes on Physical and Mental Health to be associated with changes on the EQ-VAS, as shown by a cross-sectional correlation of $r = 0.562$ . However, since a part of the content is lost when focusing on a single domain of the SarQoL <sup>®</sup> , it was decided to expect a weaker correlation than hypothesis 5, and to adopt at least 0.3 as the threshold
7. $\Delta$ SarQoL domain 2 (Locomotion) and $\Delta$ SF-36 Physical Functioning domain are correlated	$r > 0.4$	The ability to walk and the ease with which a person can walk are an important factor that influences the totality of how a person functions physically, demonstrated by a cross-sectional correlation of $r = 0.558$ . While the domain Locomotion is a much narrower construct than Physical Functioning, we expect both domains to be significantly correlated at a strength of at least 0.4
8. $\Delta$ SarQoL domain 4 (Functionality) and $\Delta$ SF-36 Physical Functioning domain are correlated	$r > 0.5$	The underlying constructs of the domains Functionality and Physical Functioning are, in theory, similar, and it was therefore felt that a relatively strong correlation of at least 0.5 was to be expected, even if the cross-sectional correlation was lower at $r = 0.420$

**Table 1** continued

Hypotheses	Expected strength of correlation	Rationale
9. $\Delta$ SarQoL domain 5 (Activities of Daily Living) and $\Delta$ SF-36 Physical Functioning domain are correlated	$r > 0.5$	While these two domains represent different underlying constructs, we theorized that a change in physical functioning would be equally reflected in a change in a person's Activities of Daily Living, because one is a prerequisite for the other. It was felt that we should expect a relatively strong correlation of at least 0.5 since we expected these two domains to be interwoven even if the cross-sectional correlation was lower at $r = 0.460$

$\Delta$  = change in;  $r$  = correlation

36 PCS and MCS, EQ-5D Utility Index and EQ-VAS) to represent the HRQoL of the subjects and constructed correlations with usual gait speed, handgrip strength, SPPB score, ALM, ALM/BMI, SMI, TUG test and the chair stand test. The strength of the association was judged as excellent when larger than 0.81, very good when between 0.61 and 0.80, good when between 0.41 and 0.60, acceptable when between 0.21 and 0.40 and insufficient when less than 0.20 [32].

### Statistical Analysis

Data were analysed using IBM SPSS Statistics, version 24.0.0.0 for Windows (Armonk, NY: IBM Corp).

The distribution of variables was determined by examining the histogram, the quantile–quantile plot, the Shapiro–Wilk test and the difference between mean and median. Gaussian variables are reported as the mean  $\pm$  standard deviation and non-Gaussian variables as median (P25–P75). Nominal variables are reported as absolute ( $n$ ) and relative frequencies (%). The presence of significant differences between T1 and T3 was examined with the paired samples  $t$  test for variables with normal distribution, the Wilcoxon matched-pair signed-rank test for non-Gaussian variables and the chi-squared test for nominal variables. Pearson correlations were calculated when both groups/variables had

normal distributions. Spearman correlations were calculated when this was not the case.

Change scores were calculated by subtracting the scores from T1 from those obtained at T3. For quality of life, this means that a positive change score indicates an improvement and a negative change score a decline. The calculation of the SRMs, their correction with the technique from Middel and Van Sonderen and the modified jack-knife method used to detect significant differences between SRMs have been described in the preceding paragraphs.

A post hoc power analysis was conducted on the Pearson and Spearman correlations used in the primary outcome with the G\*Power software, version 3.1.9.2 [33]. This analysis computes the achieved power for a bivariate normal model with an  $\alpha$ -error of 0.05 and a sample size of 42 subjects.

Results were considered significant at  $p \leq 0.05$ .

## RESULTS

In total, 42 sarcopenic participants from the SarcoPhAge study fulfilled the inclusion criteria, which is a moderate sample size according to the COSMIN checklist [34]. The subjects had a median age of 73 (69–79) years at T1, and 25 out of 42 (59.5%) were women. The median number of drugs taken by the participants increased

significantly ( $p = 0.001$ ) from 6 (5–9) at T1 to 8 (6–10) at T3, as did the proportion of subjects who fell in the year before the study visits, from 8 (19.0%) at T1 to 16 (38.1%) at T3 ( $p = 0.017$ ). The gait speed of the participants diminished significantly from a median of 1.02 (0.80–1.21) m/s at T1 to 0.89 (0.76–1.09) m/s at T3 ( $p = 0.032$ ). In the sample as a whole, a slight but significant reduction in handgrip strength was observed, from a median of 19.75 (18.00–28.00) kg at T1 to 19.00 (16.75–22.50) kg at T3 ( $p = 0.010$ ). This change was attributable to the female subjects ( $p = 0.030$ ). No significant changes between T1 and T3 were found for BMI ( $p = 0.393$ ), number of comorbidities ( $p = 0.763$ ), proportion of subjects who experienced a fracture in the year before the study visits ( $p = 0.268$ ), independence in activities of daily living as measured by the Katz scale (0.942), SPPB score ( $p = 0.083$ ), TUG test ( $p = 0.081$ ), ALM/BMI ( $p = 0.197$ ) and SMI ( $p = 0.451$ ). The ALM of the whole sample diminished significantly ( $p = 0.035$ ), but this effect was lost when the sample was divided into men ( $p = 0.287$ ) and women ( $p = 0.072$ ).

