Sleep Breath (2011) 15:549–559
DOI 10.1007/s11325-010-0399-3

ORIGINAL ARTICLE

Visual analogical well-being scale for sleep apnea patients: validity and responsiveness

A test for clinical practice

Spanish Group of Breathing Sleep Disorders
Juan F. Masa · Antonio Jiménez · Joaquín Durán · Carmen Carmona · Carmen Monasterio · Mercedes Mayos · Joaquín Terán · Ferrán Barbé · Manuela Rubio · Inmaculada Failde · Manuel Mota · José M. Montserrat

Received: 21 April 2010 / Revised: 21 June 2010 / Accepted: 18 July 2010 / Published online: 5 August 2010

Abstract
Introduction Health-related quality-of-life (HRQL) tests used in sleep apnea–hypopnea syndrome (SAHS) are time-consuming, complicating their application in clinical practice. The objective was to examine the validity and responsiveness of a simple visual analogical well-being scale (VAWS) for the clinical use.

Method The subjects proceed from a cohort of SAHS patients treated with CPAP for 12 weeks. We correlated the VAWS with other HRQL tests, related clinical and polysomnographic measures to concurrent and construct validities. Responsiveness by: (1) comparison of HRQL tests between before and after treatment and *effect size*. (2) Association of the change with treatment between

J. F. Masa (✉)
San Pedro de Alcántara Hospital,
C/Rafael Alberti 12,
10005 Cáceres, Spain
e-mail: fmasa@separ.es

A. Jiménez
Hospital de Valdecilla,
Santander, Spain

J. Durán
Txagorritxu Hospital,
Vitoria-Gasteiz, Spain

C. Carmona
Virgen del Rocío Hospital,
Sevilla, Spain

C. Monasterio
Bellvitge Hospital,
Barcelona, Spain

M. Mayos
Sant Pau Hospital,
Barcelona, Spain

J. Terán
General Yagüe Hospital,
Burgos, Spain

F. Barbé
Arnau de Vilanova Hospital,
Lerida, Spain

M. Rubio
San Jorge Primary Care Center,
Cáceres, Spain

I. Failde
Cadiz University,
Cádiz, Spain

M. Mota
Extremadura University,
Badajoz, Spain

J. M. Montserrat
Clinic Hospital,
Barcelona, Spain

J. F. Masa · J. Durán · J. Terán · F. Barbé · M. Rubio · J. M. Montserrat
Centro de Investigación Biomédica en Red (CIBER) de enfermedades respiratorias (CIBERES),
Madrid, Spain

VAWS with other HRQL tests and between VAWS with clinical parameters.
Results At baseline, VAWS correlated with all HRQL tests but better with functional outcomes in sleep questionnaires (FOSQ) and European quality-of-life questionnaire (EuroQol) thermometer. VAWS and FOSQ correlated better with clinical variables than other HRQL tests. VAWS captures the magnitude change with treatment similarly to FOSQ but better than other HRQL tests.
Conclusion VAWS is a very simple test which measures HRQL in SAHS. It could be a useful tool in clinical practice, primarily for the responsiveness of treatment.

Keywords CPAP treatment · Sleep apnea syndrome · Quality of life

Introduction

Sleep apnea–hypopnea syndrome (SAHS) affects between 2% and 4% of the adult population [1], and is characterized by repetitive obstructions in the upper airway. The consequences of these repetitive episodes are daytime sleepiness, tiredness, traffic accidents [2, 3], increased cardiovascular risk [4–6], and even mortality [7]. Continuous positive airway pressure (CPAP) is the most effective treatment for patients with symptomatic sleep apnea syndrome [8, 9].

Most of the consequences of sleep apnea impact directly on health-related quality of life (HRQL). In fact, a number of randomized controlled studies have shown impairment in HRQL and improvement using CPAP in non-specific and specific questionnaires of HRQL [10–12].

On the other hand, the Epworth sleepiness scale (ESS) [13] is a test frequently used to evaluate daytime sleepiness in patients with suspected SAHS and to ascertain response to CPAP treatment. Although daytime sleepiness is a cardinal symptom in SAHS patients, its evaluation cannot surrogate all the dimensions which affect HRQL [14].

The HRQL tests used in SAHS are often long and time-consuming, which does not facilitate application in clinical practice. However, there are two exceptions: the EuroQol thermometer—one of two independent parts of the EuroQol Test (EQ)—[15], and a verbal analogical scale [16]. These exceptions are analogical scales that explore the subjective perception of non-specific HRQL. The EQ thermometer has been evaluated in some studies involving sleep apnea patients with conflicting results [10, 17, 18]. The verbal analogical scale has been assessed in only one short study with poor results [16].

At least in theory, specific questionnaires for SAHS could better detect changes in HRQL than non-specific HRQL [19, 20]. The specific ones have been designed to evaluate patients with SAHS and they include some queries related to SAHS, for example sleepiness. Four specific questionnaires for HRQL have been proposed for the adult population [21–24] but they are not simple, short and easily self-administered tests for clinical practice.

We sought to examine the validity and responsiveness of a new simple self-administered visual analogical well-being scale (VAWS) in SAHS patients.

