Abstract

**Background** The STarT Back Tool (Subgrouping for Targeted Treatment; SBT) was developed and validated in the United Kingdom for adults with non-specific low back pain (LBP) to provide risk stratification groups. An Arabic version has not yet been developed. Consequently our objectives were: First, to cross-culturally adapt the SBT for use in Arabic speaking adults (SBT-Ar) with LBP. Second, to assess the face, content and construct validity of SBT-Ar against relevant reference standards.

**Methods** This was a prospective, cross-sectional study carried out in the outpatient department in a tertiary care hospital. A total of 59 participants (aged 18-60) with LBP able to read Arabic completed the questionnaire SBT cross-cultural adaptation was performed according to published guidelines. Face and content validity were explored by individual interviews. Construct validity was assessed using pre-hypothesized correlations with relevant reference standards.

**Results** Following 48 individual interviews the SBT final version was reached and demonstrated face and content validity. The SBT-Ar total score and psychosocial sub-scale had acceptable internal consistency and no redundancy (Cronbach α=0.7). Moderate Spearman's correlations were found between the SBT-Ar total score and reference standards (Arabic Pain Numeric Rating Scale NRS-Ar r=0.50 and Arabic Oswestry Disability Index ODI-ar r=0.51). As expected the SBT-Ar psychosocial subscale had medium to high correlations with the psychosocial reference measures (Arabic Fear-Avoidance Beliefs Questionnaire Physical Activity FABQPA-Ar r=0.41, Arabic Hospital Anxiety and Depression Scale-Anxiety HADSA-Ar r=0.58, Arabic Hospital Anxiety and Depression Scale-Depression HADSD-Ar r=0.45 and Arabic Pain Catastrophizing Scale PCS-Ar r=0.69).The SBT-Ar showed no significant floor or ceiling effects.

**Conclusion** This study culturally adapted and preliminary validated SBT into Arabic.

# INTRODUCTION

Low back pain (LBP) is the most common cause of activity limitation and absence from work throughout the world [1]. Consequently, it imposes a high economic burden on individuals, families, societies and governments [1]. It is a recurring condition with 60-80% of patients in the United Kingdom consulting their physician a year later for the same problem [2]. The prevalence among adults in Saudi Arabia is estimated at 18.8% [3]. Non-specific LBP with no known disease or pathology can range from 20 to 85% [4]. The estimated lifetime prevalence of non-specific LBP is 60-70% in industrialized countries with one-year prevalence of 15-45% and adult incidence of 5% per year [1].

The STarT Back Tool (Subgrouping for Targeted Treatment; SBT) is a self- reported prognostic questionnaire, which can be used by primary care providers [2]. It was validated to identify individuals with LBP in primary care with prognostic indicators for persistent disabling pain [2]. It is a useful component of stratified care, where patient prognostic subgroups are matched with appropriate treatment plans [2].In a randomized controlled trail published in 2011, Hill et al. found that the prognostic stratified care was cost effective and resulted in increased health outcomes when compared with usual care [5]. Health outcomes included significant improvement in disability (at 4 and 12 months), days off work (for 12 months follow up), and patient satisfaction (for 4 months follow up) [5].

Beaton et al. defines cross-cultural adaptation of self-reported measures as “a process that looks at both language translation and cultural adaptation in the process of preparing the questionnaire for another setting” [6]. The SBT was cross-culturally adapted and validated for use for adults with LBP in several countries. These countries include Denmark [7], China [8], Japan [9], Germany [10, 11], Belgium [12, 13], Spain [14], Iran [15], Brazil [16], Finland [17] and Sweden [18]. There is a need for cross-cultural adaptation of the SBT for use in Arabic speaking individuals with LBP. Thus, the aims of this study were: 1) to linguistically and culturally adapt the SBT and 2) to assess its face, content and construct validity for use in adults with LBP in Saudi Arabia.

# MATERIALS AND METHODS

## This study was approved by the International Review Board (IBR) at King Fahad Specialist Hospital Dammam (KFSH-D) and the Research Ethics Board (REB) at the University of Toronto. All participants in this study signed an informed consent form.