The three different questionnaires obtained different results for quality of life. The SarQoL<sup>®</sup> questionnaire measured a significant reduction for three domains (Body Composition,  $p = 0.023$ ; Functionality,  $p = 0.002$ ; Activities of Daily Living,  $p < 0.001$ ) and the Overall score, which diminished from a median of 61.15 (51.15–71.76) at T1 to 54.56 (42.31–68.44) at T3 ( $p = 0.002$ ). The SF-36 PCS and MCS, the EQ-5D Utility Index and the EQ-VAS, however, did not detect a significant change (respectively,  $p = 0.679$ ,  $p = 0.062$ ,  $p = 0.231$  and  $p = 0.716$ ). The complete clinical characteristics and the evolution of quality of life can be found in Table 2.

### Responsiveness

Of the nine formulated hypotheses, 8 (89%) were confirmed. Hypothesis 9 was rejected when a correlation of  $r = 0.467$  was found, just under the threshold of  $r > 0.5$ . In total, three very good correlations were found, five good correlations and two acceptable correlations.

The results of this evaluation as well as of the power analysis are reported in Table 3.

According to the criteria by De Boer et al., the SarQoL<sup>®</sup> questionnaire possesses high responsiveness because fewer than 25% of hypotheses are refuted [26].

### Standardized Response Means

The magnitude of change observed in the sample was examined by calculating SRMs. The SarQoL<sup>®</sup> questionnaire had three domains with SRMs below 0.20, indicating that no change was observed, two domains with an SRM between 0.20 and 0.49 (small change) and three domains with an SRM between 0.50 and 0.79 (moderate change). In contrast, only one domain of the SF-36 had a moderate SRM (Physical Functioning; SRM =  $-0.50$ ), and six domains reported an SRM indicating small change. A further three domains of the SF-36 had SRMs indicating no change had occurred. For the EQ-5D, small SRMs were observed for two domains, with the remaining five domains having SRMs indicating no change. All obtained SRMs can be found in Table 4.

The SRM of the SarQoL<sup>®</sup> Overall score was significantly larger than the SF-36 PCS ( $p = 0.005$ ), the EQ-5D Utility Index ( $p < 0.001$ ) and the EQ-VAS ( $p = 0.003$ ). The SRMs of the SarQoL<sup>®</sup> Overall score and the SF-36 MCS were not significantly different ( $p = 0.150$ ). The results of this analysis are reported in Table 5.

### Correlations Between Physical Parameters and QoL

Good correlations were found between change in the SarQoL<sup>®</sup> Overall score and change in gait speed ( $r = 0.50$ ), SPPB score ( $r = 0.47$ ) and the chair stand test ( $r = -0.42$ ). Good correlations were also found between change in ALM/BMI and change on the EQ-VAS ( $r = -0.48$ ) as well as between change on the timed up-and-go test and change on the SF-36 PCS ( $r = -0.44$ ). Acceptable correlations were found between change in gait speed and change on the SF-36 PCS ( $r = 0.39$ ), between change on the chair stand test and change on the SF-36 PCS



