Validation of a single item to assess daytime sleepiness for the Swiss Transplant Cohort Study

Context—Daytime sleepiness in kidney transplant recipients has emerged as a potential predictor of impaired adherence to the immunosuppressive medication regimen. Thus there is a need to assess daytime sleepiness in clinical practice and transplant registries.

Objective—To evaluate the validity of a single-item measure of daytime sleepiness integrated in the Swiss Transplant Cohort Study (STCS), using the American Educational Research Association framework.

Methods—Using a cross-sectional design, we enrolled a convenience sample of 926 home-dwelling kidney transplant recipients (median age, 59.69 years; 25%-75% quartile [Q25-Q75], 50.27-59.69), 63% men; median time since transplant 9.42 years (Q25-Q75, 4.93-15.85). Daytime sleepiness was assessed by using a single item from the STCS and the 8 items of the validated Epworth Sleepiness Scale. Receiver operating characteristic curve analysis was used to determine the cutoff for the STCS daytime sleepiness item against the Epworth Sleepiness Scale score.

Results—Based on the receiver operating characteristic curve analysis, a score greater than 4 on the STCS daytime sleepiness item is recommended to detect daytime sleepiness. Content validity was high as all expert reviews were unanimous. Concurrent validity was moderate (Spearman ϱ, 0.531; $P < .001$) and convergent validity with depression and poor sleep quality although low, was significant (ϱ, 0.235; $P < .001$ and ϱ, 0.318, $P = .002$, respectively). For the group difference validity: kidney transplant recipients with moderate, severe, and extremely severe depressive symptom scores had 3.4, 4.3, and 5.9 times higher odds of having daytime sleepiness, respectively, as compared with recipients without depressive symptoms.

Conclusion—The accumulated evidence provided evidence for the validity of the STCS daytime sleepiness item as a simple screening scale for daytime sleepiness. (*Progress in Transplantation.* 2013;23:xxx-xxx)

doi: http://dx.doi.org/10.7182/pit2013XXX

Hanna Burkhalter, RN, MSc, Anna Wirz-Justice, PhD, Christian Cajochen, PhD, Terri Weaver, RN, PhD, Jürg Steiger, MD, Thomas Fehr, MD, Reto Martin Venzin, MD, Sabina De Geest, RN, PhD
University of Basel, Switzerland (HB, AW-J, CC, TW, SDG), University of Illinois at Chicago College of Nursing (TW), University Hospital Basel, Switzerland (HB, JS), University Hospital Zürich, Switzerland (TF), University Hospital Bern, Switzerland (RMV), Katholieke Universiteit Leuven, Belgium (SDG)

Corresponding author: Sabina De Geest, RN, PhD, Institute of Nursing Science, University of Basel, Bernoullistrasse 28, CH-4056 Basel, Switzerland (e-mail: Sabina.degeest@unibas.ch)

To purchase electronic or print reprints, contact:
American Association of Critical-Care Nurses
101 Columbia, Aliso Viejo, CA 92656
Phone (800) 899-1712 (ext 532) or
(949) 448-7370 (ext 532)
Fax (949) 362-2049
E-mail reprints@aacn.org

The Swiss Transplant Cohort Study (STCS) is a nationwide prospective multicenter interdisciplinary cohort study including all patients receiving organ transplants at 1 of the 6 transplant centers in Switzerland (Lausanne, Geneva, Basel, Zürich, Bern, St Gallen).[1,2] The STCS began enrollment of patients and data collection on May 2, 2008. In this study, only 1 item assessing sleep quality was integrated. In view of the evidence described in the following paragraphs, we developed a daytime sleepiness item to be included, after validation, in the STCS and used for this same format as the 1-item sleep quality measure already included in this nationwide cohort study.

It is commonly recognized that lack of nocturnal sleep increases the tendency to fall asleep during the day. Daytime sleepiness is the subjective report of an increased desire to fall asleep and lack of energy during the day even after an adequate night's sleep.[3] Daytime sleepiness is not a disorder in and of itself, yet it is an important symptom of many other sleep disorders. Daytime sleepiness is associated with poor performance,[4] cognitive slowing, attention failures, errors,[5] and accidents.[6] It is also a known predictor of increased morbidity and mortality in patients with cardiovascular disease[7,8] and diabetes.[9]

So far, published evidence of the presence and impact of daytime sleepiness in kidney transplant recipients is limited. In a previous study,[10] we found that 34.1% of kidney transplant recipients suffered from poor daytime functioning, a measure similar to daytime sleepiness,

although daytime sleepiness was not directly assessed as functioning was more related to performance. The prevalence of daytime sleepiness reported in the general population ranges from 2.5%[11] to 25.7%[12] and in hemodialysis patients from 15%[13] to 27.3%.[14]

Because of its impact on clinical outcome, daytime sleepiness is emerging as a relevant parameter to assess in research and clinical practice. We therefore decided to integrate the assessment of daytime sleepiness into the STCS.[2] Many instruments exist to measure daytime sleepiness objectively (the Multiple Sleep Latency Test,[15] the Maintenance of Wakefulness Test,[16] and the Psychomotor Vigilance Task[17,18]) or subjectively, using self-report (the Epworth Sleepiness Scale [ESS]).[19,20] The commonality among these instruments is that they are time-consuming and labor-intensive, and some are in a laboratory situation not reflecting daily life. Although the ESS is easy to administer, clinicians do not commonly use it outside the field of sleep medicine, perhaps because of a lack of awareness of the importance of daytime sleepiness problems or a reluctance to add 8 questions to an already lengthy medical assessment.

