Validity and reliability of the french version of the start back screening tool for patients with low back pain

Bruyère, Olivier; Demoulin, Maryline; Beaudart, Charlotte; Hill, Jonathan C.; Maquet, Didier; Genevay, Stéphane; Mahieu, Geneviève; Reginster, Jean Yves; Crielaard, Jean Michel; Demoulin, Christophe

Published in:
Spine

DOI:
10.1097/BRS.0000000000000062

Publication date:
2014

Document Version
Early version, also known as pre-print

Validity and reliability of the French version of the STarT Back Screening Tool for patients with low back pain

O. Bruyère (1, 2, 3), M. Demoulin (1), C. Beaudart (1, 3), J.C. Hill (4), D. Maquet (2), S. Genevay (5), G. Mahieu (6), J.-Y. Reginster (1, 3), J.-M. Crielaard (2) and C. Demoulin (2)

(1) *Department of Public Health, Epidemiology and Health Economics, University of Liege, Belgium*

(2) *Department of Sport and Rehabilitation Sciences, University of Liege, Belgium*

(3) *Reseau d'Epidémiologie Clinique International Francophone (RECIF), Lyon, France*

(4) *Arthritis Research UK Primary Care Centre, Keele University, Stoke-on-Trent, UK*

(5) *Rheumatology Department, Geneva University Hospital, Switzerland*

(6) *Physical Rehabilitation, Hospital of Dinant, Belgium*

Corresponding author:

Olivier BRUYERE, PhD

University of Liège, Department of Public Health, Epidemiology and Health Economics

CHU Sart-Tilman, Bât B23, 4000 Liège, Belgium

Tel.: +32 (0)4 366 25 81, Fax: +32 (0)4 366 28 12

E-mail: olivier.bruyere@ulg.ac.be

Abstract

Objective: The STarT Back Tool (SBT) is a recently validated tool developed to identify subgroups of patients with low back pain (LBP) to guide early secondary prevention in primary care. Our objective was to assess the reliability and validity of the French version of the SBT.

Methods: Outpatients with LBP aged between over 18 years attending rehabilitation centre, back school, private physiotherapists unit or fitness centre were included. The SBT, the Roland-Morris Disability Questionnaire (RMDQ), the Orebro Musculoskeletal Pain Screening Questionnaire (ÖMPSQ), the Medical Outcomes Survey Short Form 36 (SF-36) questionnaire and a pain visual analogic scale (VAS) were completed by patients. Test-retest reliability was assessed with Kappa score or the intraclass coefficient correlation (ICC), internal consistency of the psychological subscale with the Cronbach α coefficient, construct validity with the Spearman coefficient correlation and floor and ceiling effects by percentage frequency of lowest or highest possible score achieved by respondents.

Results: 108 patients with LBP were included. The reliability of the total score was excellent between the test and the retest, with an ICC of 0.90 (0.81-0.95). The Cronbach α coefficient was 0.73 showing a good internal consistency for the psychological subscale. High correlation coefficients of 0.74 between SBT and RMDQ and 0.74 between the SBT and ÖMPSQ were observed. As expected, low to moderate correlations were observed between the SBT total score and some dissimilar measures of the SF-36. The lowest possible SBT score was found in 8 patients (7.4%) whereas only three patients (2.8%) had the highest possible SBT score.

Conclusions: The French version of the SBT is a reliable and valid questionnaire as is the original English version. Therefore, this new version may help French-speaking clinicians and scientists to stratify patient with LBP.

Keywords: low back pain, questionnaire, psychometric properties, validity, reliability

INTRODUCTION

Low back pain (LBP) is the most prevalent and costly musculoskeletal problem in today's economically advanced societies, and may lead to long-term disability combined with frequent use of health services (1). Although LBP is a common condition, the outcomes of physical therapy care for patients with LBP appear to be variable and, at times, suboptimal, with many patients failing to experience significant reductions in pain and disability. Considering the prevalence of patients receiving therapy for LBP, improving the care provided for this condition could have a significant impact on the overall quality of care provided (2). Therefore, it is of high importance to identify patients at risk for developing persisting LBP at an early stage. To detect these patients, prognostic factors of chronicity must be known. Screening instruments are needed to assess these influencing factors and to foretell the course of LBP. Several back pain screening tools exist to aid clinicians in identifying patients either 'at risk' of chronicity or to improve targeting of treatment (1).