### Cross Cultural Adaptation and Assessment of Face and Content Validity

We obtained permission from the author of the original SBT for cross-cultural adaptation for use in Arabic speaking adults with LBP. The process of developing the Arabic Version of the SBT was done according to the following stages described by Beaton et al. [6]:

###### Stage I: Initial Translation

Two translators translated the SBT from English to Arabic. Translator 1 was aware of the concept being examined. A physical therapist performed this translation to create the Translation Version 1 (T1). Translator 2 had no medical background and created the Translation Version 2 (T2).

###### Stage II: Synthesis of the Translations

The two translators met with the presence of an observer and synthesized the T1 and T2 versions to resolve any disagreements in the translations. The Translation 12 (T12) was created at this stage.

###### Stage III: Back Translation

Two translators without a medical background but are native English Speakers back translated the T12 from Arabic to English to create the Back Translation 1 (BT1) and the Back Translation 2 (BT2).

###### Stage IV: Expert Committee

An expert committee was created to review all the translations (T1, T2, T12, BT1, and BT2) and developed the pre-final version of the SBT-Ar. The committee consisted of the Developer of the Tool, an associate professor of physical therapy, an English language professor, a language professional, physical therapists, and the translators involved in the previous stages. Consensus was reached on all items and the committee ensured that the final version was comprehensible at the level of a 12 year old child or a 6th grader.

###### Stage V Pretesting/Pilot of the Final Version

We pretested the questionnaire with a sample of 48 adults with LBP. The participants completed the SBT-Ar and were asked probing questions about what they thought was meant by the question items and responses. Any hesitations or ambiguities were recorded including linguistic and cultural differences.

###### Stage VI Submission of the Document to the Developer

The final SBT-Ar and all versions of the translation, minutes of the meetings, and reports from the previous stages were submitted to the Developer for appraisal of the adaptation process.

### Assessment of the Construct Validity of the Arabic Version of the SBT

Similar to methods used by Hill et al. [2], the developed Arabic version SBT-Ar scores and sub-scales were compared to relevant reference measures, previously developed and validated for use in Arabic speaking individuals. The SBT-Ar has two sub-scales: the Physical and the Psychosocial. The Physical Sub-scale includes the pain and disability items; and the Psychological Sub-scale includes the kinesiophobia, anxiety, and catastrophizing, depression and bothersomeness items. All reference measures were validated for use in Arabic speaking individuals: the Arabic Pain Numeric Rating Scale (NRS-Ar) (for pain) [19, 20], the Arabic Oswestry Disability Index (ODI-Ar) (for disability) [21, 22], the Arabic Fear Avoidance-Beliefs Questionnaire- Physical Activity (FABQPA-Ar) (for kinesiophobia) [23], the Arabic Hospital Anxiety and Depression Scale (HADS-Ar) (for anxiety and depression) [24-26] and the Arabic Pain Catastrophizing Scale (PCS-Ar) (for catastrophizing) [27].

## Recruitment and Participants

Participants were recruited at KFSH-D in Saudi Arabia from Out- Patient Departments: Physical Therapy and Medical /Surgical Clinics. Data were collected between April 1st and June 23rd, 2016. A participant was eligible for the study if he/she met the following criteria: a) was an adult, aged 18 to 60; b) had a clinical diagnosis of acute, sub-acute or chronic LBP with or without radiculopathy [5]; and c) was able to read and understand Arabic with at least elementary school educational level. Participants with red flags (e.g. cauda equina syndrome, tumor, infection or inflammatory disease) and/or previous spinal surgery in the past 12 months were excluded [5].

According to the Beaton et alcross-cultural adaptation guidelines, pretesting of face and content validity of the final version of SBT-Ar should be conducted with a sample size of 30 to 40 [6].