Taking this into account, a simpler more direct question about daytime sleepiness might better serve as a screening tool. A single-item questionnaire has been previously published that asks "Please measure your sleepiness on a typical day" (0 = none, 10 is highest; cutoff ≥7); however, this item is validated only in patients who have a diagnosis of a sleep disorder and no specific recall period is used, making it unclear to patients which time frame to take into consideration when completing this item.[21]

Daytime sleepiness, sleep quality, and depression are interrelated and have been explored and explained in different models and approaches.[22-24] Poor sleep quality and daytime sleepiness are included in the criteria for the diagnosis of depression. The *Diagnostic and Statistical Manual of Mental Disorders*[25] defines the criteria for a major depressive disorder: "Difficulty falling or staying asleep (insomnia), or sleeping more than usual (hypersomnia)" (criterion 1d) and feeling tired or having little energy (1e). For dysthymic disorder, criterion 2 is set as "Sleeping too much or having difficulty sleeping" and criterion 3 as "Low energy or fatigue." Depression is common among kidney transplant recipients; reported cumulative incidences were 5.05%, 7.29%, and 9.10% at 1, 2, and 3 years after transplant, respectively.[26] The prevalence of poor sleep quality in kidney transplant recipients ranges from 30% to 62%.[27-29]

Patients thought to have obstructive sleep apnea who had a high score on the ESS also had a significantly higher score on the Center for Epidemiological Studies Depression Scale.[30] In a cross-sectional study[31] with 3045 community-dwelling women aged 70 years and older, depression was associated with poorer sleep quality and daytime sleepiness. Another cross-sectional study including 67 patients diagnosed with a depressive episode showed that daytime sleepiness measured with the ESS correlated highly with scores on the Hamilton Depression Rating Scale ($r = 0.69$, $P < .001$).[32]

Given that transplant registries do not allow the use of lengthy measures because of the burden on participants, we developed a single item to measure daytime sleepiness prospectively as part of the STCS.[2] It needs to be determined, however, which cutoff with respect to other existing measures is appropriate to identify daytime sleepiness by a single item (ie, cross-validation). Moreover, the validity of this single-item STCS daytime sleepiness instrument needs to be demonstrated before transplant outcomes can be analyzed.

Table 1 Hypotheses and research questions guiding the validation process

Basis of evidence	Question or hypothesis
Content	Question 1: Does STCS-DS item reflect the concept of daytime sleepiness based on expert review?
Relationship to other variables	Hypothesis 1: STCS-DS (low score means no DS) is positively correlated with ESS sum score (low score means no DS). (Concurrent validity) Hypothesis 2: There is a positive correlation between STCS-DS and the DASS depression score. (Convergent validity) Hypothesis 3: There is a positive correlation between STCS-DS and sleep quality measured by the PSQI score. (Convergent validity) Hypothesis 4: Higher levels of depressive symptoms are associated with higher degree of daytime sleepiness. (Group difference validity)

Abbreviations: DASS, Depression, Anxiety and Stress Scale; DS, daytime sleepiness; ESS, Epworth Sleepiness Scale; PSQI, Pittsburgh Sleep Quality Index; STCS-DS, Swiss Transplant Cohort Study–Daytime Sleepiness.

Theoretical Background: Validity

Assessing the validity of an instrument implies gathering as much evidence as possible to support validity related to test content, response processes, internal structure, relation to other variables, and consequences of testing, as outlined by the American Educational Research Association's framework.[33,34] For the purpose of our study, we focused on evidence related to test content and relationships to other variables. In order to guide our validation process, we developed 4 hypotheses based on empirical evidence that outlines the relationships between daytime sleepiness, sleep quality, and depression to be tested as part of the validation process (Table 1).[33]

In absence of an established validated cutoff of the STCS daytime sleepiness item, the first aim of this study was to identify the optimal cutoff point for classifying patients as having daytime sleepiness, using the ESS as reference. However the main aim of this study was to assess the validity of the STCS daytime sleepiness item by using the American Education Research Association's framework to assess evidence based on content and evidence based on relationship to other variables.

Methods

Design, Setting, and Sample

This study used a cross-sectional multicenter correlational design. A convenience sample of 926 home-dwelling kidney transplant recipients treated at 3 Swiss kidney transplant centers participated in the study. Patients were included if at least 6 months had elapsed since they had received the transplant, they had the ability to understand and read German, they were 18 years of age or older, and they had provided written informed consent. Individuals were excluded if they were unable to complete the study questionnaire by themselves for any reason.

Variables and Measurements

Age (in years), sex, and time since transplant (in years) were retrieved from the patients' hospital charts. Daytime sleepiness was assessed by 2 measures, the STCS daytime sleepiness (STCS-DS) item[2] and the score on the ESS.[19] The STCS-DS item asks subjects to rate their overall daytime sleepiness in the past 4 weeks on a scale of 0 (no sleepiness) to 10 (extreme sleepiness). This item is similar to a widely used but not validated item in sleep diaries carried out with actigraphy measurements.[21] The layout was made congruent with the STCS sleep quality item[10] derived from the Kidney Disease Quality of Life Short Form.[35] Receiver operating characteristic (ROC) curve analysis[36] has been used to establish an appropriate cutoff for the STCS-DS item.