The Keele STarT Back Tool (SBT) is a recently validated tool developed to identify subgroups of patients needing specific rehabilitation in early secondary prevention in primary care (3). The conceptual purposes of the SBT (Hill et al., 2008) were to identify patients with potentially treatment modifiable prognostic indicators using a brief, user-friendly tool and to validate cut-off scores for subgrouping patients into 1 of 3 a priori initial treatment options in primary care. Currently, the SBT is translated in nearly twenty languages. Recently, the French translation and cross-cultural adaptation of the SBT has been validated (4). However, besides the validation of the translation in itself, it is highly recommended that, after the translation and adaptation process, the investigators ensure that the new version has appropriate psychometric properties needed for the intended application (5). Indeed an instrument should retain both the item-level characteristics such as item-to-scale correlations and internal consistency, as well as reliability and construct validity at the score-level.

The objective of the present study is to assess, over a large number of subjects, the reliability and the validity of the previous cross-culturally adapted French version of the SBT.

METHODS

Patients

Outpatients attending a rehabilitation centre, a back school, six private physiotherapists units or two fitness centres were enrolled between July and September 2012. The inclusion criteria were patients with chronic nonspecific LBP (i.e., lasting more than 3 months) aged at least 18 years. Our study was accepted by the Ethical committee of the University of Liege and written consent was signed by all patients.

Outcome assessments

The SBT

The SBT is a validated questionnaire containing 9-items; they were selected as predictive of 'poor prognosis' following a literature review and a secondary analysis to identify strong independent predictors for persistent disabling back pain (3). All items use a dichotomised response format ('disagree' = 0 point, 'agree' = 1 point), except for the item related to bothersomeness, which uses a 5-point Likert scale. The overall SBT score ranges from 0 to 9 and is produced by summing all positive items; a psychosocial subscale score ranging from 0 to 5 is produced by summing bothersomeness, fear, catastrophising, anxiety, and depression items (items 5 to 9).

Pain intensity

Pain intensity was assessed by means of a self-administered visual analogic scale (VAS) ranging from 0 (no pain at all) to 10 (the worst imaginable pain) (6).

The Roland-Morris Disability Questionnaire (RMDQ)

The RMDQ is a self-administered specific questionnaire designed to assess disability caused by LBP consisting of 24 yes-no items. The total score ranges from 0 (no disability) to 24 (maximum disability) (7). The validated French version of the RMDQ (i.e. the EIFEL questionnaire) was used in the present study (8).

The Orebro Musculoskeletal Pain Screening Questionnaire (ÖMPSQ)

The ÖMPSQ consists of 24 self-report items (21 items are scored); they were selected following a literature review to identify strong independent risk factors for work absence (9). The authors defined 'poor prognosis' as the accumulation of 30 days or more of sick leave at a six months follow-up. The ÖMPSQ's 21 scored items use an 11-point response format, apart from item 1 (pain sites), which has five descriptive components that are double weighted. The instrument therefore provides a potential score ranging from two to 210 points. The French version of the tool was used in this study (10).

The Medical Outcomes Survey Short Form 36 (SF-36) questionnaire

The SF-36 is a widely used generic outcome measure, which consists of eight domains; physical function, role-physical, pain, general health, vitality, social functioning, role-emotional and mental health. The SF-36 is self-explanatory and takes about 10 min to complete. The SF-36 score ranges from 0 to 100, 0 indicating extreme problems and 100 indicating no problem. We used the validated French version of the questionnaire (11)

Psychometric scale properties

Sample size

As suggested by Terwee et al. (12), i.e., the study sample size consisted of at least 50 patients for reliability and construct validity and of at least 100 responders for internal consistency.

Test-retest reliability

Test-retest stability was analysed by asking the patients to fill in again the questionnaire 2 weeks later. This test was only performed among the patients that reported no change in back pain over the week to the question “did you experience some change in you low back pain during the past two weeks?”. The intraclass coefficient correlation (ICC) was used to test the reliability between the baseline and retest scores and the Kappa test was used to evaluate item-by-item agreement. Values between 0.60 and 0.80 indicated good reliability whereas values higher than 0.80 indicated excellent reliability. Ninety-five percent confidence intervals (CI) were calculated for ICC values.

Internal consistency

Internal consistency is the estimation of item homogeneity. Items of the scale should tap various aspects of a unique trait, not different attributes. The internal coherence was examined using Cronbach α, which was estimated for the psychosocial subscale of the questionnaire, with a value of more than 0.70 being considered as acceptable.