Methods

Patients

The subjects belonged to cohort of SAHS patients aged between 18 and 70 years consecutively recruited from ten sleep centers where CPAP was indicated [9, 25]. The exclusion criteria were as follows: psychophysical incapacity to perform questionnaires, patients with chronic disease (cancer, chronic pain, renal failure, moderate or severe chronic obstructive pulmonary disease etc.), drugs or alcohol addiction, Cheyne–Stokes syndrome, life-threatening SAHS, patients with previous uvulopalatopharyngoplasty (UPPP), absence of a partner at home, important chronic nasal obstruction, lack of skill in adjusting the nasal mask and refusal to participate in the study. Fixed CPAP pressure was obtained randomly by three different methods of CPAP titration, by standard polysomnography, by self-adjusted CPAP device at home and by predicted formula with domiciliary adjustment. The baseline and CPAP polysomnographic studies were analyzed manually according to standard criteria [26, 27]. The study was approved by the ethics committees of the 10 centers. Informed consent was obtained from all the patients.

Study protocol

The protocol has been previously published [9]. Patients with suspicion of SAHS were referred for polysomnography. Those meeting the criteria of CPAP treatment [25] were randomized in three CPAP titration groups, (standard polysomnography group, self-adjusted CPAP device group, and predicted formula group) to determine the fixed CPAP [9]. In the predicted formula group patient pressure was initially determined on the basis of a predicted formula and was adjusted on the second and third controls if snoring or apneas were observed by the partner. Patient outcomes (ESS, HRQL test, snoring and apneas observed, secondary effects and CPAP compliance) were evaluated at baseline and after 12 weeks of CPAP treatment. After this 3-month period of treatment, patients underwent a second polysomnography with their fixed CPAP to verify the disappearance of the respiratory events.

Outcomes

Visual analogical well-being scale

The VAWS was conceived [9] as a tool for use in clinical practice. It consists of a 120-mm straight line on which the patient indicates his or her health status with respect to the symptoms which were the motive for the consultation (e.g., SAHS). The ends of the line indicate the least favorable and the most favorable well-being status (see Annex 1). Once the patient marked his/her point, we measured the distance (in mm) from the least favorable position and transformed it into percentages. The time employed is less than 1 min.

Medical outcome survey—short form 36 (SF 36) [19, 28, 29]

It is a non-specific self-administered questionnaire that assesses eight dimensions of HRQL: physical functioning, limitation caused by physical problems, handicap due to emotional problems, social functioning, mental health, energy and vitality, pain and general health perception. These eight items are usually reduced to two, physical and mental dimensions. Values in the general population are around 80. Lower scores reflect poorer HRQL.

Functional outcomes in sleep questionnaires [22]

This consists of a 30-item self-report questionnaire designed to measure the impact of excessive sleepiness on multiple activities of daily living. It assesses five dimensions: activity level, vigilance, intimacy and sexual relationships, general productivity, and social outcomes. Values in the general population are around 110. Lower scores reflect poorer HRQL.

European quality-of-life questionnaire—EuroQol 5 D (EQ 5D) [15]

This is a non-specific self-administered HRQL questionnaire which measures five areas of health: mobility, self-care, pain/discomfort, usual activities and anxiety/depression. Values in the general population are around 85. Lower scores reflect poorer HRQL.

European quality-of-life questionnaire—EuroQol thermometer (EQ thermometer) [15]

The EuroQol also adds a linear visual analogical scale to assess the general health situation (0=the worst imaginable health to 100=the best imaginable health). Values in the general population are around 84.

Epworth sleepiness scale [13]

This is a self-administered questionnaire that measures the likelihood of falling asleep in eight daily situations.

American sleep disorders association sleepiness [30]

Group sleepiness into four categories of intensity according to interference with daily life (none, mild, moderate, and severe).

Clinical variables Restlessness and snoring in four degrees of intensity: never, sometimes, frequently and always.

Polysomnographic parameters such as AHI, sleep efficiency, arousals index, light and deep sleep.

Other additional outcomes measured are shown in Table 1. These include anthropometric data, questionnaires concerning personal habits, tobacco smoking, alcohol intake, drug consumption, education, apneas observed by the partner, work, number of hours of sleep per night and forced spirometry [31].

Statistical analysis

We included the entire sample to analyze the validity at baseline and the evaluative properties (responsiveness). This sample was also divided into three titration groups (which were equally effective [9]) to analyze responsiveness after CPAP treatment.

Characteristics of the patients

At baseline, we compared the clinical and polysomnographic characteristics between patients who withdrew from the study and patients who were available for follow-up using t test (normal distribution) or Kruskal–Wallis tests (non-normal distribution) for continuous variables and χ^2 test for qualitative variables.

"Ceiling and floor effect"

We analyzed the distribution of scores at baseline to calculate the percentage of patients with the highest and the lowest scores.

Test–retest reliability and agreement

At the end of the study, we carried out an additional study including 81 patients with the same inclusion and exclusion criteria and protocol as the present one with the same sleep centers (see Annex 2). Test–retest reliability was analyzed between a VAWS, achieved at diagnosis time, and another one 3 weeks later before CPAP treatment, by using intraclass correlation coefficient (ICC) [32].

