Using the STarT back tool: Does timing of stratification matter?

Newell D PhD*, Field J DC MSc and Pollard D BSc

Institution: Anglo European College of Chiropractic

Address of correspondence author*: Dr D Newell, AECC, 14-14 Parkwood Road, Bournemouth, BH5 2DF, Telephone 01202 436207, Fax

E mail: dnewell@aecc.ac.uk

Introduction

Attempts to target treatment for nonspecific LBP patients have proved problematic. Numerous approaches have been devised including exploring patient and condition based characteristics associated with outcomes (Leboeuf-Yde et al, 2009), formal clinical prediction rule construction (Flynn et al, 2002) and a priori screening tools including recently the STarT Back Tool (SBT) (Hill et al, 2008). This device, designed for intended use in clinical practice was based on the knowledge that much nonspecific low back pain appears to include psychological components as well as musculoskeletal aspects to the experience of pain and disability and that these can provide barriers to recovery.

The literature concerning the identification of individuals with nsLBP at risk of not improving and the wherewithal to ameliorate such risks has been disappointing with various authors reporting few and/or inconsistent modifiable baseline prognostic factors in this population. In particular for patients undergoing chiropractic care one of the only robust predictors of outcome has been shown to be early change in symptomatology (Axen et al, 2005; Larsen and LeBoeuf-Yde, 2005; Rubinstein et al, 2008; Newell and Field, 2007). Studies such as Childs et al (2004) that did demonstrate predictive model consistency in patients undergoing physiotherapy were not able to maintain that consistency outside the original study population (Hancock et al, 2008).

Current guidelines for the treatment of nsLBP describe several treatments found to be generally helpful for this condition (NICE, 2009). However they are not helpful in deciding which of these may be more appropriate for any particular patient. The SBT has been developed to help clinicians direct nsLBP patient towards appropriate and cost effective care at their initial presentation. The SBT groups patients into three risk categories termed Low, Medium and High, identifying in turn those most suited for minimal intervention, manual therapy and manual therapy plus psychological . Stratification into targeted treatment using this tool as an initial screen has been shown to result in more favourable outcomes in comparison to treatments as chosen by an experienced physiotherapist (Hill et al, 2011).Because this tool identifies modifiable risk factor there remains the potential that such factors may vary early in treatment and so the risk group categorisation will alther. The SBT is recommended in recent guidance on developing care pathways for back pain and is being increasingly used within the UK and elsewhere (British Pain Society, 2012)

However, following promising initial results some studies have highlighted potential problems when investigating the generalisability and utility of this tool in different nsLBP patient groups and settings.

For example Fritz et al, (2011) looked at the prognostic performance of this tool in nsLBP subjects in a military personnel population and found that although high risk patients displayed higher initial pain and disability scores compared to the other categories there was little if any difference in outcomes at follow up during routine physical therapy treatment.

Similarly, our previous study (Field and Newell, 2012) investigated the prognostic utility of the SBT in a population of nsLBP patients undergoing chiropractic management and found that despite initial and expected baseline differences in severity of patient symptomatology across SBT risk groups, all groups of patients subsequently recovered equally at short, medium and longer term follow up. One of the potential explanations for this result was that it may be more difficult to assess who may recover because of multiple idiosyncratic factors to the patient that are unknown before treatment starts. However, we have suggested, along with others, that early change in symptomatology may provide a better insight into the patient's likely prognostic trajectory (Axen et al, 2005; Bolton and Hurst, 2011).

In this context is is possible that SBT categorisation at baseline may fail to assign appropriate treatment wheas assignment following a short delay may more successfully predict final outcomes than when collected at presentation.

Our question therefore was, in nsLBP patients undergoing manual therapy as provided by chiropractors does administering the SBT post the first treatment provide better prediction of outcomes than administration at initial presentation?

Methodology

Data for this observational, prospective cohort study was collected between 1st February and 17th August 2012 from eleven chiropractic clinics in the UK. These clinics routinely collect clinical outcomes using an automated web based collection system (Care Response; https://www.care-response.com/CareResponse/home.aspx)

Subjects and Procedure

Consecutive patients aged over 16 presenting to one of the clinics with nonspecific LBP (nsLBP) and diagnosed as amenable to chiropractic care are routinely asked, as part of normal practice, to complete pre-examination forms including the Bournemouth Questionnaire (BQ) (Bolton and Breen, 1999). Patients can either complete these forms at the clinic or online before their first visit. For this study only those patients completing these routine forms online were invited to be part of this study.