The SBT total score correlates moderately with the Oswestry Disability Index (ODI) (Spearman’s Correlation r=0.39) [17]. Based on a 2-sided alpha value of 0.05 and a beta value of 0.2, the sample size for the construct validity is recommended as, 47 to 62 by Hulley et al [28].

## Data Analysis

Statistical analysis was conducted using Statistical Package for Social Sciences Statistics Software (Version 20). Descriptive statistics were reported and consisted of mean and standard deviation if the data were normally distributed. Otherwise, median and interquartile range (IQR) were reported. A pairwise deletion or available-case analysis was used to deal with missing data.

We examined the internal consistency (Cronbach α) of the SBT-Ar total score and psychosocial sub-score to measure the extent to which of items correlate with each other, and hence measure the same concept or construct. Since the SBT-Ar has sub-scales measuring different constructs, acceptable internal consistency is expected. A high Cronbach α value would suggest there was a redundant item within the measure. The acceptable value of Cronbach α is 0.7 to 0.9 [29]. Construct validity was evaluated by assessing expected correlations between the SBT-Ar Total Score, the Pain and Disability Sub-scale score, the Psychosocial Sub-scale score, and the risk groups with the relevant reference standards, using the Spearman’s Correlation Coefficient, because the data were not normally distributed. Correlation (r) levels were defined as; high (r ≥ 0.6), moderate (0.6 > r ≥ 0.30), or low (r < 0.30) [30]. An alpha level (or type I error level) of 0.05 was used to determine statistical significance. Based on the findings of the French [12], Finnish [17] and German [10] studies, it was hypothesized that: a) mean SBT-Ar Total and the Pain and Disability Sub-scale scores would have a moderate to high correlation with the Pain NRS-Ar and ODI-Ar with scores going in the same direction; b) mean SBT-Ar Psychosocial Sub-scale scores would have moderated to high correlation with the FABQPA-Ar, HADS-Ar and PCS-Ar; and c) the SBT-Ar risk groups would have a positive correlation with the reference standards. Floor and ceiling effects of the SBT-Ar were considered present if 15% of the participants had minimum or maximum scores [11].

**RESULTS**

### Cross Cultural Adaptation and Assessment of Face and Content Validity

The forward translators met to resolve minor linguistic discrepancies between the two versions in the directional statement, items 1, 5 and 9. The Expert Committee reviewed all the versions of the forward and backward translations (T1, T2, T12, BT1 & BT2) and reached a consensus on the pre-final version of the SBT-Ar. The back translations were compared to each other and to the original tool. This resulted in some changes to simplify the directional statement to make it comprehensible at the level of a 12 year old child or a 6th grader. All the statements in both back translations were similar except item 6 (Anxiety). The item was retranslated until a consensus was reached.

For the face and content validity, 48 individuals with LBP filled out the questionnaire, followed by individual interviews. Figure 1 shows the Recruitment Flow Diagram. The following ambiguities in items 5 and 8 were reported: 1) Item 5 (Fear-Avoidance), "It's not safe" with the Arabic literal translation was interpreted as a security issue. Therefore, it was replaced by “It is dangerous". In addition, another synonym was used for the word “physically” because it was not clear. 2) Item 8 (Depression) was misinterpreted as a matter of ‘ability’ and not ‘enjoyment’. When asked probing questions, some participants explained that they agreed with the item when they were unable to walk or sit because of their pain. They did not understand that the question was asking about enjoyment and not ability. Thus, we adapted the item by adding examples such as “enjoyment of eating, visiting friends, hobbies and work”. See Figure 2 for a graphic representation of the Stages of Cross- Cultural Adaptation of the SBT-Ar and Figure 3 represents the final version of the SBT-Ar.

## *Internal Consistency*

There was no redundancy in the SBT-Ar Total items with a Cronbach α coefficient of 0.07. The Psychosocial sub-score showed acceptable internal consistency with a Cronbach α coefficient of 0.07.

***Construct Validity***

The study sample included 59 participants. Figure 1 shows the Recruitment Flow Diagram. The baseline characteristics are shown in Table 1 and scores of the SBT-Ar and the reference measures are reported in Table 2.