The ESS is a validated questionnaire that contains 8 items that measure a subject's expectation of dozing in 8 hypothetical situations. Dozing is defined as falling into a light sleep.[37] Dozing probability ratings range from 0 (no probability) to 3 (high probability). Scores on the 8 items are summed, yielding a total dozing score between 0 and 24. An ESS sum score of 6 or greater indicates daytime sleepiness.[19] A score of 10 or greater indicates that the subject is very sleepy and should seek medical advice.[19] Total ESS scores show high test-retest reliability ($\varrho = 0.82$, $P < .001$)[38] and a high level of internal consistency (Cronbach $\alpha = 0.74$-0.88 in 4 different groups of chronically ill patients).[39] A single factor emerged when performing factor analyses on ESS item scores of 150 patients and 104 students.[38] The ESS has been validated for application in German-speaking populations.[20]

Sleep Quality

Sleep quality was assessed by using the Pittsburgh Sleep Quality Index (PSQI), a self-rated questionnaire consisting of 19 items, assessing a wide variety of factors related to sleep quality during a 1-month time interval, including estimates of sleep duration and latency and of the frequency and severity of specific sleep-related problems. These 19 items are grouped into 7 component scores, each weighted equally on a scale from 0 to 3. The 7 component scores are then summed to yield a global PSQI score, which has a range of 0 to 21; higher scores indicate worse sleep quality.

A cutoff of greater than 5 points is used to classify patients as having poor sleep quality.[40] A PSQI global score greater than 5 resulted in a sensitivity of 98.7% and specificity of 84.4% as a marker for sleep disturbances in insomnia patients versus healthy control subjects.[41] Backhaus et al[41] translated the PSQI into German by using the back-translation method. Item analysis confirmed internal consistency of the German version of the PSQI scale (Cronbach α of 0.85). The test-retest reliability for the short interval (2 days) was high for the global as well as for subscale scores (0.76-0.92). For the longer interval (mean, 45.6 days; SD, 18 days), the test-retest reliability was low for the subscores "sleep quality" ($r = 0.23$) and "sleep disturbance" ($r = 0.84$), whereas it remained moderate to high for the global score ($r = 0.86$) and 4 of 7 subscores, ranging from 0.59 to 0.83.[41] Sleep diaries show a high correlation to the PSQI,[41] indicating good validity based on relation to other relevant variables.

Depression

Depression was measured with the Depression, Anxiety, and Stress Scale (DASS): a 21-item self-report instrument, of which 7 items measure depressive symptoms (DASS-D)[42] on an ordinal 4-point Likert severity/frequency scale to rate the extent to which they have experienced each state in the past week: 0 = did not apply to me; 3 = applied to me very much in the past week. Scores are summed and multiplied by 2, resulting in a range of 0 to 42 for each subscale. The following cutoffs are used to evaluate severity of depressive symptoms: 0-9, no depressive symptoms; 10-13, mild symptoms; 14-20, moderate symptoms; 21-27, severe symptoms; and ≥28, extremely severe symptoms.[43] The Cronbach α for the DASS-D scale is high ($\alpha = 0.88$; 95% CI, 0.87-0.89) and shows good concurrent validity with the Brief Symptom Inventory ($r = 0.70$),[44] the Beck Depression Inventory ($r = 0.74$),[42] the Personal Disturbance Scale depression scale ($r = 0.78$), and the Hospital Anxiety and Depression Scale ($r = 0.66$).[45]

Evidence Based on Content: Content Validity

Evidence supporting test content includes logical or empirical analyses of how adequately that content represents the domain of interest and can include the judgment of content experts.[33] Validity-related data analysis procedures were guided by research questions and hypotheses formulated to test item validity (Table 1).

To test if the STCS-DS item indeed measures daytime sleepiness, we asked all members of the STCS Psychosocial Interest Group (transplant physicians, psychologists, nurses, physicians, and epidemiologists) to evaluate whether the STCS-DS item captures daytime sleepiness, expressed as the percentage of agreement. First the item was evaluated and then it was voted on at one of the in-person meetings of the STCS Psychosocial Interest Group.

Evidence Based on Relationships With Other Variables

Evidence of relationships to other variables is commonly evaluated by assessing associations among variables. If the observed relationships match the hypothesized relationships, then the evidence supports the validity of the interpretation.[33]

Concurrent Validity. Concurrent validity (or criterion-related validity) requires that both variables measuring the same concept be captured at 1 point in time. As the strength of the correlation increases, the probability that the variables measure the same concept is increased.[46] Daytime sleepiness is the subjective report of an increased desire to fall asleep and lack of energy during the day even after an adequate night's sleep.[3] Excessive daytime sleepiness is sleepiness in a situation when an individual would be expected to be awake and alert.[47] Following these 2 definitions, the ESS asks about 8 very concrete situations when an individual would be expected to be awake and alert. In contrast, the STCS-DS asks for sleepiness in the past 4 weeks unrelated to a specific situation. The ESS is more precise in asking for these situations whereas the STCS-DS asks for a metacognitive process transforming the concept of daytime sleepiness into a daily life situation.