These participants, via a web page, were presented with background information on the study and a consent form when they completed the pre-examination forms described above. Baseline data consisted of patient characteristics and condition specific parameters as well as the SBT. Two days after the first appointment these subjects were asked via e-mail to complete a second SBT online.

The SBT (Appendix I) contains nine questions related to physical and psychosocial factors that have been identified as strong independent predictors for persistent disabling LBP. SBT overall scores (ranging from 0 to 9) are determined by summing all positive responses and SBT psychosocial subscale scores (ranging from 0 to 5) are determined by summing items related to bothersomeness, fear, catastrophising, anxiety, and depression. Based on overall and psychosocial subscale scoring, the SBT categorizes patients as ‘high-risk’ (psychosocial subscale scores ≥4) in which high levels of psychosocial prognostic factors are present with or without physical factors, ‘medium-risk’ (overall score >3; psychosocial subscale score <4) in which physical and psychosocial factors are present, but not a high level of psychosocial factors, or ‘low-risk’ (overall score 0-3) in which few prognostic factors are present (Hill et al, 2008).

Practitioners were blinded to patients STarT Back scores and their participation in the study and provided chiropractic care as they considered appropriate and not as defined by the SBT categorization.

Outcomes

In these practices, patients who start treatment are emailed outcome assessment questionnaires consisting of the BQ and a Patient's Global Impression of Change (PGIC), at 14, 30 and 90 days following their initial visit. In this study the dichotomised PGIC was the primary outcome measure.

The BQ is a validated patient reported outcome measure (PROM) consisting of seven 11-point numerical rating scales (0–10) each covering a different aspect of the back pain

experience. These were (i)pain; (ii) disability in activities of daily living; (iii) disability insocial activity; (iv) anxiety; (v) depression; (vi) fear avoidance behavior; and (vii) locus of control. Subscales are summed to produce a total BQ score (maximum of 70). (Bolton and Breen, 1999)

Using the Patients' Global Impression of Change (PGIC) Scale (Appendix II), patients are asked 'How would you describe your pain/complaint now, compared to how you were when you completed the questionnaire before your first visit to this clinic?' The scale ranges from 1 (worse than ever) to 7 (very much improved). This outcome was dichotomized for each of the follow up points with improvement being defined by a PGIC response of better or much better (score of ≥6) (Newell and Bolton, 2010).

The BQ and PGIC have been recommended as preferred measurements by the 'Any Qualified Provider Resource Centre' (UK, NHS) for monitoring outcomes in low back pain patients (UK DoH, 2012)

We also collected data on the number of visits completed at each follow up time point.

Analysis

Descriptive statistics were calculated for baseline characteristics as a group and across SBT categories pre and post the initial visit. Comparisons across SBT categories for each stratification point were achieved using a Kruskal Wallace Test for number of visits, pain and total BQ scores, ANOVA for age and Pearson χ^2 for all categorical variables.

To determine any associations between SBT categorisation and the primary outcome univariate logistic regression analysis was carried out using the SBT categorisation as the independent variable and the dichotomised PGIC as dependent variables at each of the follow up time points.

Adjusted models for predicting favourable outcome as defined by the PGIC were constructed with an entry criterion for significant baseline and follow up variables of $p<0.15$ and retention at $p<0.05$ using a binary logistic analysis forward LR procedure. This was carried out for all follow up points.

Descriptive statistics were also used to show the proportion of patients that had changed risk groups in the two days between SBT sub grouping at baseline and SBT sub grouping two days after the initial visit. Odds for improving for those patients that deteriorated, improved or a combination of both (changed) versus those that did not change SBT categorisation post initial treatment were also calculated.

All statistical analyses were performed using statistical software SPSS (v20.0, SPSS Inc., Chicago IL).

Ethics

Ethics for this study were sought and approved by the Research and Ethics subcommittee of the Anglo-European College of Chiropractic.

Results

Seven hundred and forty-nine subjects consented to and filled out baseline questionnaires.

After initial categorisation at baseline, 39%, 37% and 24% were defined as low, medium or high-risk subgroups respectively. Two days after the initial visit, SBT categorisation resulted in near identical figures in each subgroup respectively, although as shown later (Table 6) these individuals may not have been the same from one categorisation to the next.