Table 1 Clinical and polysomnographic characteristics of patients who abandoned and completed the study protocol

	Abandoned protocol *N*=45	Completed protocol *N*=315	*p* value
Age, years (mean and SD)	48.6 (10.1)	51.4 (9.7)	ns
Gender,			
% Men	91.1	88.6	ns
% Women	8.9	11.4	ns
BMI, kg. m^{-2} (mean and SD)	33.4 (5.3)	33.5 (7.1)	ns
Active worker,%	88.6	79.0	ns
Primary school only,%	53.3.1	56.2	ns
Alcohol, g/day (mean and SD)	22.8 (26.3)	24.5 (23.8)	ns
Active smoker,%	55.0	40.8	ns
Blood hypertension,%	50.0	57.8	ns
FEV1,% predicted (mean and SD)	94.3 (18.5)	94.4 (18.2)	ns
Sleep, hours per night (mean and SD)	6.9 (1.4)	7.0 (1.6)	ns
ESS (mean and SD)	15.5 (4.5)	15.7 (3.5)	ns
ASDA sleepiness,% of severe	21.4	37.5	ns
Habitual snoring,%	83.7	90.2	ns
Apneas observed,%	46.5	60.1	0.03
AHI (mean and SD)	58.3 (23.0)	62.7 (22.9)	ns
Sat O2 below 90%,%	23.4 (23.9)	28.5 (26.7)	ns
Sleep efficiency	82.0 (12.8)	78.9 (12.9)	ns
Arousals index	50.2/22.4)	56.6 (20.2)	ns
Light sleep	74.9 (15.5)	77.6 (12.3)	ns
Deep sleep	9.3 (8.9)	8.2 (9.1)	ns

ESS Epworth sleepiness scale, *ASDA* ASDA sleepiness scale, *AHI* apnea and hypopnea index

Validity

Cross-sectional validity at baseline was examined in two ways: concurrent criterion and construct validities. We correlated the VAWS with HRQL tests to concurrent validity and the VAWS with other related clinical and polysomnographic measures to construct validity.

Evaluative properties

To determine whether the VAWS achieved changes in HRQL after treatment intervention (responsiveness), the following analyses were performed:

1. Comparison of the values of the HRQL test before and after CPAP treatment using the *paired t test* (or equivalent non-parametric) and *effect size* calculation [29].
2. Determination of the ability to change in patients with or without adequate CPAP treatment (<2 h/day, between 2 and 4 h/day, and >4 h/day).
3. Longitudinal validity:
   (a) Concurrent: correlation of the changes with CPAP treatment between the VAWS and HRQL questionnaires.
   (b) Construct: comparison of the change with CPAP treatment between the VAWS and HRQL tests with clinical variables. We made variables with differences before and after treatment in qualitative variables (distributed in 4 degrees of intensity: never, sometimes, frequently, and always). We classified the results into four groups of intensity to achieve a more homogeneous distribution of change: (a) non improvement or worsening; (b) light improvement; (c) moderate improvement; and (d) significant improvement. We compared the means of the differences in the VAWS and HRQL tests (dependent variables) with the change in the 4 groups of intensity using qualitative variables (factor), by ANOVA (normal distribution) or Kruskal–Wallis (non-normal distribution). Statistical significance indicated the presence and the level of association.
   (c) We also measured the improvement in VAWS and HRQL tests by comparison among percentiles of Epworth sleepiness scale (ESS) improvement to determine the level of association. Then, we compared the improvement in ESS value among VAWS percentiles to examine the level of improvement in VAWS which resulted in significant

change in ESS. We initially used one-way ANOVA. Where appropriate, differences between individual means were tested using the least significant difference (SPSS 14.0; SPSS Inc., Chicago, IL, USA). If the variables were not distributed normally, a non-parametric test was used (Kruskal–Wallis and Dunn post hoc to identify differences between individual means).

For the correlation analysis in cross-sectional and longitudinal validity, we applied Pearson's test to determine the magnitude and direction of the correlation provided that the continuous variables were normally distributed. Otherwise Spearman's test was used.

Results

A total of 466 patients were initially evaluated and 106 were excluded for the following reasons: chronic disease 40 (8.6%), severe nasal obstruction 13 (2.8%), refusal to participate in the study protocol 12 (2.6%), psychophysical inability to answer the questionnaires, ten (2.1%), absence of partner at home, ten (2.1%), alcohol addiction, nine (1.9%), previous UPPP, six (1.3%), lack of skill in adjusting the nasal mask, five (1.1%), and life-threatening SAHS, one (0.2%).

Out of 360 patients finally included, 45 (12.5%) abandoned the study, 18 (5%) did not tolerate the treatment, eight (2.2%) underwent titration failure, 13 (3.6%) were lost in the follow-up period, and six (1.7%) were lost other motives.

The general characteristics of the subjects who abandoned and those that completed the study protocol are shown in Table 1. Only the variable (apneas observed by the partner) was more frequent in patients who completed the protocol. The efficacy, use and secondary effects of CPAP treatment were similar in the CPAP groups in a previous report [9].