Therefore we expected a moderate to high correlation between the STCS-DS item and the ESS total score, given that STCS-DS asks for sleepiness in general in the past 4 weeks, and the ESS asks the same, but in specific circumstances.[46]

Convergent Validity. Convergent validity is when 2 measures of a construct that theoretically should be related to each other are, in fact, observed to be related to each other.[48] In order to provide evidence for convergent validity, we expect to observe (1) a moderate to high correlation between the STSC-DS item and sleep quality, as measured by the PSQI, and (2) a low to moderate positive correlation between the STSC-DS item and depressive symptoms, measured by the DASS-D.[48]

Known Group Difference Validity. Known group difference validity is when data are collected from 2 groups that have expected differences on the measure of interest. If the measure is able to discriminate between the groups through statistically significant findings, this provides evidence for the validity of the measure.[49] The evidence states that depression is associated with poorer sleep quality and daytime sleepiness.[31] We hypothesize therefore that stratifying our sample on the DASS-D score will split the sample in a similar way for daytime sleepiness. If DASS-D scoring (normal, mild, moderate, severe, extremely severe) is able to discriminate between the groups through statistically significant findings, this provides evidence for the validity of the daytime sleepiness measure.[49]

Data Collection

Addresses of all patients who fulfilled the eligibility criteria were extracted from the centers' transplant databases by the responsible physician and the head outpatient nurse. Each potential participant received a package containing an information letter, the informed consent documents, prestamped envelopes, and the questionnaires. Participants who consented to participate completed the study questionnaires and returned them to the researcher by mail. Patients not responding were contacted once by telephone and invited to participate again. If they agreed to do so, a new package was sent. The study was approved by the ethics committee of Basel, Bern, and Zürich. Data were deidentified and stored in an electronic databank.

Statistical Analysis

Data were entered once and randomly checked for discrepancies with original data (<1%). The Package IBM SPSS Statistics 19 (Version 19.0.0, IBM Corporation) was used for statistical analysis, setting 5% for all critical probability levels. Descriptive statistics included mean, standard deviation (SD), median and interquartile ranges (IQR), and frequencies as appropriate based on measurement levels and distributions of variables. The Mann-Whitney U test and χ^2 tests were used to explore whether recipients' sex, age, and years since transplant differed between those who responded and those who did not send back the questionnaires.

Establishing the Best Cutoff for the STCS-DS. We used ROC curve analysis to establish an appropriate cutoff for the STCS-DS item. We plotted the true-positive rate (sensitivity) as a function of the false-positive rate (1 minus specificity) for different cutoff values for the

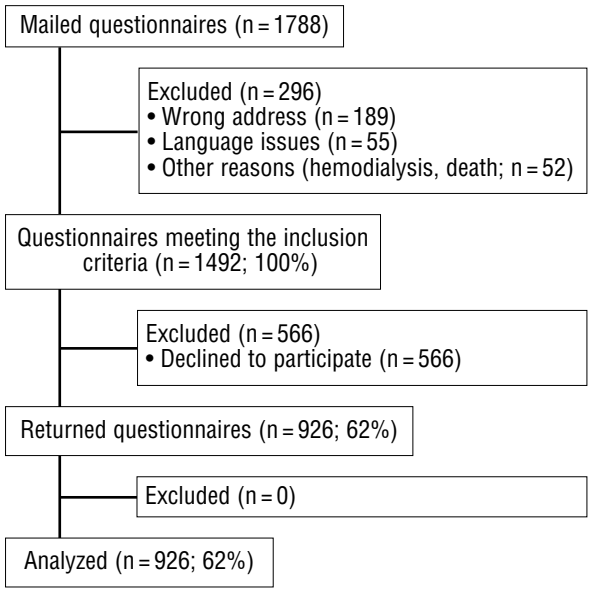

Figure Flow diagram of sample.

Table 2 Description of the sample

Characteristic	All
No. of patients	926
Males, % (No.)	63.3 (586)
Age, median (Q25-Q75), y	59.69 (50.27-59.69)
Years since transplant, median (Q25-Q75)	9.42 (4.93-15.85)
Daytime sleepiness measured by the ESS	
ESS Score	6 (3-9)
Mean (SD)	50.9 (471)
DS, % (No.)	21.3 (197)
Excessive DS, % (No.)	
Daytime sleepiness measured with the STCS-DS	
STCS-DS score, median (Q25-Q75)	3 (1-5)
STCS-DS, % (No.)	32.4 (300)
Sleep quality measured with the PSQI	
PSQI score, median (Q25-Q75)	6 (4-10)
PSQI, % (No.)	49.5 (456)
Depressive symptoms measured with the DASS	
DASS score, median (Q25-Q75)	3 (1-6)
DASS, % (No.)	33.9 (310)

Abbreviations: DASS, Depression, Anxiety and Stress Scale (>9); DS, daytime sleepiness (cutoff ≥6); ESS, Epworth Sleepiness Scale; Excessive DS, daytime sleepiness (cutoff ≥10); PSQI, Pittsburgh Sleep Quality Index (>5); STCS-DS, Swiss Transplant Cohort Study Daytime Sleepiness Item (≥4).

STCS-DS item, relative to the ESS sum score (ie, cutoff ≥10, where a score of 10 or more indicates very sleepy and should seek medical advice and a cutoff ≥6, whereas a score of 6-9 suggests daytime sleepiness[19]).