**Table 2** Clinical characteristics and quality of life scores for sarcopenic sample ( $n = 42$ )

	<b>T1</b>	<b>T3</b>	<b>Change</b>	<b>P value</b>
Age (years)	72.90 (68.85–78.81)	NA	NA	NA
Gender				
Male	17 (40.5%)	NA	NA	NA
Female	25 (59.5%)	NA	NA	NA
BMI (kg/m <sup>2</sup> )	23.25 (20.35; 25.68)	23.09 (20.06; 25.84)	– 0.03 (– 0.67; 0.58)	0.393 <sup>a</sup>
Number of drugs	6.00 (5.00; 9.00)	8.00 (6.00; 10.00)	1.00 (0.00; 3.00)	0.001 <sup>a</sup>
Number of comorbidities	4.00 (3.00; 6.25)	4.00 (2.75; 7.00)	0.00 (0.00; 0.00)	0.763 <sup>a</sup>
Fall in last year				
Yes	8 (19.0%)	16 (38.1%)	NA	0.017 <sup>b</sup>
No	34 (81.0%)	26 (61.9%)	NA	
Fracture in last year				
Yes	4 (9.5%)	4 (9.5%)	NA	0.268 <sup>b</sup>
No	38 (90.5%)	38 (90.5%)	NA	
Katz score	8.00 (8.00; 9.00)	8.00 (8.00; 9.00)	0.00 (0.00; 0.00)	0.942 <sup>a</sup>
SPPB score	9.50 (8.00; 11.00)	8.00 (6.75; 11.00)	– 0.50 (– 2.00; 0.25)	0.083 <sup>a</sup>
Gait speed (m/s)	1.02 (0.80; 1.21)	0.89 (0.76; 1.09)	– 0.10 (– 0.26; 0.14)	0.032 <sup>a</sup>
Chair stand test (s)	14.57 (11.97; 18.29)	16.07 (11.06; 20.94)	1.06 (– 0.86; 3.34)	0.083 <sup>a</sup>
Timed up-and-go (s)	10.67 (8.66; 13.31)	12.23 (9.15; 16.27)	0.88 (– 1.18; 3.14)	0.081 <sup>a</sup>
Hand grip strength (kg)	19.75 (18.00; 28.00)	19.00 (16.75; 22.50)	– 1.50 (– 5.25; 1.00)	0.010 <sup>a</sup>
HGS men (kg)	28.00 (21.00; 37.00)	25.00 (20.50; 31.50)	– 1.00 (– 7.50; 1.00)	0.146 <sup>a</sup>
HGS women (kg)	19.00 (14.50; 20.25)	18.00 (12.00; 19.25)	– 2.00 (– 4.25; 1.00)	0.030 <sup>a</sup>
ALM (kg)	14.31 (13.09; 18.73)	14.02 (12.94; 18.14)	– 0.29 (– 0.57; 1.16)	0.035 <sup>a</sup>
ALM men	18.92 (17.44; 20.26)	19.10 (16.79; 19.97)	– 0.30 (– 0.63; 0.30)	0.287 <sup>a</sup>
ALM women	13.40 (12.47; 14.15)	13.21 (12.03; 14.00)	– 0.29 (– 0.48; 0.16)	0.072 <sup>c</sup>
ALM/BMI	0.69 (0.57 (0.74)	0.67 (0.58–0.74)	– 0.01 (– 0.04; 0.02)	0.197 <sup>a</sup>
ALM/BMI men	0.74 (0.70; 0.88)	0.74 (0.69; 0.89)	– 0.01 (– 0.03; 0.03)	0.940 <sup>a</sup>
ALM/BMI women	0.60 (0.53; 0.70)	0.60 (0.51; 0.67)	– 0.02 (– 0.04; 0.02)	0.141 <sup>a</sup>
SMI (kg/m <sup>2</sup> )	5.57 (5.25–6.70)	5.51 (5.14–6.60)	– 0.02 (– 0.18; 0.15)	0.451 <sup>a</sup>
SMI men	6.86 (6.37–7.16)	6.84 (6.26; 7.30)	– 0.01 (– 0.20; 0.17)	0.454 <sup>a</sup>
SMI women	5.26 (5.11; 5.52)	5.32 (4.94; 5.47)	– 0.04 (– 0.19; 0.10)	0.440 <sup>c</sup>
SarQoL D1				
Physical and Mental Health	58.87 (45.53; 69.15)	51.09 (41.37; 67.19)	– 5.00 (– 12.51; 4.72)	0.107 <sup>a</sup>

**Table 2** continued

	T1	T3	Change	P value
SarQoL D2				
Locomotion	55.56 (46.53; 72.22)	55.56 (38.20; 70.14)	− 2.78 (− 11.81; 5.55)	0.331 <sup>c</sup>
SarQoL D3				
Body Composition	58.33 (45.83; 67.71)	50.00 (41.67; 60.63)	− 4.16 (− 12.92; 4.17)	0.023 <sup>c</sup>
SarQoL D4				
Functionality	70.24 (59.49; 82.85)	63.46 (47.60; 75.89)	− 4.55 (− 10.70; 1.78)	0.002 <sup>a</sup>
SarQoL D5				
Activities of Daily Living	61.61 (43.33; 75.00)	48.22 (37.29; 65.42)	− 6.43 (− 20.00; − 3.12)	< 0.001 <sup>a</sup>
SarQoL D6				
Leisure activities	33.25 (29.09; 49.88)	33.25 (16.62; 66.50)	0.00 (− 16.62; 16.62)	0.645 <sup>a</sup>
SarQoL D7				
Fears	87.50 (75.00; 100.00)	87.50 (75.00; 100.00)	0.00 (− 12.50; 0.00)	0.382 <sup>a</sup>
SarQoL Overall score	61.15 (51.15; 71.76)	54.56 (42.31; 68.44)	− 5.23 (− 12.46; 1.61)	0.002 <sup>a</sup>
SF-36 PCS	42.08 (31.86; 49.14)	37.65 (30.47; 48.24)	1.40 (− 5.36; 4.78)	0.679 <sup>a</sup>
SF-36 MCS	44.71 (33.86; 53.31)	38.91 (30.55; 50.40)	− 2.18 (− 10.13; 3.77)	0.062 <sup>a</sup>
EQ-5D Utility Index	0.800 (0.517–0.827)	0.800 (0.708–0.827)	0.00 (− 0.193; 0.1557)	0.231 <sup>a</sup>
EQ-VAS	70.00 (60.00–75.00)	70.00 (60.00–75.00)	0.00 (− 7.50; 5.00)	0.716 <sup>a</sup>

NA not applicable, PCS Physical Component Summary, MCS Mental Component Summary

<sup>a</sup> Wilcoxon matched-pair signed-rank test

<sup>b</sup> Chi-squared test

<sup>c</sup> Paired samples *t* test

( $r = -0.37$ ) and the SF-36 MCS ( $r = -0.36$ ). No other correlations were statistically significant. The full analysis can be found in Table 6.