The *ceiling and floor* effect was low because the percentage of subjects with a higher score was 1.9% and the percentage of subjects with a lower score was 0.3%.

Test–retest reliability of the VAWS was good since ICC was 0.83.

Validity at baseline

The correlations at baseline of the VAWS and HRQL tests to assess concurrent validity are shown in Table 2. The VAWS was correlated with all the HRQL tests. The VAWS correlated better with the EuroQol thermometer and with the functional outcomes in sleep questionnaires (FOSQ). Good correlation was observed with FOSQ dimensions (activity=0.412, vigilance=0.345, general productivity=0.328, social outcome=0.337; $p<0.001$) except with intimacy and sexual relation (0.142; $p<0.05$). The non-specifics HRQL tests (except EuroQol Thermometer) had higher correlations with the FOSQ than with VAWS.

As regards the construct validity (Table 2), higher correlations were observed between clinical variables with VAWS and FOSQ than between clinical variables with other HRQL tests. In contrast to the FOSQ, the VAWS adds a weak but statistically significant correlation to related polysomnographic variables (such as sleep efficiency, arousals index, light and deep sleep). No statistically significant correlations were observed between the apnea and hypopnea index with VAWS and HRQL tests.

Table 2 Correlations at baseline among VAWS, HRQL tests and clinical and polysomnographic related variables

	VAWS	FOSQ-G	SF36-P	SF36-M	EQ-5D	EQ-T
FOSQ-G	0.402[a]					
SF36-P	0.278[a]	0.532[a]				
SF36-M	0.315[a]	0.542[a]	–			
EQ-5D	0.292[a]	0.523[a]	0.585[a]	0.488[a]		
EQ-T	0.521[a]	0.482[a]	0.513[a]	0.427[a]	0.499[a]	–
ESS	−0.330[a]	−0.381[a]	−0.242[a]	−0.093	−0.204[a]	−0.097
ASDA	−0.291[a]	−0.316[a]	−0.216[a]	−0.132[c]	−0.189[b]	−0.148[b]
Restlessness	−0.373[a]	−0.389[a]	−0.326[a]	−0.361[a]	−0.392[a]	−0.317[a]
Snoring	−0.158[b]	−0.175[b]	−0.096	−0.054	−0.124[c]	−0.088
AHI	−0.074	−0.039	−0.01	0.064	−0.010	−0.038
Sleep efficiency	0.132[c]	0.063	0.067	0.020	0.112[c]	0.099
Arousals index	−0.142[c]	−0.088	−0.069	0.011	−0.047	−0.077
Light sleep	−0.131[c]	−0.079	−0.065	0.048	−0.024	−0.080
Deep sleep	0.146[b]	0.029	0.002	−0.041	−0.016	0.057

FOSQ-G FOSQ global punctuation, *SF36-P* SF 36 Physical, *SF·&-M* SF 36 Mental, *EQ-5D* EuroQuol 5D, *EQ-T* EuroQuol Thermometer, *ESS* Epworth sleepiness scale. *ASDA* ASDA sleepiness. *AHI* apnea and hypopnea index

[a] p value <0.001

[b] p value <0.01

[c] p value <0.05

Evaluative properties

The ability of the VAWS to identify changes with CPAP treatment is shown in Table 3. Statistical improvement after treatment occurred with the VAWS and HRQL tests but the magnitude of the change evaluated by the *effect size* was higher in the VAWS and the FOSQ. The difference before and after treatment and effect size in the three CPAP groups was slightly lower in the self-adjusted group than in the other two CPAP groups (see "Discussion" section).

The ability of the VAWS and HRQL tests to distinguish between patients with and without adequate CPAP treatment is shown in Table 4. In contrast to HRQL tests, the difference in the VAWS before and after treatment and the *effect size* increased progressively according to the treatment use, as it should be expected (lower values in incorrectly treated, moderate in partially treated and higher in correctly treated patients). In this last group, the 95% confidential interval of the difference between before and after treatment ranges from 17 to 23.

The analysis to evaluate the longitudinal validity is shown in Tables 5 and 6 and Fig. 1. The correlations in rating of change among the VAWS and HRQL tests can be observed in Table 5. The VAWS correlated better with the EuroQol thermometer and the FOSQ. Significant correlation was observed with FOSQ dimensions (activity=0.377, vigilance=0.271, general productivity=0.215, social outcome=0.242; $p<0.001$) except with intimacy and sexual relation (0.117; $p<0.05$). The non-specifics HRQL tests (except SF 36 Physical) had similar or lower correlations with the FOSQ and/than with VAWS.

The change in the VAWS proved to be better associated than HRQL tests as far as the improvement in clinical variables was concerned (Table 6). The FOSQ resembles the VAWS in some variables such as restlessness and American sleep disorders association (ASDA) sleepiness. All HRQL tests improved according to the improvement in EES percentiles, but specially the FOSQ and the VAWS (Fig. 1a). Significant improvement in EES value occurred on comparing the first and second percentiles of VAWS with a mean of change of 20 between percentiles (Fig. 1b).