We found significant correlations between the SBT-Ar Sub-scores and the relevant reference standards, ranging from 0.33 to 0.71 (see Table 3). While Table 4 shows a comparison of the percentage of positive responses to SBT and the SBT-Ar items, Table 5 shows a comparison of the total sub-scores and the percentage risk groups. In addition, the Kruskal Wallis test showed significant difference between the risk groups (p= 0.05). This confirms the ability of the SBT-Ar to stratify the patients into low, medium and high risk groups.

There was no significant floor or ceiling effect. Only 2 (3.4%) participants scored the minimum

score “0” and one participant (1.7%) scored the maximum score “9”.

# DISCUSSION

We have cross-culturally adapted and completed a preliminary validation of the SBT for use in Arabic Speaking Individuals. The face and content validation of the pre-final version of the SBT-Ar identified some ambiguities. The tool was adapted to its final version, and showed acceptable face and content validity. There was no redundancy in the SBT-Ar Total items. The Psychosocial Sub-scale internal consistency was acceptable. We established the construct validity of the newly developed SBT-Ar by demonstrating a significant correlation to the following reference measures: pain NRS-Ar, ODI-Ar, FABQPA-Ar, HADS-Ar and PCS-Ar.

With respect to face and content validity, some linguistic issues came up during the process. Similar linguistic issues were common among other studies including the Danish [7], Chinese [8], Japanese [9], German [10, 11], French [12, 13], Spanish [14], Iranian [15], Brazilian [16], Finnish [17] and Swedish [18].In our study, we modified items 5 (Fear Avoidance) and 8 (Depression), as the items were not clearly understood. In other studies similar issues were noted. For example, in the Danish [7] and German [10] study, changes were made to item 5 (Fear Avoidance) to clarify the items. The participants in the French version study wondered if items 6 (Anxiety) and 8 (Depression) were related to LBP or to general health. The items were adapted to ensure they were LBP-related [13].

Similar to other validation studies of the SBT, we found significant correlations between the SBT-Ar Total Score, Psychosocial Sub-scale Score, Risk Groups and the relevant outcome measures. Correlations were moderate between: the SBT-Ar Pain and Disability Subscore with the ODI-Ar (r=0.33); SBT-Ar Psychosocial Subscore with the FABQ-PA (r=0.41), HADS-Ar (anxiety scale) (r=0.58) and HADS-Ar (depression scale) (r=0.45). In the German version, there were correlations ranging from r=0.35 to r= 0.56 between the German SBT, its psychosocial sub-score, risk groups and other measures the Visual Analogue Scale (VAS), the Ronald Morris Disability Questionnaire (RMDQ), the Tampa Scale of Kinesiophobia (TSK), the Pain Catastrophizing Scale (PCS), the Hospital Anxiety and Depression Scale (HADS) and the 2- item Patient Health Questionnaire (PHQ-2) [10]. The French version of SBT was found to have good convergence validity when compared to other back pain questionnaires. High correlation was found between the French version, and the RMDQ (r=0.74) and the Orebo Musculoskeletal Pain Screening Questionnaires (OMPSQ) (r=0.74) [12]. Strong correlation (r=0.81) was found between the Iranian Version and the ODI confirming convergent validity [15]. The Finnish SBT version and its Psychosocial Sub-score correlated moderately with the reference scales, intensity of LBP, intensity of leg pain, ODI, OMPSQ and Beck Depression Inventory (r= 0.31 to 0.45) except a weak correlation between the psychosocial subscale and the ODI (r=0.29) [17]. Finally there was a strong correlation between SBT Swedish Version and its psychosocial subscale with the OMPSQ-Short r=0.61 and r=0.60 respectively [18]. A wide range of correlations was found which is possibly due to the fact that different set of reference measures were used in each study. However, all the studies including this one found significant correlations which indicated that the cross-culturally SBT is valid in all the different languages.