Validity-Related Data Analysis. At an in-person meeting in January 2011, the daytime sleepiness concept was presented by the author in a 10-minute PowerPoint presentation, showing different arguments and perspectives. After the presentation, the members (N = 20) of the Psychosocial Interest Group of the STCS were invited to comment and vote. The percentage of agreement was assessed by counting the participating members present at the meeting and dividing the number of members agreeing to the statement "The STCS-DS item reflects the concept of daytime sleepiness" by that total.

Spearman rank-order correlation was used to examine the association between the STCS-DS item and the ESS sum score (Table 1, H1), depressive symptoms (Table 1, H2), and poor sleep quality (Table 1, H3) on the PSQI. The group difference validity was done with a binary logistic regression with the DASS-D score (cutoff <9 normal; cutoff ≥13 mild depression; cutoff ≥20 moderate depression; cutoff ≥27 severe depression; cutoff ≥28 extremely severe depression) as a predictor of STCS-DS (Table 1, H4).

Results

Out of 1788 kidney transplant recipients, 1492 met the eligibility criteria and had a valid home address, and 926 of them returned a completed questionnaire (62% response rate; see Figure). Responders and nonresponders did not differ except for age: nonresponders (n = 509) were significantly younger than responders ($t = 2.51$, $df = 1039$, $P = .01$). The analyses were based on the 926 participants who had complete data. The median age was 59.69 years (25%-75% quartile [Q25-Q75], 50.27-59.69), 586 were men (63.3%), and the median time since transplant was 9.42 years (Q25-Q75, 4.93-15.85; Table 2).

Cutoff Values for the STCS-DS Item

Based on ROC curve analyses, a cutoff value of 4.5 on the STCS-DS item yielded the highest levels of sensitivity and specificity in predicting the ESS cutoff of at least 10 (excessive daytime sleepiness) with a sensitivity of 67% and a specificity of 84%. A cutoff value of 3.5 for the STCS-DS item also had the highest sensitivity (57%) and specificity (77%) relative to the ESS cutoff of 6 or greater (daytime sleepiness; Table 3).

Evidence Based on Content

All experts (100%) agreed that the STCS-DS item reflects the concept of daytime sleepiness (Table 1, Q1).

Evidence Based on Relations to Other Variables

The concurrent validity between the STCS-DS item and the ESS score (Table 1, H1) showed a moderate

No sleepiness (ESS score <6)	DS (ESS score >5)	No sleepiness (ESS score <10)	Excessive DS (ESS score <9)
455	471	729	197
60.9 (277)	65.6 (309)	62.6 (456)	66.0 (130)
61.17 (51.06-68.65)	58.68 (49.29-67.32)	60.61 (50.85-68.23)	58.17 (48.19-65.65)
9.32 (4.92-15.59)	9.58 (4.96-16.15)	9.81 (5.01-16.29)	8.51 (4.85-14.99)
3 (2-4)	9 (7-11)	4 (3-7)	12 (11-14)
0 (0)	41.8 (197)	37.6 (274)	100 (197)
2 (1-3)	4 (2-6)	2 (1-4)	6 (3.5-7)
16.5 (75)	47.8 (225)	31.4 (229)	75.1 (148)
5 (3-8)	6 (4-10)	5 (3-8)	8 (5-11)
40.1 (181)	58.5 (275)	43.8 (317)	70.6 (139)
2 (0-5)	3 (1-7)	2 (1-5)	5 (2-9)
26.3 (118)	41.2 (192)	29.3 (211)	50.3 (99)

but significant correlation (Spearman ϱ, 0.531; $P < .001$). Convergent validity (Table 1, H2) showed that the STCS-DS item was correlated minimally but significantly with the DASS-D score (Spearman ϱ, 0.235; $P < .001$) and the STCS-DS of 4 or greater was correlated minimally but significantly (Table 1, H3) with the PSQI score (Spearman ϱ, 0.318; $P = .002$). Table 4 shows the group difference validity results. Higher levels of depressive symptoms are associated with higher odds of daytime sleepiness. Kidney transplant recipients with moderate to extremely severe depressive symptoms show 3.4, 4.3, and 5.9 times higher odds of having daytime sleepiness, respectively, as compared with recipients without depressive symptoms.

Discussion

This study examined the validity of using 1 item incorporated in the STCS to measure daytime sleepiness. The STCS is a nationwide prospective cohort study that uniquely also assesses selected psychosocial and behavioral variables from before transplant to lifelong after transplant, including sleep quality and daytime sleepiness.[2] As the STCS does not allow extensive assessment of each variable (in an effort to limit

Table 3 Receiver operating characteristic curve for the single daytime sleepiness item

Variable	No. (%)	Property	STCS-DS	STCS-DS
Daytime sleepiness[a]	469 (50.7)	Area under curve (95% CI)	0.75 (0.71-0.78)	0.75 (0.71-0.78)
		Optimal cutoff	3.5	4.5
		Sensitivity, %	58	48
		Specificity, %	77	84
		Area under curve (95% CI)	0.80 (0.77-0.83)	0.80 (0.77-0.83)
Excessive daytime sleepiness[b]	197 (21.2)	Optimal cutoff	3.5	4.5
		Sensitivity, %	75	67
		Specificity, %	69	87

Abbreviation: STCS-DS, Swiss Transplant Cohort Study Daytime Sleepiness Item.
[a] Daytime sleepiness score on the Epworth Sleepiness Scale ≥ 6.
[b] Daytime sleepiness score on the Epworth Sleepiness Scale ≥ 10.