Discussion

This study was designed to validate a simple and rapid well-being test for its use in clinical practice. To our knowledge, this study about HRQL is one with the largest number of patients suffering SAHS treated with CPAP. The main results were as follows:

1. At baseline the VAWS correlated with all the HRQL tests. The VAWS and the FOSQ correlated more favorably with clinical variables.
2. The VAWS captured changes in the HRQL with CPAP treatment in a way similar to the FOSQ but showed an improvement with respect to the other HRQL tests. The VAWS showed better ability than HRQL tests to distinguish between patients with and without adequate CPAP treatment and VAWS and FOSQ showed the best association with the improvement in ESS. A change on 20 in VAWS value with treatment can be clinically relevant.

At baseline the FOSQ had higher correlations than VAWS with non-specifics HRQL tests, except with EuroQol Thermometer. The absence of several scores in the VAWS could explain the lower correlations with some HRQL tests.

The EuroQol Thermometer resembled VAWS at base line but with worse association to clinical symptoms. However, the ability to identify changes with CPAP treatment and to distinguish patients with adequate CPAP compliance was lower with EuroQol Thermometer than VAWS. Regarding longitudinal validity, the rating change with treatment on clinical symptoms was better associated with VAWS than with EuroQol Thermometer. Supposedly,

Table 3 Pre and post-treatment values of VAWS (and three CPAP treatment groups with similar efficacy) and HRQL tests

	Pre-treatment Mean (SD)	Post-treatment Mean (SD)	*p* value	Effect size
VAWS	50.8 (22.2)	70.0 (19.7)	<0.001	0.87
Standard	48.3 (21.6)	69.6 (21.1)	<0.001	0.99
Autoadjusted	55.5 (22.4)	71.1 (18.6)	<0.001	0.69
Predicted formula	48.4 (22.0)	69.2 (19.4)	<0.001	0.94
FOSQ	88.2 (21.8)	106.4 (15.3)	<0.001	0.90
SF 36 physical	44.2 (8.8)	47.2 (7.9)	<0.001	0.36
SF 36 mental	46.2 (11.5)	49.9 (10.3)	<0.001	0.32
EuroQol 5D	75.0 (19.4)	81.9 (17.5)	<0.001	0.35
EuroQol thermometer	64.6 (19.3)	73.8 (15.5)	<0.001	0.46

Table 4 Rating of change in VAWS and HRQL tests in groups of patients with and without adequate CPAP treatment

	Incorrectly treated[a]		Partially treated[b]		Correctly treated[c]	
	N=21		*N*=56		*N*=238	
	Difference (95 CI)	Effect size	Difference (95 CI)	Effect size	Difference (95 CI)	Effect size
VAWS	13.0 (0.9 to 26.0)	0.60	17.4 (10.2 to 24.6)	0.79	20.2 (16.6 to 23.4)	0.92
FOSQ	17.3 (10.0 to 24.6)	0.85	13.6 (8.7 to 18.5)	0.67	19.6 (17.2 to 22.1)	0.96
SF physical	4.8 (1.0 to 8.0)	0.52	1.3 (-0.5 to 2.9)	0.14	3.4 (2.4 to 4.4)	0.39
SF mental	4.4 (-0.4 to 9.1)	0.37	5.0 (2.4 to 7.6)	0.43	3.3 (2.0 to 4.6)	0.29
EuroQuol 5D	10.6 (2.5 to 18.8)	0.54	7.2 (2.8 to 11.6)	0.37	6.5 (4.4 to 8.7)	0.33
EuroQuol Thermometer	10.9 (1.9 to 19.9)	0.57	7.9 (2.3 to 13.4)	0.39	9.1 (6.9 to 11.4)	0.47

CI confidence interval

[a] Compliance <2 h/day

[b] Compliance of 2–4 h/day

[c] Compliance >4 h/day

the EuroQol Thermometer could detect similar dimensions of HRQL than the VAWS but perhaps in a more general facet, since the first assesses well-being for general health situation and the second according to symptoms of suspected SAHS.

The intimacy and sexual relationship dimension of the FOSQ showed a weaker correlation with the VAWS than other FOSQ dimensions. Other studies have found poor association with the same dimension in SAHS subjects [22] and no change with CPAP treatment [33].

Even the highest levels of correlations among tests observed in the present study were moderate. It is not surprising because other studies showed similar correlations levels among HRQL tests and between HRQL tests and clinical symptoms [26, 29, 33]. Likely, the HRQL tests detect different dimensions into quality of life.

In the light of the foregoing discussion our large sample permits the detection of significant correlations to better characterize the VAWS with respect to HRQL tests. For instance, in contrast to HRQL tests, the VAWS showed at baseline weak but significant correlations (and in the expected direction) with snoring, sleep efficiency, arousals index and light and deep sleep. This lends support to the association of the VAWS measurement with SAHS. As expected according the results of previous studies [16], no HRQL tests were correlated with the apnea and hypopnea index. Although there is no clear explanation, it could be caused by a cognitive function defect or tolerance to the disease or its symptoms.