Just over half of the cohort was female as was the percentage that reported pain for more than 30 days in the year. The duration of the present episode was largely constituted by those presenting with < 1 month or greater than 3 months pain with only 10% between 1 and 3 months duration. Around a third reported pain above the knee while 12% reported pain below the knee (Table 1)

The numbers of treatments received at each of the follow up outcome points compared between SBT categories as defined post visit are shown in Table 2. In general the high-risk patients received a significantly greater mean number of treatments at 14 and 30 days follow up despite the fact that in this study the practitioners were not aware of the SBT categorisation. By 90 days however their were no statistical differences in treatment numbers

Table 3 shows the characteristics of the SBT groups as stratified at the initial and post initial visits. Although there were no differences between the two categorisation points there were, not surprisingly, significant differences between SBT categories within categorisation time points with high risk groups being older, with higher severity scores, more leg pain and a greater proportion of acute presentation.

The clinical progression of these groups over the course of clinical management is similar whether they were stratified by the SBT at initial or post the initial visit (Figure 1). Both in terms of pain and total BQ scores the low risk groups changed the least while the medium and high-risk groups changed the most. Patients categorized by the SBT at 2 different time points (initial and post initial visit) behaved differently across the risk groups in terms of change scores Interestingly it was high risk group patients that displayed the biggest change scores when categorized by the SBT at the initial visit while in contrast the biggest change was seen in the medium risk group when categorized post the initial visit although this was not statistically significant in either case.

Table 4 shows the odds of improvement of patients compared between SBT subgroups as defined at initial and post initial visit time points. Generally there was no difference in the prognostic ability of the SBT regardless of whether categorisation was before or after the initial visit with both explanation of variation in outcome (Nagelkerke) and ability to discriminate between those improving from those that did not (ROC) being below 5% or

below 0.75 respectively. Overall, the medium risk group faired the best in this cohort being around twice as likely to improve than low risk groups. High-risk groups on the other hand were no less likely to improve than low risk groups.

A multivariate analysis of all baseline and follow up PGIC categorisations (Table 5) revealed different predictors for improvement at each of the follow up points. These consisted of shorter duration, absence of pain above the knee and less than 30 days pain in the previous year predicting favourable outcomes at 14 days follow up. In contrast, at 30 days follow up, improvement at 14 days was strongly associated with improvement together with being female and being ranked in the SBT medium risk group 2 days following the initial treatment

At 90 days however, only past improvement at 14 and 30 days were associated with favourable outcomes suggesting that early change dominated the likelihood of improving at the 3-month follow up. Interestingly the only SBT contribution to predicting improvement at 14 days follows up was the post initial visit categorisation, again perhaps indicating early change as being better indicators of a favourable outcome.

Finally an exploration was carried out to ascertain the potential lability of SBT subgroup categorisation over the time period from categorisation at the initial visit compared with those 2 days following the initial visit (Table 6).

Around the same proportion of subjects deteriorated and improved during this time period with, in total, over a third of patients changing SBT subgroups during the period between just before and 2 days post the initial treatment. This may reflect the lability of the SBT, the condition itself, some impact of the first visit or all three, although treatment effects must remain entirely speculative with this design.

However, there was no consistent difference between those that deteriorated or improved in their subsequent improvement at each of the follow up times although in general those that deteriorated one SBT category did slightly better at follow up than those that stayed the same compared to those that improved one SBT category.

During the course of the study, there was a 58% drop out of respondents at 90 days. An analysis of baseline characteristics of respondents compared to non-respondents found no significant differences at 14 and 30 days follow up. However, at 90 days some characteristics were significantly different with respondents being slightly older (46.2 versus 49.9), more likely to be a returning patient and less chronic than non respondents.

Discussion

With spiraling health costs in chronic conditions generally (The Health Foundation 2011) and little remittance in the cost of LBP specifically (Becker et al, 2010), there remains a need for guidance as to which patients might benefit from specific targeted intervention, despite general guidance concerning the range of treatments available (NICE, 2009). Given that a large number of LBP patient are routinely categorised as non specific in nature,

ascertaining cause is problematic as a guide to targeted treatment whereas broad screening using tools such as the SBT may prove more useful.