## DISCUSSION

The aim of this study was to evaluate the responsiveness of the SarQoL<sup>®</sup> questionnaire in a population of older, community-dwelling, sarcopenic subjects by formulating hypotheses on the correlations between change scores, and by calculating the standardized response means. Additionally, we examined the correlations between changes in physical parameters and the evolution of the quality-of-life scores.

The results from the hypotheses reveal that the SarQoL<sup>®</sup> questionnaire has high

responsiveness according to the criteria of De Boer et al., with only one hypothesis out of nine (11%) refuted [26]. The most notable results are the strong correlations found for the Overall score and domain 4 (Functionality) of the SarQoL<sup>®</sup> questionnaire, and the Physical Functioning domain of the SF-36. These correlations, respectively  $r = 0.669$  and  $r = 0.680$ , were larger than the expected correlation of  $r = 0.5$  but make sense in light of the similarity of their content and the relatively important weight of domain 4 in the calculation of the Overall score of the SarQoL<sup>®</sup> questionnaire.

The SRMs show that the change measured by the Overall score of the SarQoL<sup>®</sup> questionnaire was significantly larger than that measured by the SF-36 PCS, the EQ-5D utility index and the

**Table 3** Evaluation of responsiveness with hypotheses

Hypothesis	Expected strength of correlation	Observed correlation		Confirmation/rejection	Power ( $1 - \beta$ )
		<i>r</i>	<i>p</i> value		
1. $\Delta$ SarQoL Overall score and $\Delta$ SF-36 General Health domain are correlated	$r > 0.4$	0.442 <sup>a</sup>	0.005	Confirmed	0.851
2. $\Delta$ SarQoL Overall score and $\Delta$ SF-36 Vitality domain are correlated	$r > 0.3$	0.454 <sup>b</sup>	0.004	Confirmed	0.872
3. $\Delta$ SarQoL Overall score and $\Delta$ SF-36 Physical Functioning domain are correlated	$r > 0.5$	0.669 <sup>a</sup>	< 0.001	Confirmed	0.999
4. $\Delta$ SarQoL Overall score and $\Delta$ EQ-VAS are correlated	$r > 0.4$	0.404 <sup>a</sup>	0.009	Confirmed	0.773
5. $\Delta$ SarQoL domain 1 (Physical & Mental Health) and $\Delta$ SF-36 General Health domain are correlated	$r > 0.3$	0.610 <sup>a</sup>	< 0.001	Confirmed	0.994
6. $\Delta$ SarQoL domain 1 (Physical & Mental Health) and $\Delta$ EQ-VAS are correlated	$r > 0.3$	0.312 <sup>a</sup>	0.047	Confirmed	0.531
7. $\Delta$ SarQoL domain 2 (Locomotion) and $\Delta$ SF-36 Physical Functioning domain are correlated	$r > 0.4$	0.412 <sup>a</sup>	0.010	Confirmed	0.791
8. $\Delta$ SarQoL domain 4 (Functionality) and $\Delta$ SF-36 Physical Functioning domain are correlated	$r > 0.5$	0.680 <sup>a</sup>	< 0.001	Confirmed	0.999
9. $\Delta$ SarQoL domain 5 (Activities of Daily Living) and $\Delta$ SF-36 Physical Functioning domain are correlated	$r > 0.5$	0.467 <sup>a</sup>	0.003	Rejected	0.893

$\Delta$  = change in; *r* = correlation

<sup>a</sup> Spearman correlation

<sup>b</sup> Pearson correlation

EQ-VAS, but not the SF-36 MCS. The absence of a significant difference between the SRM of the Overall score and the SF-36 MCS indicates a very large 95% confidence interval of the latter. The SRM obtained for the SarQoL<sup>®</sup> Overall score is in accordance with the change in physical parameters of the subjects. Participants lost approximately 10% of their original gait speed (from a median of 1.02 m/s to 0.89 m/s), and the female participants lost a median of 2 kg of grip strength in the 2-year interval. It is also interesting to note that the number of falls experienced in the year preceding the administration of the test doubled from 8 (19.0%) to 16 (38.1%). The SarQoL<sup>®</sup> Overall score more accurately reflects these changes, more so than the SF-36 and the EQ-5D.

The SarQoL<sup>®</sup> questionnaire measured an SRM indicating moderate change for domain 4 (Functionality) and domain 5 (Activities of Daily Living), highlighting that the effects of diminished muscle strength and physical performance manifest themselves most in all the physical tasks performed on a regular basis. SRMs indicating small change were reported for domain 1 (Physical and Mental Health) and domain 3 (Body Composition). The smaller SRM for domain 1 may result from the way the questions are formulated, with many more abstract concepts (energy, physical capacity, muscle mass, etc.) instead of the very relatable examples from domains 4 and 5 (climbing a flight of stairs, opening a bottle or jar, etc.). Subjects may have more difficulty finding the