The effect size was slightly lower in the autoadjusted group than in the other two CPAP groups. A probable explanation was that the baseline values were higher in autoadjusted compared with the other groups (55±22 in autoadjusted, 48±22 in both standard and predicted formula groups; $p<0.05$) [9].

VAWS specifically measures “well being status”. This term is semantically different from HRQL but obviously much related. Since VAWS has only one item is improbable that it can have the same sensibility to detect multiple dimensions from HRQL as questionnaires with several items. Although the significant association pre-treatment among the VAWS with HRQL questionnaires (specifically

Table 5 Correlations in rating of change among HRQL tests

	VAWS	FOSQ Global punctuation	SF Physical	SF Mental	EuroQol 5D
FOSQ global punctuation	0.328				
SF 36 physical	0.220	0.351			
SF mental	0.262	0.284	–		
EuroQol 5D	0.252	0.200	0.326	0.157[a]	
EuroQol thermometer	0.458	0.370	0.315	0.213	0.252

[a] $p<0.01$, all the rest $p<0.001$

Table 6 Comparisons of rating of change (improvement) among HRQL tests and clinical related qualitative variables

	Restlessness					Snoring[a]				ASDA sleepiness			
	No/worsen	Light	Moderate	Significant	*P* value	No/light	Moderate	Significant	*P* value	No/worsen	Light	Moderate	Significant	*P* value
	N=68	*N*=71	*N*=81	*N*=94		*N*=14	*N*=62	*N*=336		*N*=45	*N*=109	*N*=109	*N*=49	
VAWS	10.7 (25.7)	14.6 (25.1)	19.5 (26.)	28.4 (30.1)	<0.001	−11.4 (26.8)	15.9 (24.0)	21.8 (27.7)	<0.001	7.0 (27.0)	14.3 (22.7)	23.7 (29.0)	31.2 (30.0)	<0.001
FOSQ	9.4 (10.7)	13.9 (14.4)	17.5 (19.8)	29.1 (20.6)	<0.001	15.3 (16.5)	16.1 (19.1)	19.1 (18.9)	ns	9.8 (15.9)	10.7 (13.6)	24.2 (19.0)	29.7 (20.5)	<0.001
SF Physical	1.0 (5.4)	2.1 (7.2)	2.3 (7.7)	5.9 (8.1)	<0.001	0.2 (6.1)	2.7 (7.1)	3.3 (7.7)	ns	2.0 (6.6)	1.3 (7.0)	4.1 (8.0)	5.4 (7.6)	<0.01
SF Mental	0.7 (8.4)	1.5 (8.3)	4.1 (11.1)	6.8 (10.3)	<0.01	−3.2 (9.7)	4.3 (10.9)	3.8 (9.7)	<0.05	2.2 (11.8)	1.7 (8.4)	5.1 (10.6)	5.9 (9.4)	<0.05
EuroQuol 5D	3.5 (17.1)	2.7 (13.8)	6.3 (16.0)	13.0 (17.3)	<0.001	0.5 (15.7)	3.5 (16.0)	8.1 (16.8)	ns	5.8 (16.0)	4.4 (15.0)	8.3 (18.2)	10.0 (16.9)	ns
EuroQuol Thermometer	3.1 (14.1)	7.2 (19.1)	12.0 (20.0)	12.1 (17.7)	<0.01	−6.0 (19.9)	9.6 (16.8)	9.7 (18.2)	<0.05	4.0 (18.9)	6.6 (16.6)	11.1 (18.7)	14.4 (18.9)	<0.01

Values showed as means and SD

No/worsen non improvement or worsening, *No/light* non improvement or light

[a] No patients worsen, therefore the improvement intensity with CPAP treatment was distributed in three categories

◀ **Fig. 1** **a** Improvement of VAWS and HRQL tests according to ESS percentiles; the level of improvement in the VAWS and the FOSQ was higher than the other HRQL tests. **b** Improvement of Epworth Sleepiness Scale (EES) according to VAWS percentiles; significant EES improvement occurs between first and second percentiles suggesting clinical significance

with the FOSQ) suggest that VAWS captures the main dimensions of HRQL in SAHS, especially about the responsiveness to CPAP treatment. So, we understand that the main use of VAWS should be to value the rate of change with treatment.

EES and VAWS values were associated at baseline and after improvement with treatment. An interesting question is whether VAWS could be a viable alternative to EES. Our study cannot demonstrate this, but VAWS values were better associated with HRQL tests than EES at baseline and after treatment response, although Epworth sleepiness scale probably detects daytime sleepiness better. Therefore, since VAWS is a quick test (less than a minute), our proposal is to adopt VAWS in clinical practice to gain additional information on HRQL, specifically concerning responsiveness to CPAP treatment.

One limitation of the present study was the number of patients excluded. Most of them were excluded because they did not evaluate HRQL questionnaires to adjust the sample to patients susceptible of improving in HRQL with CPAP treatment (exclusion of patients with relevant chronic diseases) and to obtain correct CPAP treatment. Another limitation was the restriction of the results for patients requiring CPAP although this is the most demanding population of HRQL test for clinical practice. Finally, having at our disposal a control group without treatment enabled us to better assess the ability of the VAWS to distinguish between treated and untreated patients. However, this could give rise to an ethical conflict in patients requiring CPAP. Considering the effect size between groups with adequate CPAP treatment and the increase in ESS with VAWS percentiles, a clinically significant improvement with CPAP treatment can be on 20.