Construct validity

Construct validity represents the extent to which the results of the questionnaire are related to the theoretical concept to be measured. Construct validity includes the degree of correlation between an instrument and other measures that assess similar concepts (e.g. the

SBT versus the VAS, the RMDQ or the ÖMPSQ = convergent validity) and the divergence from measures that are dissimilar (SBT versus some items of the SF-36 = discriminant validity). There is no consensus in the literature on the criterion to determine when two measures should be considered correlated. Validity was evaluated by the Spearman correlation coefficient according to the following criteria: excellent (r = 0.81–1.0), very good (r = 0.61–0.80), good (r = 0.41–0.60), acceptable (r = 0.21–0.40), and fair (r = 0–20).

Floor and ceiling effects

The floor and ceiling effects were analysed by calculating the percentage frequency of lowest or highest possible score achieved by respondents. Floor and ceiling effects higher than 15% were considered to be significant.

Statistical analysis

All statistical analyses have been made using the SPSS 16.0 for Windows. A Shapiro-Wilk test verified the normal distribution for all parameters. Quantitative variables that were normally distributed were expressed as mean ± standard deviation (SD) and quantitative variables that were not normally distributed were expressed as median (percentile 25, percentile 75). Qualitative variables were reported as absolute and relative frequencies (%).

RESULTS

Subjects

The study included 108 patients with LBP: 60 women (55.6%) and 48 men (44.4%) with a mean age of 49.5 years (range: 34-59). A substantial proportion (75.9%) of the population experienced LBP for more than 6 months. Table 1 summarizes other clinical characteristics. All questionnaires were correctly answered without missing data.

Test-retest reliability

Among the whole study population, 60 were asked to fill in again the SBT one week after the first one. Among the patients that reported no change in back pain over the week (n=35), there was an excellent agreement between the test and the retest total score, with an ICC of 0.90 (95% CI 0.81-0.95). Item-by-item agreement as well as psychosocial subscale score agreement is presented in Table 2.

Internal consistency

The Cronbach α coefficient was 0.74 for the psychosocial subscale, showing a good internal consistency.

Construct validity

For the convergent validity, very good correlation coefficients (0.74) were found between SBT and RMDQ and between SBT and ÖMPSQ, as shown in Table 3. As expected, correlation coefficients were also high, between the item 3 of the SBT and the item 17 of the RMDQ (r=0.71) as well as between the item 4 of the SBT and the item 9 of the RMDQ (r=0.70). The very low correlation coefficients found between the total score of the SBT and the vitality, role-emotional and mental health scores of the SF-36 (r=-0.28, -0.31, -0.32, respectively) confirmed the discriminant validity of the SBT.

Floor and ceiling effects

The lowest and highest possible SBT scores were found only in 8 (7.4%) and in 3 patients (2.8%), respectively. For the psychological subscale lowest and highest possible scores were found only in X (X.X%) and in X patients (X.X%), respectively.

DISCUSSION

The objective of the present study was to assess the reliability and the validity of the French version of the SBT. According to the present study, the test-retest reliability of the French translation of the SBT is excellent when considering the measure of the whole score (ICC>0.900) and the psychosocial subscale (ICC>0.800).This is in line with the reliability data published about the original validation of the tool (3). The internal consistency of the SBT psychological subscale is good (Cronbach α coefficient of 0.74) and also in line with the validated paper of the original tool. The convergent validity of the French version of the SBT was analysed by comparing it with other back pain screening tools. As expected, we found a very good correlation between the SBT and both the RMDQ and the ÖMPSQ. High correlations were also observed between similar items of the SBT and the RMDQ. The moderate correlation observed between the SBT and some health dimensions of the SF-36 (e.g. role-physical, physical function, general health) suggests that the theoretical construct of the measures are not completely the same. The divergent validity was shown by the low correlation observed between the SBT and the vitality, role-emotional and mental health scores of the SF-36. These results were as expected because these tools are supposed to measure different conceptual constructs.

Back specific instruments, such as the RMDQ and the ÖMPSQ but also the Core Outcome Measures Index (13), the Quebec back pain disability scale (14) and the Dallas Pain Questionnaire (15) have been cross-culturally adapted to French. However, there is no “gold standard” for which questionnaire should be used, and different doctors and physicians have their own choices of the questionnaire. Although the SBT has been validated recently, SBT is increasingly used and seems to be of potential interest for the management of patients with LBP (2, 16-21). Indeed, this tool has the specific interest to classify patients according to the risk of chronicity. A recent large randomised controlled trial involving 851 adults followed for 12 months showed the clinical effectiveness and cost-effectiveness of stratified primary care

(using the SBT questionnaire) compared with non-stratified current best practice (22). The results demonstrated that the stratified care approach significantly reduced levels of disability and was cost-saving compared to the current best practice management approach. Unfortunately, the SBT has only been fully validated in English (original version) (3), Danish (23), and Spanish (24).