**Table 4** Standardized response means

Domains	SRM	Corrected SRM	Interpretation <sup>a</sup>
1. $\Delta$ SarQoL D1 Physical and Mental Health	- 0.31	- 0.34	Small change
2. $\Delta$ SarQoL D2 Locomotion	- 0.15	- 0.19	No change
3. $\Delta$ SarQoL D3 Body Composition	- 0.37	- 0.47	Small change
4. $\Delta$ SarQoL D4 Functionality	- 0.50	- 0.62	Moderate change
5. $\Delta$ SarQoL D5 Activities of Daily Living	- 0.57	- 0.56	Moderate change
6. $\Delta$ SarQoL D6 Leisure activities	0.04	- 0.04	No change
7. $\Delta$ SarQoL D7 Fears	- 0.01	- 0.01	No change
8. $\Delta$ SarQoL Overall score	- 0.54	- 0.72	Moderate change
9. $\Delta$ SF-36 Physical Functioning	- 0.44	- 0.50	Moderate change
10. $\Delta$ SF-36 Social Functioning	- 0.41	- 0.48	Small change
11. $\Delta$ SF-36 Role Limitations due to Physical Health	0.02	- 0.02	No change
12. $\Delta$ SF-36 Role Limitations due to Emotional Problems	- 0.28	- 0.26	Small change
13. $\Delta$ SF-36 Mental Health	- 0.27	- 0.35	Small change
14. $\Delta$ SF-36 Vitality	- 0.03	- 0.03	No change
15. $\Delta$ SF-36 Bodily Pain	- 0.17	- 0.15	No change
16. $\Delta$ SF-36 General Health	- 0.23	- 0.28	Small change
17. $\Delta$ SF-36 Physical Component Summary	- 0.18	- 0.20	Small change
18. $\Delta$ SF-36 Mental Component Summary	- 0.29	- 0.34	Small change
19. $\Delta$ EQ-5D Mobility	0.10	- 0.08	No change
20. $\Delta$ EQ-5D Autonomy	- 0.36	NA <sup>b</sup>	Small change
21. $\Delta$ EQ-5D Usual activities	0.20	- 0.33	Small change
22. $\Delta$ EQ-5D Pain	- 0.07	- 0.06	No change
23. $\Delta$ EQ-5D Anxiety	- 0.19	- 0.17	No change
24. $\Delta$ EQ-5D Utility Index	0.19	0.18	No change
25. $\Delta$ EQ-VAS	- 0.11	- 0.09	No change

<sup>a</sup> Interpretation of corrected SRMs:  $0.20 \leq \text{SRM} < 0.49$  = small change;  $0.50 \leq \text{SRM} < 0.79$  = moderate change;  $\text{SRM} \geq 0.80$  = large change

<sup>b</sup> Correction for SRM of EQ-5D Autonomy cannot be computed because  $\Delta$  EQ-5D Autonomy at T3 is constant (all subjects responded with the same answer)

right answers for them because these changes are much less perceptible in absolute terms. The SRM for domain 3 (Body Composition) covers an area where drastic change is not necessarily expected given that the median age in the sample is 73 years old and that many of the age-

related changes to the way one looks have already manifested themselves. Finally, three domains reported SRMs that indicate no change has occurred. Domain 6 (Leisure Activities) and domain 7 (Fears) are represented by, respectively, two and four items in the questionnaire

**Table 5** Exploration of significant differences between SRMs

Hypothesis	Intercept	<i>p</i> value	Interpretation	Larger SRM
The SRMs of SarQoL Overall score and SF-36 PCS score are significantly different	− 0.326	0.005	Different	SarQoL
The SRMs of SarQoL Overall score and SF-36 MCS score are significantly different	− 0.236	0.150	Not different	None
The SRMs of SarQoL Overall score and EQ-5D Utility Index are significantly different	− 0.724	< 0.001	Different	SarQoL
The SRMs of SarQoL Overall score and EQ-VAS are significantly different	− 0.443	0.003	Not different	SarQoL

Calculation of *p* values carried out with modified jack-knife method [30]

and may be much less sensitive than domains with more items. For domain 2 (Locomotion), this reasoning does not apply. This domain asks pointed questions connected to walking (length, frequency, difficulties, tiredness, etc.), and given that the usual gait speed has significantly diminished, one would expect to see an effect in this domain. However, the questions in this domain may be affected by the phenomenon of response shift, whereby the internal standards of measurement of the subject are recalibrated.

The SF-36 reported moderate change for the domain Physical Functioning, and small change for the domains Social Functioning, Role limitations due to emotional problems, Mental Health, and General Health, and reported no change for the other domains. These results are in line with our hypothesis that the SarQoL<sup>®</sup> questionnaire, being specific to sarcopenia, should detect a greater change than generic questionnaires such as the SF-36. The EQ-5D reported a small change for the domains Autonomy and Usual Activities and no change for all other scores. This should not be surprising given the distance between the response options for the EQ-5D, which means a significant change needs to occur in real life for it to be registered in the change scores.

Lastly, the correlations between changes in physical parameters and the changes on the different overall/composite scores revealed three good correlations for the SarQoL<sup>®</sup> Overall

score, one good and two acceptable correlations for the SF-36 PCS, one acceptable correlation for the SF-36 MCS, no correlations for the EQ-5D Utility Score and one good correlation for the EQ-VAS. In general, the SarQoL<sup>®</sup> Overall score correlates well with physical performance, with good correlations for change in gait speed, SPPB and CST. However, these results should be interpreted with caution given the multidimensional nature of sarcopenia, which is unlikely to be covered in a single test.