Table 4 Predictors of Swiss Transplant Cohort Study–Daytime Sleepiness in the simple logistic regression analysis

Symptoms of depression[a]	Odds (95% CI)	*P*
Normal (reference)		
Mild	1.377 (0.91-2.09)	.13
Moderate	3.411 (2.25-5.165)	>.001
Severe	4.260 (2.29-7.93)	>.001
Extremely severe	5.990 (2.86-12.54)	>.001

[a] Based on Depression, Anxiety, and Stress Scale–Depression.

the burden on subjects), daytime sleepiness is assessed with 1 item derived from previous research. The validity of the STCS-DS item was tested by using a large dataset of kidney transplant patients participating in a research program on sleep in kidney transplant recipients, a study separate from the STCS. Validity was tested in view of evidence related to test content and evidence related to relationships with other variables.[33]

Cutoff Score for STCS-DS Item

As a cutoff for the STCS-DS item has not yet been established, we used the ROC curve analysis to determine it. This analysis showed useable results based on 2 cutoffs of the ESS, a validated and established instrument to assess daytime sleepiness. We did this analysis for both of the cutoffs of the ESS, thus allowing 2 suggested cutoffs for our STCS-DS item. The ideal cutoff for daytime sleepiness is 3.5 and the ideal cutoff for excessive daytime sleepiness is 4.5. To avoid screening for a diagnosis of a sleep disorder where one does not exist, we recommend restrictive positive screening using values greater than or equal to 4 in the STCS-DS item. These findings support its use for a general screening tool followed by in-depth assessments for sleep disorders.

Evidence Based on Test Content

The agreement of our experts showed that the STCS-DS measures the daytime sleepiness concept. For further studies, we suggest assessing the content validity index for the validity based on test content.[50]

Evidence Based on Relationships With Other Variables

We provided first evidence for validity of the STCS-DS item. H1 showed a significant moderate correlation, demonstrating the similarity of the concepts. We had hoped to find a higher correlation; however, this hope was unrealistic, considering that the aim of replacing 8 items with 1 item is very ambitious.

Depressive symptoms (H2) and poor sleep quality (H3) were significantly correlated; this result may be, in part, due to the large sample. These results highlight the expected interaction of these variables with daytime sleepiness and demonstrate that the variables measure different things. The low correlation may indicate that other factors that we did not correct for and did not assess affect the variability. For example, we did not assess for history of insomnia; this factor is a predictor for future development of depression in older persons as well as young adults.[51] Further, we had no polysomnography measurements to establish changes in sleep architecture. We will in future have actigraphy results in a selected subgroup that may provide detailed information about the nature of the diverse sleep disturbances in kidney transplant recipients. In addition, kidney transplant recipients are subjected to other strong factors limiting the correlations and affecting sleep, for example, the consequences of immunosuppressive drug therapy,[52] impaired immune system,[53] and high vulnerability for infections.[54]

Group difference validity has been shown with the DASS-D scale. Higher levels of depressive symptoms were significantly associated with a higher odds ratio of daytime sleepiness. These validities show that STCS-DS measures a concept similar to that measured by the ESS, that STCS-DS is related to but measures a different concept than depression and sleep quality; and that a positive score on the DASS-D scale is associated with an increased odds ratio for daytime sleepiness. This finding suggests that careful use of a simple screening is beneficial and—in the case of a positive screening value—should be followed up by a thorough assessment for a sleep disorder.

We suggest further testing all sources of validity,[34,48] especially evidence based on response processes and evidence based on consequences of testing. Evidence based on internal structure is in this case irrelevant, as there is only 1 item. Validity based on response processes should be assessed with more types of assessments (interview techniques, verbal protocol methods, think-aloud techniques) at different time points.[33] Validity based on consequences of testing could be assessed by assessing the alertness and performance (the impact of daytime sleepiness score results) of the kidney transplant recipients and the consequences on the nurses in charge in view of higher services. For example, a study could be done to measure the difference in adherence to immunosuppressive medications between kidney transplant recipients screened with daytime sleepiness receiving no intervention and kidney transplant recipients screened with daytime sleepiness receiving light therapy.

Limitations of the Study

This study has several limitations. The questionnaire responders were significantly older than were the nonresponders, perhaps because they are retired[55] and thus have more time to answer the questionnaires. Next, the study aimed to include a broad sample with

low selection criteria and low questionnaire and measurement burden. Therefore we had a big sample, only 1 time point measure, and few items included in the questionnaire, limiting the analysis for testing other validity dimensions. A further limitation is that the questionnaire did not permit investigation of self-rated daytime sleepiness as related to specific sleep disorders. Last, testing the validity and reliability of a 1-item scale is limited (eg, reliability [internal consistency] is only possible with more than 1 item) as is the study design (eg, for test-retest validity and criterion validity a longitudinal study is needed).[46]

Conclusion

The "Standards for Educational and Psychological Tests" proposes that validity is a unitary concept supported by theory, and accumulating evidence provides a sound scientific basis for the proposed score interpretation. Our results support the importance of assessment of daytime sleepiness, sleep quality, and depressive symptoms in kidney transplant patients. Our validity testing of "evidence based on content" and "evidence based on relationship to other variables" provided supporting evidence for the validity of a single daytime sleepiness item in the STCS-DS.