In summary, the VAWS, which is a simple and rapid test, measures some components of HRQL related to SAHS and its responsiveness to CPAP treatment. It could therefore prove to be a useful tool in routine clinical practice, primarily to assess the change with treatment.

Acknowledgments We are indebted to Verónica Rodríguez and Vanessa Iglesias for assistance in the translation of the manuscript and Asunción Martín and Carmen Lorenzana for technical assistance.

Conflict of interest statement No authors have any conflicts of interest.

Supported by ISCIII-RTIC-03-11, JUNTAEX-IPR00A064 and SEPAR.

There was no influence from promoters about design, data collection, analysis, results, paper writing, and journal choice for publication of the study.

Annex 1

Visual analogical well-being scale

The patient must choose one point on the line that corresponds to his or her health-related quality of life regarding the symptoms for the consultation; in this case suspected of sleep apnea–hypopnea syndrome.

Question

If the line below expresses the most favorable (on the right) and the least favorable (on the left) well-being status with respect to the symptoms which are motive of the consultation, the question to be answered is the following: whereabouts on this line do you think you are?

Annex 2

Reliability is an important measurement in health-related quality-of-life tests. In the "original" study this analysis was not performed.

As a consequence, at the end of the study the same sleep centers collected 81 new patients subsidiary of CPAP treatment after complete polysomnography with the same inclusion and exclusion criteria and protocol than in the previous study. At the inclusion time and 3 weeks later (previous to CPAP treatment) a visual analogical well-being scale (VAWS) was carried out.

Statistically, we firstly compared the new sample (81 patients) with the previous (315 patients) on clinical, anthropometric and polysomnographic variables using t test (normal distribution) or Mann–Whitney (non-normal distribution) for continuous and χ^2 for qualitative variables. Secondly, we analyzed the test–retest reliability between VAWS measurements by intraclass correlation coefficient (ICC).

Table 7 Clinical, anthropometric and polysomnographic characteristics of patients from "original" and additional studies

	Original sample N=315	Additional sample N=81	p value
Age, years (mean and SD)	51.4 (9.7)	52.5 (11.5)	ns
Gender,% men	88.6	82.2	ns
BMI, kg. m^{-2} (mean and SD)	33.5 (7.1)	32.9 (6.0)	ns
Active worker,%	79.0	74.1	ns
Primary school only,%	56.2	50.6	ns
Alcohol, g/day (mean and SD)	24.5 (23.8)	18.4 (21.6)	0.04
Active smoker,%	40.8	26.6	0.02
FEV1,% predicted (mean and SD)	94.4 (18.2)	87.8 (16.8)	0.02
Sleep, hours per night (mean and SD)	7.0 (1.6)	6.9 (1.3)	ns
ESS (mean and SD)	15.7 (3.5)	15.0 (4.7)	ns
Habitual snoring,%	90.2	89.6	ns
VAWS (mean and SD)[a]	50.8 (22.2)	47.0 (25.4)	ns
AHI (mean and SD)	62.7 (22.9)	60.4 (23.8)	ns
Sat O2 below 90%,%	28.5 (26.7)	24.3 (23.3)	ns
Sleep efficiency	78.9 (12.9)	81.0 (11.2)	ns
Arousals index	56.6 (20.2)	52.0 (22.2)	ns
Light sleep	77.6 (12.3)	75.7 (14.5)	ns
Deep sleep	8.2 (9.1)	10.2 (11.7)	ns
REM sleep	14.0 (6.8)	13.0 (6.5)	ns

ESS Epworth sleepiness scale, *ASDA* ASDA sleepiness scale, *VAWS* visual analogical well-being scale, *AHI* apnea and hypopnea index

[a] VAWSs registered before treatment in original study and the first measure in the additional one

The two samples (original and additional) were similar because only statistical differences between them were observed in non relevant variables such as alcohol intake, active smoker and FEV1 (Table 1). Good ICC (0.83) was detected between inclusion time and 3-weeks-later measurements of VAWS.