In addition, effective targeting may help to curb unnecessary and inappropriate use of high cost pathways for those that need minimal intervention, and in this respect the SBT has been shown to provide a method of guiding a potentially large group of nsLBP patients (low risk) toward low cost management. In the case of this study for example, nearly two fifths of the patients fall into the low risk category. It also potentially provides further guidance by invoking the differential of increasing psychological overlay to define high from medium risk patients.

However, the prediction of outcome in the nsLBP population under care has been disappointing when restricted to baseline characteristics with shorter duration of condition being one of the few consistent predictors of favourable outcome (Leboeuf-Yde et al, 2009). This is also apparent in this study but only for short-term prognosis. Emerging evidence however, suggests that early changes in condition specific characteristics maybe be more helpful in determining the eventual improvement or otherwise of patients attending for chiropractic treatment (Axen et al, 2005; Bolton and Hurst, 2011). Given the absence of any strong association between SBT stratification categories and follow up patient status when categorised at baseline (Field and Newell, 2012), this study explored the possibility that stratifying patients early after treatment had started might prove to be more useful in predicting outcome.

The results suggest that although the majority of patients did well irrespective of the subgroup they were placed in by the SBT, univariate analysis indicated that medium risk groups as categorised at baseline and post initial visit do better at short to medium follow up than the other risk groups. However, after adjusting for other baseline and follow up variables, only the post visit SBT categorisations display significant association with differential outcomes, with again the medium risk group improving more than the other groups at 30 days follow up.

Bolstering results reported previously (Axen et al, 2011), both 30 and 90 days improvement was dominated by favourable change in the previous follow up points, although at 30 days follow up females tended to do better than males.

The change in SBT categorisation of patients over a short time period is the first to our knowledge to be reported. Surprisingly, in the 2 days between initial and post visit categorisation over 1/3 of patients swapped risk groups with around equal numbers improving or deteriorating. When these groups were followed up there was little difference in improvement status compared to those individuals that maintained their SBT risk categorisation. A further unexpected results was that those who had deteriorated were more likely to improve at 30 days, albeit only small numbers of patients. This raises questions as to when the SBT is best administered if it is to be used as a clinical decision making aid.

In this population, SBT categorisation was associated with the number of treatment sessions patients were likely to receive during the first month of care but not at 90 days.

Given the theory underlying the SBT subgrouping that those in the medium risk group have largely physical barriers to recovery whilst those in the high risk group have more complex barriers including psychological factors, this study's findings that those in the medium risk group did better than the high risk group when presenting to manual therapists is unsurprising. That the medium risk group did better than those categorised as having few barriers to recovery (low risk) is unexpected. Reasons for this can only remain conjecture, however, althought the lower number of treatments provide to them may be significant.

As with the previous paper by these authors this study found no difference between individuals categorised as high and low risk. One possible reason suggested is that the psychological risk factors contributing to the high-risk score are being effectively targeted by the chiropractors (Foster et al, 2013)

Study Limitations

This population of patients were those self referring for chiropractic care and may not be enitriely reperesentative of the wider nsLBP population. In addition, despite large numbers this was a geographically focused set of clinics in the south of the UK, again limiting the generalizability of these results.

In addition, drop out rates over the course of the study, particularly at 90 days follow up may bias the outcomes, limiting the interpretation of this time point.

Conclusion

During chiropractic treatment for nsLBP patients the medium risk category patients were more likely to improve in the short to medium term compared to the other risk groups regardless of the timing of the SBT stratification. However, following adjustment with other baseline variables only the post visit SBT categorisation remained as a predictor of outcome, albeit only at 30 days follow up.

Multivariate models were dominated by condition status at previous time points indicating that early change in symptomatology has a far greater influence on future prognosis than status at baseline.

Stratification using the SBT is somewhat unstable over the very short term with over one third of patients changing SBT status in a short time window.

Further work is indicated to increase understanding of the impact of timing of SBT categorisation on its usefulness in stratifying patient's to differing care pathways.

295

296 References

297 Axen, I. et al., 2005. The Nordic Back Pain Subpopulation Program: Can Patient Reactions
298 to the First Chiropractic Treatment Predict Early Favorable Treatment Outcome in Non-
299 persistent Low Back Pain? Journal of Manipulative and Physiological Therapeutics, 28(3),
300 pp.153-158.

301 Axen, I., Bodin, L. & al, E., 2011. Clustering patients on the basis of their individual course
302 of low back pain over a six month period. BMC ….