Acknowledgments

The authors thank all the volunteers who helped with data collection and the ambulatory care teams for the excellent collaboration with information transfer.

Financial Disclosures

This study was funded by the International Transplant Nurse Society and by research grants from the Swiss Renal Foundation (Alfred and Erika Bär-Spycher Foundation).

Terri E. Weaver has consulted for Philips Respironics, Inc, and TEVA, Inc. She has received FOSQ License Agreements from Nova Som, GlaxoSmithKline, Philips Respironics, Inc, Cephalon, Inc, and Nova Nordsk. Dr Weaver is on the Board of Directors for ViMedicus, Inc.

References

1. Koller MT, Van Delden C, Pascual M, Müller N, Baumann P, Steiger J. Die Schweiterische Transplantationskohortenstudie (STCS). *Schweiz Arzteztg.* 2009;90(24):953-955.
2. Steiger J, Van Delden C, Pascual M. Swiss Transplant Cohort Study. 2012. http://stcs.ch/. Accessed December 21, 2012.
3. Johns MW. What is excessive daytime sleepiness? In: Fulke P, Vaughan S, eds. *Sleep Deprivation: Causes, Effects and Treatment.* Nova Science Publishers; 2009:59-94.
4. Ulfberg J, Carter N, Talback M, Edling C. Excessive daytime sleepiness at work and subjective work performance in the general population and among heavy snorers and patients with obstructive sleep apnea. *Chest.* 1996;110(3):659-663.
5. Melamed S, Oksenberg A. Excessive daytime sleepiness and risk of occupational injuries in non-shift daytime workers. *Sleep.* 2002;25(3):315-322.
6. Czeisler CA. Impact of sleepiness and sleep deficiency on public health—utility of biomarkers. *J Clin Sleep Med.* 2011; 7(5 suppl):S6-8.
7. Ayas NT, White DP, Manson JE, et al. A prospective study of sleep duration and coronary heart disease in women. *Arch Intern Med.* 2003;163(2):205-209.
8. King CR, Knutson KL, Rathouz PJ, Sidney S, Liu K, Lauderdale DS. Short sleep duration and incident coronary artery calcification. *JAMA.* 2008;300(24):2859-2866.
9. Knutson KL, Spiegel K, Penev P, Van Cauter E. The metabolic consequences of sleep deprivation. *Sleep Med Rev.* 2007;11(3): 163-178.
10. Burkhalter H, Sereika SM, Engberg S, Wirz-Justice A, Steiger J, De Geest S. Validity of 2 sleep quality items to be used in a large cohort study of kidney transplant recipients. *Prog Transplant.* 2011;21(1):27-35.
11. Kaneita Y, Ohida T, Uchiyama M, et al. Excessive daytime sleepiness among the Japanese general population. *J Epidemiol.* 2005;15(1):1-8.
12. Buysse DJ, Hall ML, Strollo PJ, et al. Relationships between the Pittsburgh Sleep Quality Index (PSQI), Epworth Sleepiness Scale (ESS), and clinical /polysomnographic measures in a community sample. *J Clin Sleep Med.* 2008;4(6):563-571.
13. Merlino G, Piani A, Dolso P, et al. Sleep disorders in patients with end-stage renal disease undergoing dialysis therapy. *Nephrol Dial Transplant.* 2006;21(1):184-190.
14. Sabry AA, Abo-Zenah H, Wafa E, et al. Sleep disorders in hemodialysis patients. *Saudi J Kidney Dis Transplant.* 2010; 21(2):300-305.
15. Wise MS. Objective measures of sleepiness and wakefulness: application to the real world? *J Clin Neurophysiol.* 2006;23(1): 39-49.
16. Arzi L, Shreter R, El-Ad B, Peled R, Pillar G. Forty- versus 20-minute trials of the maintenance of wakefulness test regimen for licensing of drivers. *J Clin Sleep Med.* 2009;5(1):57-62.
17. Sunwoo BY, Jackson N, Maislin G, Gurubhagavatula I, George CF, Pack AI. Reliability of a single objective measure in assessing sleepiness. *Sleep.* 2012;35(1):149-158.
18. Basner M, Dinges DF. Maximizing sensitivity of the psychomotor vigilance test (PVT) to sleep loss. *Sleep.* 2011;34(5):581-591.
19. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep.* 1991;14(6):540-545.
20. Bloch KE, Schoch OD, Zhang JN, Russi EW. German version of the Epworth Sleepiness Scale. *Respiration.* 1999;66(5): 440-447.
21. Zallek SN, Redenius R, Fisk H, Murphy C, O'Neill E. A single question as a sleepiness screening tool. *J Clin Sleep Med.* 2008;4(2):143-148.
22. Strober LB, Arnett PA. An examination of four models predicting fatigue in multiple sclerosis. *Arch Clin Neuropsychol.* 2005;20(5):631-646.
23. Hajak G, Landgrebe M. Time and depression: when the internal clock does not work. In: Mendlewicz J, ed. In: *Search of Lost Time: A New Dimension in Antidepressant Efficacy.* Vol 32. Brussel, Belgium: Medicographia; 2010.
24. Riemann D, Spiegelhalder K, Feige B, et al. The hyperarousal model of insomnia: a review of the concept and its evidence. *Sleep Med Rev.* 2010;14(1):19-31.
25. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV).* 4 ed. Arlington, VA: American Psychiatric Publishing; 2000.
26. Dobbels F, Skeans MA, Snyder JJ, Tuomari AV, Maclean JR, Kasiske BL. Depressive disorder in renal transplantation: an analysis of Medicare claims. *Am J Kidney Dis.* 2008;51(5): 819-828.
27. Eryilmaz MM, Ozdemir C, Yurtman F, Cilli A, Karaman T. Quality of sleep and quality of life in renal transplantation patients. *Transplant Proc.* 2005;37(5):2072-2076.
28. Burkhalter H, Sereika SM, Engberg S, Wirz-Justice A, Steiger J, De Geest S. Sleep quality description based on the Pittsburgh Sleep Quality Index in renal transplant recipients. In: Weimar W, Bos MA, Busschbach JJ, eds. *Organ Transplantation: Ethical, Legal and Psychosocial Aspects. Expanding the European Platform.* Vol 2. Lengerich, Germany: PABST Science Publishers; 2011.
29. Kachuee H, Ameli J, Taheri S, et al. Sleep quality and its correlates in renal transplant patients. *Transplant Proc.* 2007; 39(4):1095-1097.