This study has a major strength, all the reference measures used for the construct validation were previously validated for use in Arabic speaking individuals. On the other hand, this study has some limitations. Our sample was recruited at KFSH-D, which is a tertiary care center serving individuals with complex medical diseases. This sampling may explain the higher percentage of the medium risk group for the SBT-Ar 44% versus 34% for the SBT. Therefore, our tool may not be generalizable for use in primary or secondary health care center populations.

# Conclusion

The current study found that the SBT-Ar demonstrated good face, content and construct validity. Clinicians working with Arabic speaking individuals with LBP are now able to use the SBT-Ar in their practice. Future multicenter (primary, secondary and tertiary) studies are needed to further investigate the psychometric properties of the SBT-Ar including predictive validity and reliability.

**Conflicts of Interest and Sources of Funding:** The authors declare no conflict of interest and no funds were received in support of this work.

The protocol for this study was approved by the Research Ethics Board (REB) at the University of Toronto, Canada and the International Review Board (IRB) at King Fahad Specialist Hospital- Dammam, Saudi Arabia.

Permission for cross-cultural adaptation and validation of the STarT Back Screening Tool was obtained from Jonathan C Hill, Keele University.

#

#

#  REFERENCES

[1] World Health Organization. Low back pain. World Health Organization.2013 [cited 2015 15 August, 2015]. Available from: <http://www.who.int/medicines/areas/priority_medicines/Ch6_24LBP.pdf>.

[2] Hill JC, Dunn KM, Lewis M, Mullis R, Main CJ, Foster NE, Hay EM. A primary care back pain screening tool: identifying patient subgroups for initial treatment. Arthritis Rheum 2008 May;59(5):632-41.

[3] Al-Arfaj AS, Al-Saleh SS, Alballa SR, Al-Dalaan AN, Bahabri SA, Al-Sekeit MA, Mousa MA. How common is back pain in Al-Qaseem region. Saudi Med J 2003 Feb;24(2):170-3.

[4] Riihimaki H. Low-back pain, its origin and risk indicators. Scand J Work Environ Health 1991;17(2):81-90.

[5] Hill J, Whitehurst D, Lewis M, Bryan S, Dunn K, Foster NE, Konstantinou K, Main CJ, Mason E, Somerville S, Sowden G, Vohora K, Hay EM. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. Lancet 2011 Oct;378(9802):1560-71.

[6] Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine 2000 Dec;25:3186- 91.

[7] Morso L, Albert H, Kent P, Manniche C, Hill J. Translation and discriminitive validation of the StarT Back Screening Tool into Danish. Eur Spine J 2011 Dec;20:2166- 73.

[8] Luan S, Min Y, Li G, Lin C, Li X, Wu S, Ma C, Hill JC. Cross-cultural Adaptation, Reliability, and Validity of the Chinese Version of the STarT Back Screening Tool in Patients With Low Back Pain. Spine 2014 Jul;39(16):E974-E9.

[9] Matsudaira K, Oka H, Kikuchi N, Haga Y, Sawada T, Tanaka S. Psychometric properties of the Japanese Version of the STarT Back Tool in patients with low back pain. PLoS ONE 2016 Mar;11(3):e0152019.

[10] Aebischer B, Hill JC, Hilfiker R, Karstens S. German translation and cross-cultural adaptation of the STarT back screening tool. PLoS ONE 2015 Jul;10(7).

[11] Karstens S, Krug K, Hill JC, Stock C, Steinhaeuser J, Szecsenyi J, Joos S. Validation of the German version of the STarT-Back Tool (STarT-G): a cohort study with patients from primary care practices. BMC Musculoskelet Disord 2015 Nov;16:346.

[12] Bruyère O, Demoulin M, Beaudart C, Hill JC, Maquet D, Genevay S, Mahieu G, Reginster JY, Crielaard JM, Demoulin C. Validity and reliability of the french version of the start back screening tool for patients with low back pain. Spine 2014 Jan;39(2):E123-E8.