References

1. Young T, Palta M, Dempsey J (1993) The occurrence of sleep-disordered breathing among middle aged adults. N Engl J Med 328:1230–1235
2. Masa JF, Rubio M, Findley LJ, Cooperative Group (2000) Habitually sleepy drivers have a high frequency of automobile crashes associated with respiratory disorders during sleep. Am J Respir Crit Care Med 162:1407–1412
3. Terán J, Jiménez A, Cordero J (1999) The association between sleep apnea and the risk of traffic accidents. N Engl J Med 340:847–851
4. Peppard PE, Young T, Palta M (2000) Prospective study of the association between sleep-disordered breathing and hypertension. N Engl J Med 342:1378–1384
5. Parra O, Arboix A, Bechich S (2000) Time course of sleep-related breathing disorders in first-ever stroke or transient ischemic attack. Am J Respir Crit Care Med 161(2):375–380
6. Masa JF, Rubio M, Pérez P (2006) Association between habitual naps and sleep apnea. Sleep 11:1463–1468
7. Marin JM, Carrizo SJ, Vivente E (2005) Long-term cardiovascular outcomes in men with obstructive sleep apnoea–hypopnoea with or without treatment with continuous positive airway pressure: an observational study. Lancet 365:1046–1053
8. Ballester E, Badia JR, Hernández L (1999) Evidence of the effectiveness of continuous positive airway pressure in the treatment of sleep apnea/hypopnea syndrome. Am J Resp Crit Care Med 159:495–501
9. Masa JF, Jiménez A, Durán J, the Cooperative Group (2004) Alternative methods of titrating continuous positive airway pressure. Am J Respir Crit Care Med 170:1218–1224
10. Chakravorty I, Cayton RM, Szczepura A (2002) Health utilities in evaluating intervention in the sleep apnoea/hypopnoea syndrome. Eur Respir J 20:1233–1238
11. Jenkinson C, Davies RJO, Mullins R (1999) Comparison of therapeutic and subtherapeutic nasal continuous positive airway pressure for obstructive sleep apnoea: a randomised prospective parallel trial. Lancet 353:2100–2105
12. Engleman HM, Kingshott RN, Wraith PK (1999) Randomized placebo-controlled crossover trial of continuous positive airway pressure for mild sleep apnea/hypopnea syndrome. Am J Respir Crit Care Med 159:461–467
13. Johns MW (1991) A new method for measuring daytime sleepiness: the Epworth Sleepiness Scale. Sleep 14:540–545
14. Lacasse Y, Godbout C, Series F (2002) Health-related quality of life in obstructive sleep apnea. Eur Respir J 19:499–503
15. EuroQol Group (1990) EuroQol—a new facility for measurement of health-related quality of life. Health Policy 16:199–208
16. Sanner BM, Klewer J, Trumm A (2000) Long-term treatment with continuous positive airway pressure improves quality of life in obstructive sleep apnoea syndrome. Eur Respir J 16:118–122
17. Jenkinson C, Stradling J, Peterasen S (1998) How should we evaluate health status? A comparison of three methods in patients presenting with obstructive sleep apnoea. Qual Life Res 7:95–100
18. Schmidlin M, Fritsch K, Matthews F, Thurnheer R, Senn O, Bloch KE (2010) Utility indices in patients with the obstructive sleep apnea syndrome. Respiration 79(3):200–8, Epub May 27
19. Guyatt GH, Feeny DH, Patrick DL (1993) Measuring health-related quality of life. Ann Intern Med 118:622–629

20. Mara J, Rivero-Arias O, Durán-Cantolla J, Alonso-Álvarez ML, Gaminde I, De la Torre MG (2005) Efecto del tratamiento con presión positiva continua en la vía respiratoria durante la noche en la calidad de vida de los pacientes con apnea del sueño. Med Clin (Barc) 125(16):611–615
21. Lacasse Y, Godbout C, Series F (2002) Independent validation of the sleep apnoea quality of life index. Thorax 57:483–488
22. Weaver TE, Laizner AM, Evans LK (1997) An instrument to measure functional status outcomes for disorders of excessive sleepiness. Sleep 20:835–843
23. Piccirillo JF, Gates GA, White DL (1998) Obstructive sleep apnea treatment outcomes pilot study. Otolaryngol Head Neck Surg 102:1855–1864
24. Lacasse Y, Bureau MP, Series F (2004) A new standardised and self-administered quality of life questionnaire specific to obstructive sleep apnoea. Thorax 59:494–499
25. Barbé F, Amilibia J, Capote F (1995) Diagnóstico del Síndrome de Apneas Obstructivas Durante el Sueño. Arch Bronconeumol 31:460–462
26. Rechtschaffen A, Kales A (eds) (1968) A manual of standardized terminology and scoring system for sleep stages of human subjects. Brain Information Service/Brain Research Institute, University of California at Los Angeles
27. American Sleeep Disorders Association (1992) EEG arousals: scoring rules and examples. Sleep 15:174–183
28. Ware J, Sherbourne C (1992) The Mos 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 30:473–483
29. Liang MH, Larson MG, Cullen KE (1985) Comparative measurement efficiency and sensitivity of five health status instruments for arthritis research. Arthitis Rheum 28:542–547
30. Diagnostic classification steering committee (1990) International classification of sleep disorders: diagnostic and coding manual. American Sleep Disorders Association, Rochester
31. American Thoracic Society (1987) Standardization of spirometry, 1987 update. Am Rev Respir Dis 136:1285–1298
32. Kramer MS, Fenstein AR (1981) Clinical biostatistics. LIV. The biostatistics of concordance. Clin Pharmacol Ther 29:111–123
33. Ferrer M, Vilagut G, Monasterio C (1999) Measurement of the perceived impact of sleep problems: the Spanish version of the functional outcomes sleep questionnaire and the Epworth Sleepiness Scale. Med Clin (Barc) 113(7):250–255