30. Jacobsen JH, Shi L, Mokhlesi B. Factors associated with excessive daytime sleepiness in patients with severe obstructive sleep apnea. *Sleep Breath.* 2013;17(2):629-635.
31. Maglione JE, Ancoli-Israel S, Peters KW, et al. Depressive symptoms and subjective and objective sleep in community-dwelling older women. *J Am Geriatr Soc.* 2012;60(4): 635-643.
32. Mume CO. Excessive daytime sleepiness among depressed patients. *Libyan J Med.* 2010;5(4626).
33. AERA, APA, NCME. *Standards for Educational and Psychological Testings.* Washington, DC 20036: American Educational Research Association; 1999.
34. Moss PA. Reconstructing validity. *Educ Res.* 2007;36(8): 470-476.
35. Research ANd Development (RAND). Kidney Disease and Quality of Life Short Form (KDQOL-SF). 1995. http:// www .rand.org. Accessed December 21, 2012.
36. Fawcett T. An introduction to ROC analysis. *Pattern Recognit Lett.* 2006;27(8):861-874.
37. Gove PB, ed. *Webster's Third New International Dictionary of the English Language.* Springfield, Mass: Philip Babcock Gove; 2008.
38. Johns MW. Reliability and factor analysis of the Epworth Sleepiness Scale. Sleep. 1992;15(4):376-381.
39. Johns MW. Sleepiness in different situations measured by the Epworth Sleepiness Scale. *Sleep.* 1994;17(8):703-710.
40. Buysse DJ, Reynolds CF 3rd, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res.* 1989; 28(2):193-213.
41. Backhaus J, Junghanns K, Broocks A, Riemann D, Hohagen F. Test-retest reliability and validity of the Pittsburgh Sleep Quality Index in primary insomnia. *J Psychosom Res.* 2002;53(3):737-740.
42. Lovibond PF, Lovibond SH. The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behav Res Ther.* 1995;33(3):335-343.
43. Lovibond SH, Lovibond PF. *Manual for the Depression Anxiety Stress Scales.* 2nd ed. Sydney: Psychology Foundation of Australia; 1995.
44. Derogatis LR. *Brief Symptom Inventory.* Baltimore, MD: Clinical Psychometric Research; 1975.
45. Crawford JR, Henry JD. The Depression Anxiety Stress Scales (DASS): normative data and latent structure in a large non-clinical sample. *Br J Clin Psychol.* 2003;42(pt 2):111-131.
46. Kimberlin CL, Winterstein AG. Validity and reliability of measurement instruments used in research. *Am J Health Sys Pharm.* 2008;65(23):2276-2284.
47. Arand DL, Bonnet MH. The multiple sleep latency test and maintenance of wakefulness test: use and procedures. In: Pandi-Perumal SR, Verster JC, Monti JM, Lader M, Langer SZ, eds. *Sleep Disorders: Diagnosis and Therapeutics.* London, United Kingdom: Informa Healthcare; 2008:658-668.
48. Messick S. Validity of Psychological Assessment. *Am Psychol.* 1995;50(9):741-749.
49. DeVellis RF. *Scale Development: Theory and Applications.* 2nd ed. Thousand Oaks, CA: Sage Publications; 2003.
50. Polit DF, Beck CT. The content validity index: are you sure you know what's being reported? Critique and recommendations. *Res Nurs Health.* 2006;29(5):489-497.
51. Breslau N, Roth T, Rosenthal L, Andreski P. Sleep disturbance and psychiatric disorders: a longitudinal epidemiological study of young adults. *Biol Psychiatry.* 1996;39(6):411-418.
52. Pinsky BW, Takemoto SK, Lentine KL, Burroughs TE, Schnitzler MA, Salvalaggio PR. Transplant outcomes and economic costs associated with patient noncompliance to immunosuppression. *Am J Transplant.* 2009;9(11):2597-2606.
53. Besedovsky L, Lange T, Born J. Sleep and immune function. *Pflugers Arch.* 2012;463(1):121-137.
54. Razonable RR, Eid AJ. Viral infections in transplant recipients. *Minerva Med.* 2009;100(6):479-501.
55. Swisstransplant. Swisstransplant Year Report. 2011. http:// www.swisstransplant.org/l1/. Accessed December 21, 2012.