This study has several strengths. The methodology we adopted supplied us with evidence from different sources and allowed us to show both the quality and quantity of responsiveness. We were able to draw upon the data collected within the SarcoPhAge study, which allowed us to have a moderate sample size ( $n = 42$ ) despite the relatively low prevalence of sarcopenia. Furthermore, the SarcoPhAge study collected muscle mass data with DXA, which is, in practice, the most reliable method, and collected data on a number of tests for physical performance, which allowed us to compare the changes on several physical parameters [35].

There are, however, several limitations in this study. The SarcoPhAge study was not specifically designed to allow the evaluation of the responsiveness of the SarQoL<sup>®</sup> questionnaire, lacking both a known intervention and a transition question. A second limitation is that the primary methodology used in this study, the testing of hypotheses, has only been

**Table 6** Correlations between changes in physical parameters and evolution of quality of life

Domains	<i>r</i>	<i>p</i> value	Interpretation
1. $\Delta$ Gait speed and $\Delta$ SarQoL Overall	0.50	0.001	Good
2. $\Delta$ Gait speed and $\Delta$ SF-36 PCS	0.39	0.017	Acceptable
3. $\Delta$ Gait speed and $\Delta$ SF-36 MCS	0.02	0.926	NS
4. $\Delta$ Gait speed and $\Delta$ EQ-5D Utility Index	− 0.09	0.560	NS
5. $\Delta$ Gait speed and $\Delta$ EQ-VAS	0.16	0.324	NS
6. $\Delta$ Grip strength and $\Delta$ SarQoL Overall	0.08	0.592	NS
7. $\Delta$ Grip strength and $\Delta$ SF-36 PCS	0.27	0.104	NS
8. $\Delta$ Grip strength and $\Delta$ SF-36 MCS	− 0.14	0.393	NS
9. $\Delta$ Grip strength and $\Delta$ EQ-5D Utility Index	0.22	0.165	NS
10. $\Delta$ Grip strength and $\Delta$ EQ-VAS	0.08	0.626	NS
11. $\Delta$ SPPB and $\Delta$ SarQoL Overall	0.47	0.002	Good
12. $\Delta$ SPPB and $\Delta$ SF-36 PCS	0.30	0.068	NS
13. $\Delta$ SPPB and $\Delta$ SF-36 MCS	0.25	0.131	NS
14. $\Delta$ SPPB and $\Delta$ EQ-5D Utility Index	0.12	0.450	NS
15. $\Delta$ SPPB and $\Delta$ EQ-VAS	0.12	0.450	NS
16. $\Delta$ ALM and $\Delta$ SarQoL Overall	0.15	0.355	NS
17. $\Delta$ ALM and $\Delta$ SF-36 PCS	0.04	0.829	NS
18. $\Delta$ ALM and $\Delta$ SF-36 MCS	0.19	0.264	NS
19. $\Delta$ ALM and $\Delta$ EQ-5D Utility Index	0.03	0.832	NS
20. $\Delta$ ALM and $\Delta$ EQ-VAS	< − 0.01	0.986	NS
21. $\Delta$ ALM/BMI and $\Delta$ SarQoL Overall	− 0.02	0.901	NS
22. $\Delta$ ALM/BMI and $\Delta$ SF-36 PCS	− 0.14	0.807	NS
23. $\Delta$ ALM/BMI and $\Delta$ SF-36 MCS	− 0.11	0.537	NS
24. $\Delta$ ALM/BMI and $\Delta$ EQ-5D Utility Index	− 0.06	0.726	NS
25. $\Delta$ ALM/BMI and $\Delta$ EQ-VAS	− 0.48	0.002	Good
26. $\Delta$ ALM/Ht <sup>2</sup> and $\Delta$ SarQoL Overall	0.11	0.477	NS
27. $\Delta$ ALM/Ht <sup>2</sup> and $\Delta$ SF-36 PCS	0.10	0.570	NS
28. $\Delta$ ALM/Ht <sup>2</sup> and $\Delta$ SF-36 MCS	0.22	0.192	NS
29. $\Delta$ ALM/Ht <sup>2</sup> and $\Delta$ EQ-5D Utility Index	0.01	0.964	NS
30. $\Delta$ ALM/Ht <sup>2</sup> and $\Delta$ EQ-VAS	< 0.01	0.989	NS
31. $\Delta$ TUG and $\Delta$ SarQoL Overall	− 0.17	0.279	NS
32. $\Delta$ TUG and $\Delta$ SF-36 PCS	− 0.44	0.007	Good
33. $\Delta$ TUG and $\Delta$ SF-36 MCS	− 0.23	0.174	NS

**Table 6** continued

Domains	<i>r</i>	<i>p</i> value	Interpretation
34. $\Delta$ TUG and $\Delta$ EQ-5D Utility Index	– 0.02	0.923	NS
35. $\Delta$ TUG and $\Delta$ EQ-VAS	– 0.02	0.882	NS
36. $\Delta$ CST and $\Delta$ SarQoL Overall	– 0.42	0.013	Good
37. $\Delta$ CST and $\Delta$ SF-36 PCS	– 0.37	0.032	Acceptable
38. $\Delta$ CST and $\Delta$ SF-36 MCS	– 0.36	0.040	Acceptable
39. $\Delta$ CST and $\Delta$ EQ-5D Utility Index	– 0.13	0.470	NS
40. $\Delta$ CST and $\Delta$ EQ-VAS	– 0.11	0.546	NS