[13] Bruyère O, Demoulin M, Brereton C, Humblet F, Flynn D, Jonathan CH, Marquet D, Van Beveren J, Reginster J-Y, Crielaard JM, Demoulin C. Translation validation of new back pain screening questionnaire (the STarT Back Screening Tool) in French. Arch Public Health 2012 Jun;70(12).

[14] 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. Aten Primaria 2011 Jul;43(7):356-61.

[15] Azimi P, Shahzadi S, Azhari S, Montazeri A. A validation study of the Iranian version of STarT Back Screening Tool (SBST) in lumbar central canal stenosis patients. J Orthop Sci 2014 Mar;19(2):213-7.

[16] Pilz B, Vasconcelos RA, Marcondes FB, Lodovichi SS, Mello W, Grossi DB. The Brazilian version of start back screening tool - translation, cross-cultural adaptation and reliability. Braz J Phys Ther 2014 Sep-Oct;18(5):453-61.

[17] Piironen S, Paananen M, Haapea M, Hupli M, Zitting P, Ryynanen K, Takala EP, Korniloff K, Hill JC, Hakkinen A, Karppinen J. Transcultural adaption and psychometric properties of the STarT Back Screening Tool among Finnish low back pain patients. Eur Spine J 2015 Feb;25(1):287-9.

[18] Betten C, Sandell C, Hill JC, Gutke A. Cross-cultural adaptation and validation of the Swedish STarT Back Screening Tool. Eur J Physiother 2015 Feb;17(1):29-36.

[19] Williamson A, Hoggart B. Pain: a review of three commonly used pain rating scales. J Clin Nurs 2005 Aug;14(7):798-804.

[20] Alghadir AH, Anwer S, Iqbal ZA. The psychometric properties of an Arabic numeric pain rating scale for measuring osteoarthritis knee pain. Disabil Rehabil 2016 Jan;38(24):2392-7.

[21] Guermazi M, Mezghani M, Ghroubi S, Elleuch M, Med AO, Poiraudeau S, Mrabet F, Dammak J, Fermanian J, Baklouti S, Sellami S, Revel M, Elleuch MH. The Oswestry index for low back pain translated into Arabic and validated in a Arab population. Ann Readapt Med Phys 2005 Feb;48(1):1-10.

[22] As Algarni SG, J G Jones, M Guermazi. Validation of an Arabic version of the Oswestry index in saudi Arabia. Ann Phys Rehabil Med 2014 Aug;57:653-63.

[23] Laufer Y, Elheiga-Na'amne BA, Rozen N. Translation and validation of the Arab version of the fear avoidance beliefs questionnaire. J Back Musculoskelet Rehabil 2012 Aug;25(3):201-8.

[24] Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale: An updated literature review. J Psychosom Res 2002 Feb;52(2):69-77.

[25] El-Rufaie OE, Absood G. Validity study of the Hospital Anxiety and Depression Scale among a group of Saudi patients. Br J Psychiatry 1987 Nov;151(5):687-8.

[26] El-Rufaie OE, Absood GH. Retesting the validity of the Arabic version of the Hospital Anxiety and Depression (HAD) scale in primary health care. Soc Psychchiatry Psychiatr Epidemiol 1995 Jan;30(1):26-31.

[27] Huijer HA, Fares S, French DJ. The Development and Psychometric Validation of an Arabic-language Version of the Pain Catastrophizing Scale (PCS-A). Pain Res Manag 2017 Jan:1472792.

[28] Designing clinical research. 4th ed. Hulley SB, editor: Philadelphia, PA : Lippincott Williams & Wilkins; 2013.

[29] Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Edu 2011 Jun;2:53-5.

[30] Andresen EM. Criteria for assessing the tools of disability outcomes research. Arch Phys Med Rehabil 2000 Dec;81(12 Suppl 2):S15-20.

# Figure Captions

Figure 1. Recruitment Flow Diagram

Figure 2. Stages of Cross-Cultural Adaptation of the Start Back Tool Arabic (SBT-Ar)

Figure 3. STarT Back Tool Arabic