The lack of validated translations of the SBT making difficult the comparison of our results with others. Moreover, it should be acknowledged that there is no clear recommendation for the assessment of reliability and validity of a translated questionnaire to the opposite of the translation process where clear recommendations exist (5). It should also be pointed out that, as there is no gold standard screening tool and because all these tools are not validated in every language, authors wishing to validate a translated questionnaire often used their own reference tool. Furthermore, studying the validity of such questionnaire remains difficult considering the absence of a “gold standard”. However, in our study, we have used the most widely used tools already validated in French. At last, the statistics used in the present study to assess reliability and validity are the most widely used in the literature. Moreover, as recommended (12), our study population included at least the requested 50 subjects for reliability and construct validity and at least 100 subjects for internal consistency.

In conclusion, the French version of the SBT is reliable and valid and can therefore be recommended for clinical and research purposes. The next step will be to assess the effectiveness of this version to identify subgroup of patients at different risk of progression to the chronicity in a French-speaking population.

Competing interests

J. Hill was involved in the development of the initial English questionnaire.

Authors' contributions

OB, MD and CD were involved in the design of the study. OB drafted the manuscript. Statistical analyses were performed by OB, MD and CB. All authors were involved in the interpretation of the results and critically reviewed the manuscript. All authors participated to the interpretation of the data and approved the final version of the manuscript.

References

1. Melloh M, Elfering A, Egli Presland C, Roeder C, Barz T, Rolli Salathe C, et al. Identification of prognostic factors for chronicity in patients with low back pain: a review of screening instruments. International orthopaedics. 2009;33(2):301-13.
2. Fritz JM, Beneciuk JM, George SZ. Relationship between categorization with the STarT Back Screening Tool and prognosis for people receiving physical therapy for low back pain. Physical therapy. 2011;91(5):722-32.
3. Hill JC, Dunn KM, Lewis M, Mullis R, Main CJ, Foster NE, et al. A primary care back pain screening tool: identifying patient subgroups for initial treatment. Arthritis and rheumatism. 2008;59(5):632-41.
4. Bruyere O, Demoulin M, Brereton C, Humblet F, Flynn D, Hill JC, et al. Translation validation of a new back pain screening questionnaire (the STarT Back Screening Tool) in French. Arch Public Health. 2012;70(1):12.
5. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine. 2000;25(24):3186-91.
6. Huskisson EC. Measurement of pain. Lancet. 1974;2(7889):1127-31.
7. Roland M, Morris R. A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. Spine. 1983;8(2):141-4.
8. Coste J, Le Parc JM, Berge E, Delecoeuillerie G, Paolaggi JB. [French validation of a disability rating scale for the evaluation of low back pain (EIFEL questionnaire)]. Rev Rhum Ed Fr. 1993;60(5):335-41.
9. Linton SJ, Hallden K. Can we screen for problematic back pain? A screening questionnaire for predicting outcome in acute and subacute back pain. The Clinical journal of pain. 1998;14(3):209-15.
10. Nonclercq O, Berquin A. Predicting chronicity in acute back pain: validation of a French translation of the Orebro Musculoskeletal Pain Screening Questionnaire. Ann Phys Rehabil Med. 2012;55(4):263-78.
11. Leplege A, Ecosse E, Verdier A, Perneger TV. The French SF-36 Health Survey: translation, cultural adaptation and preliminary psychometric evaluation. Journal of clinical epidemiology. 1998;51(11):1013-23.
12. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. Journal of clinical epidemiology. 2007;60(1):34-42.
13. Genevay S, Cedraschi C, Marty M, Rozenberg S, De Goumoens P, Faundez A, et al. Reliability and validity of the cross-culturally adapted French version of the Core Outcome Measures Index (COMI) in patients with low back pain. European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society. 2012;21(1):130-7.
14. Yvanes-Thomas M, Calmels P, Bethoux F, Richard A, Nayme P, Payre D, et al. Validity of the French-language version of the Quebec back pain disability scale in low back pain patients in France. Joint, bone, spine : revue du rhumatisme. 2002;69(4):397-405.
15. Marty M, Blotman F, Avouac B, Rozenberg S, Valat JP. Validation of the French version of the Dallas Pain Questionnaire in chronic low back pain patients. Revue du rhumatisme. 1998;65(2):126-34.
16. Morso L, Kent P, Albert HB, Manniche C. Is the psychosocial profile of people with low back pain seeking care in Danish primary care different from those in secondary care? Manual therapy. 2013;18(1):54-9.