$\Delta$  = change in; *r* = correlation

NS not significant, *SPPB* Short Physical Performance Battery, *ALM* appendicular lean mass, *ALM/BMI* ALM divided by body mass index, *ALM/H<sup>2</sup>* ALM divided by height squared, *TUG* timed up-and-go test, *CST* chair stand test

introduced a few years ago and that several questions about this process have not yet found a consensus, such as how many hypotheses should be tested, what percentage should be confirmed for good responsiveness and how to set the strength of the expected correlations. We have tried to address these issues by using pre-defined, specific and challenging hypotheses but recognize that this methodology should be considered an ongoing process and hope that other studies can re-evaluate our hypotheses and add their own. Lastly, the SF-36 PCS and MCS scores were used in the evaluation of the SRMs but not in the hypotheses. We acknowledge that the PCS and MCS would have made good targets for the formulation of hypotheses, but unfortunately, the choice to calculate these scores was made after the hypotheses were formulated and after the statistical manipulations had started. It was therefore impossible for us to include the PCS and MCS scores in the hypotheses. It is our hope that future responsiveness studies will include the PCS and MCS in their hypotheses.

## CONCLUSIONS

This study contributed data on the last major psychometric property of the SarQoL<sup>®</sup> questionnaire not yet studied. The questionnaire has good responsiveness, measured both in an

evaluation with hypotheses (8/9 confirmed) and by the strength of its standardized response means. The SarQoL<sup>®</sup> questionnaire appears to be the optimal tool for the assessment of quality of life in sarcopenic populations. Its use in clinical trials assessing biochemical entities for the management of sarcopenia should be recommended, as patient-related outcomes are encouraged to be included as co-primary endpoints in such studies [36].

## ACKNOWLEDGEMENTS

We thank the participants of the SarcoPhAge study, whose continued participation made this article possible.

**Funding.** No funding or sponsorship was received for this study.

**Authorship.** All named authors meet the International Committee of Medical Journal Editors (ICMJE) criteria for authorship for this article, take responsibility for the integrity of the work as a whole, and have given their approval for this version to be published.

**Authorship Contributions.** AG, CB, OB and JYR developed the study design. CB and ML collected the data. AG was responsible for data analysis and drafted the article. All authors

provided revision of the draft article, and read and approved the final manuscript.

**Prior Presentation.** The data in this article have been presented as a poster at the ICFSR 2018 congress (Miami, USA) and the WCO-IOF-ESCEO 2018 congress (Krakow, Poland).

**Disclosures.** OB, CB and J-YR are shareholders of SarQoL<sup>®</sup> sprl. J-YR is the president of the European Society for Clinical and Economic Aspects of Osteoporosis, Osteoarthritis and Musculoskeletal Diseases (ESCEO), which has endorsed the SarQoL<sup>®</sup> questionnaire. AG and ML declare they have no conflicts of interest.

**Compliance with Ethics Guidelines.** Participants in the SarcoPhAge study all provided written informed consent. The research protocol and its amendments were approved by the Ethics Committee of the University Teaching Hospital of Liège (no. 2012-277).

**Data Availability.** The dataset used for this study is available from the corresponding author on reasonable request.

**Open Access.** This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## REFERENCES

- Cruz-Jentoft AJ, Baeyens JP, Bauer JM, et al. Sarcopenia: European consensus on definition and diagnosis. *Age Ageing*. 2010;39:412–23.
- Cao L, Morley JE. Sarcopenia is recognized as an independent condition by an International Classification of Disease, tenth revision, clinical modification (ICD-10-CM) code. *J Am Med Dir Assoc*. 2016;17:675–7.
- Beaudart C, Zaaria M, Pasleau F, Reginster J-Y, Bruyère O. Health outcomes of sarcopenia: a systematic review and meta-analysis. *PLoS One*. 2017;12:e0169548.
- Chang K-V, Hsu T-H, Wu W-T, Huang K-C, Han D-S. Is sarcopenia associated with depression? A systematic review and meta-analysis of observational studies. *Age Ageing*. 2017;46:738–46.
- Beaudart C, Reginster J-Y, Geerinck A, Locquet M, Bruyère O. Current review of the SarQoL<sup>®</sup>: a health-related quality of life questionnaire specific to sarcopenia. *Expert Rev Pharmacoecon Outcomes Res*. 2017;17:335–41.
- Rizzoli R, Reginster JY, Arnal JF, et al. Quality of life in sarcopenia and frailty. *Calcif Tissue Int*. 2013;93:101–20.
- Beaudart C, Biver E, Reginster J-Y, et al. Development of a self-administrated quality of life questionnaire for sarcopenia in elderly subjects: the SarQoL. *Age Ageing*. 2015;44:960–6.
- Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63:737–45.
- Guyatt GH, Kirshner B, Jaeschke R. Measuring health status: what are the necessary measurement properties? *J Clin Epidemiol*. 1992;45:1341–5.
- Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chronic Dis*. 1987;40:171–8.
- Beaudart C, Biver E, Reginster J-Y, Rizzoli R, Rolland Y, Bautmans I, et al. Validation of the SarQoL, a specific health-related quality of life questionnaire for sarcopenia. *J Cachexia Sarcopenia Muscle*. 2017;8:238–44.
- Beaudart C, Edwards M, Moss C, et al. English translation and validation of the SarQoL<sup>®</sup>, a quality of life questionnaire specific for sarcopenia. *Age Ageing*. 2017;46:271–7.
- Ildiko GA, Mihai G, Beaudart C, et al. Psychometric performance of the Romanian version of the SarQoL<sup>®</sup>, a health-related quality of life questionnaire for sarcopenia. *Arch Osteoporos*. 2017;12:103.
- Geerinck A, Scheppers A, Beaudart C, et al. Translation and validation of the Dutch SarQoL<sup>®</sup>, a quality of life questionnaire specific to sarcopenia. *J Musculoskelet Neuronal Interact*. 2018. [http://www.ismni.org/jmni/accepted/jmni\\_aa\\_BAUTMANS.pdf](http://www.ismni.org/jmni/accepted/jmni_aa_BAUTMANS.pdf).



15. Beaudart C, Reginster J-Y, Petermans J, et al. Quality of life and physical components linked to sarcopenia: the SarcoPhAge study. *Exp Gerontol*. 2015;69:103–10.
16. Locquet M, Beaudart C, Bruyère O, Kanis JA, Delandsheere L, Reginster J-Y. Bone health assessment in older people with or without muscle health impairment. *Osteoporos Int*. 2018;29:1057–67.
17. Beaudart C, Reginster J-Y, Petermans J, Bruyère O. Qualité de vie du patient sarcopénique: apport de l'étude liégeoise SarcoPhAge. *Geriatr Psychol Neuropsychiatr Vieil*. 2015;13:391–5.
18. Locquet M, Beaudart C, Reginster J-Y, et al. Prevalence of concomitant bone and muscle wasting in elderly women from the SarcoPhAge cohort: preliminary results. *J Frailty Aging*. 2017;6:18–23.
19. The SarQoL questionnaire [Internet]. <http://www.sarqol.org>. Accessed 7 Sep 2018.
20. Ware JJ, Sherbourne C. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care*. 1992;30:473–83.
21. Ware JJ, Kosinski M. SF-36 physical and mental health summary scales: a manual for users of version 1. Lincoln: QualityMetric; 2001.
22. Laucis NC, Hays RD, Bhattacharyya T. Scoring the SF-36 in orthopaedics: a brief guide. *J Bone Jt Surg*. 2015;97:1628–34.
23. Van Reenen M, Oppe M. EQ-5D-3L User guide. Basic information on how to use the EQ-5D-3L instrument. 2015. EuroQol Group. [https://euroqol.org/wp-content/uploads/2016/09/EQ-5D-3L\\_User\\_Guide\\_2015.pdf](https://euroqol.org/wp-content/uploads/2016/09/EQ-5D-3L_User_Guide_2015.pdf).
24. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med*. 2001;33:337–43.
25. De Vet HCW, Knol DL, Terwee CB, Mokkink LB. *Measurements in medicine*. Cambridge: Cambridge University Press; 2011.
26. De Boer M, Moll A, de Vet H, Terwee CB, Völker-Dieben HJM, van Rens GHMB. Psychometric properties of vision-related quality of life questionnaires: a systematic review. *Ophthalmic Physiol Opt*. 2004;24:257–73.
27. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. *J Clin Epidemiol*. 2000;53:459–68.
28. Middel B, Van Sonderen E. Statistical significant change versus relevant or important change in (quasi) experimental design: some conceptual and methodological problems in estimating magnitude of intervention-related change in health services research. *Int J Integr Care*. 2002;2:e15.
29. Cohen J. Quantitative methods in psychology. *Psychol Bull*. 1992;112:155–9.
30. Bessette L, Sangha O, Kuntz K, et al. Comparative responsiveness of generic versus disease-specific and weighted versus unweighted health status measures in carpal tunnel syndrome. *Med Care*. 1998;36:491–502.
31. Angst F, Verra ML, Lehmann S, Aeschlimann A. Responsiveness of five condition-specific and generic outcome assessment instruments for chronic pain. *BMC Med Res Methodol*. 2008;8:26.
32. Evans JD. *Straightforward statistics for the behavioral sciences*. Pacific Grove: Brooks/Cole; 1996.
33. Faul F, Erdfelder E, Lang A-G, Buchner A. G\*Power: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods*. 2007;39:175–91.
34. Terwee CB, Mokkink LB, Knol DL, Ostelo RWJG, Bouter LM, De Vet HCW. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res*. 2012;21:651–7.
35. Buckinx F, Landi F, Cesari M, et al. Pitfalls in the measurement of muscle mass: a need for a reference standard. *J Cachexia Sarcopenia Muscle*. 2018;9:269–78.
36. Reginster J-Y, Cooper C, Rizzoli R, et al. Recommendations for the conduct of clinical trials for drugs to treat or prevent sarcopenia. *Aging Clin Exp Res*. 2016;28:47–58.