17. Del Pozo-Cruz B, Gusi N, Del Pozo-Cruz J, Adsuar JC, Hernandez-Mocholi M, Parraca JA. Clinical effects of a nine-month web-based intervention in subacute non-specific low back pain patients: a randomized controlled trial. Clinical rehabilitation. 2013;27(1):28-39.
18. Maas ET, Juch JN, Groeneweg JG, Ostelo RW, Koes BW, Verhagen AP, et al. Cost-effectiveness of minimal interventional procedures for chronic mechanical low back pain: design of four randomised controlled trials with an economic evaluation. BMC musculoskeletal disorders. 2012;13:260.
19. Beneciuk JM, Bishop MD, Fritz JM, Robinson ME, Asal NR, Nisenzon AN, et al. The STarT Back Screening Tool and Individual Psychological Measures: Evaluation of Prognostic Capabilities for Low Back Pain Clinical Outcomes in Outpatient Physical Therapy Settings. Physical therapy. 2012.
20. Field J, Newell D. Relationship between STarT Back Screening Tool and prognosis for low back pain patients receiving spinal manipulative therapy. Chiropractic & manual therapies. 2012;20(1):17.
21. Storheim K. Targeted physiotherapy treatment for low back pain based on clinical risk can improve clinical and economic outcomes when compared with current best practice. Journal of physiotherapy. 2012;58(1):57.
22. Hill JC, Whitehurst DG, Lewis M, Bryan S, Dunn KM, Foster NE, et al. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. Lancet. 2011;378(9802):1560-71.
23. Morso L, Albert H, Kent P, Manniche C, Hill J. Translation and discriminative validation of the STarT Back Screening Tool into Danish. European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society. 2011;20(12):2166-73.
24. Gusi N, del Pozo-Cruz B, Olivares PR, Hernandez-Mocholi M, Hill JC. The Spanish version of the "STarT Back Screening Tool" (SBST) in different subgroups. Atencion primaria / Sociedad Espanola de Medicina de Familia y Comunitaria. 2011;43(7):356-61.

Table 1. Baseline characteristics of the study population

Variables	**n**	
Age (year)	108	49.5 (34.5-59.0)
Women (%)	108	60 (55.6)
Body mass index (kg/m²)	108	25.0 (21.5-27.1)
Duration of low back pain	108	
< 1 month		1 (0.9%)
1-3 months		11 (10.2%)
4-6 months		14 (13.0%)
7 months–3 years		35 (32.4%)
> 3 years		47 (43.5%)
Pain visual analogic scale (0-100)	108	43.2 ± 26.4
STarT Back	108	X
Low risk		51 (47.2%)
Moderate risk		30 (27.8%)
High risk		27 (25%)
Roland-Morris Disability Questionnaire	108	6 (3-11)
Orebro Musculoskeletal Pain Screening Questionnaire	56	84.5 (71-116)
SF-36	56	
SF-36 physical component scale		53.7 ± 21.0
SF-36 mental component scale		60.5 (37.1-75.7)

Table 2. Test-retest reliability of the SBT assessed by calculating the Kappa score and intraclass coefficient correlation (ICC)

Items	Kappa score	IC 95%
1	0.716	0.490-0.943
2	0.815	0.618-1
3	0.798	0.583-1
4	0.873	0.703-1
5	0.547	0.252-0.842
6	0.495	0.194-0.795
7	0.746	0.513-0.980
8	0.659	0.417-0.901
9	0.625	0.362-0.888
	Intraclass coefficient correlation (ICC)	
Psychosocial score	0.815	0.664-0.902
Total score	0.900	0.812-0.948

Table 3. Correlations between the SBT and the other questionnaires

Questionnaire	Coefficient correlations
pain VAS	0.66
RMDQ	0.74
ÖMPSQ	0.74
SF-36	
Physical function	-0.58
Role-physical	-0.49
Body pain	-0.71
General health	-0.46
Vitality	-0.28
Social functioning	-0.44
Role-emotional	-0.31
Mental health	-0.32
item 17 of the RMDQ and item 3 of the SBT	0.71
item 9 of the RMDQ and item 4 of the SBT	0.70