303 Becker, A. et al., 2010. Low back pain in primary care: costs of care and prediction of future
304 health care utilization. Spine, 35(18), pp.1714-1720.

305 Bolton J E and Breen AC., 1999. The Bournemouth Questionnaire: a short-form
306 comprehensive outcome measure. I. Psychometric properties in back pain patients. Journal
307 of Manipulative and Physiological Therapeutics, 22(8), pp.503-510.

308 Bolton JE and Hurst HC., 2011. Prognostic factors for short-term improvement in acute and
309 persistent musculoskeletal pain consulters in primary care. Chiropractic & Manual
310 Therapies, 19(1), p.27.

311 British Pain Society, 2012. Low back and radicular pain pathway.
312 http://bps.mapofmedicine.com/evidence/bps/low_back_and_radicular_pain1.html, Accessed
313 July 2014

314 Childs, J.D. et al., 2004. A clinical prediction rule to identify patients with low back pain most
315 likely to benefit from spinal manipulation: a validation study. Annals of Internal Medicine,
316 141(12), pp.920-928.

317 Field JR and Newell D., 2012. Relationship between STart Back Screening Tool and
318 prognosis for low back pain patients receiving spinal manipulative therapy. Chiropractic &
319 Manual Therapies, 20(1), pp.1-16.

320 Flynn, T. et al., 2002. A clinical prediction rule for classifying patients with low back pain
321 who demonstrate short-term improvement with spinal manipulation. Spine, 27(24), p.2835.

322 Foster N, Hill J, O'Sullivan P & Hancock M. 2013. Stratified models of care. Best Practice &
323 Research Clinical Rheumatology, 27(5), pp.649–661.

324 Fritz JM, Beneciuk JM and George SZ., 2011. Relationship between categorization with the
325 STarT Back Screening Tool and prognosis for people receiving physical therapy for low
326 back pain. Physical therapy, 91(5), pp.722-732.

327 Hancock, M.J. et al., 2008. Independent evaluation of a clinical prediction rule for spinal
328 manipulative therapy: a randomised controlled trial. European Spine Journal, 17(7), pp.936-
329 943.

Hill, J.C. et al., 2008. A primary care back pain screening tool: identifying patient subgroups for initial treatment. Arthritis Care & Research, 59(5), pp.632-641.

Hill, J.C. et al., 2011. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. The Lancet.

Larsen, K. & Leboeuf-Yde, C., 2005. The Bournemouth Questionnaire: Can it Be Used to Monitor and Predict Treatment Outcome in Chiropractic Patients With Persistent Low Back Pain? Journal of Manipulative and Physiological Therapeutics, 28(4), pp.219-227.

Leboeuf-Yde, C. et al., 2009. The Nordic Subpopulation Research Programme: prediction of treatment outcome in patients with low back pain treated by chiropractors- does the psychological profile matter? Chiropractic & Osteopathy, 17(1), p.14.

Lemeunier, N. et al., 2012. The natural course of low back pain: a systematic critical literature review. Chiropractic & Manual Therapies, 20(1), p.33.

National Institute for Health and Care Excellence, 2009. Low back pain: Early management of persistent non-specific low back pain. CG88. London: National Institute for Health and Care Excellence.

Newell, D & Field, J., 2007. Who will get better? Predicting clinical outcomes in a chiropractic practice. Clinical Chiropractic, 10(4), pp.179-186.

Newell, D and Bolton, JE, 2010. Responsiveness of the Bournemouth questionnaire in determining minimal clinically important change in subgroups of low back pain patients. Spine, 35(19), pp.1801-1806.

Rubinstein, S.M. et al., 2008. Predictors of a favorable outcome in patients treated by chiropractors for neck pain. Spine, 33(13), pp.1451-1458.

The Health Foundation, 2011. Snapshot. Co-creating Health., pp.1-16.

van Middelkoop, M. et al., 2011. A systematic review on the effectiveness of physical and rehabilitation interventions for chronic non-specific low back pain. European Spine Journal, 20(1), pp.19-39.

UK DoH. Extension of Patient choice of Any Qualified Provider in MSK Services for Back and Neck Pain. https://www.supply2health.nhs.uk/AQPResourceCentre/AQPServices/PTP/Pages/BackNekPain.aspx Accessed June 2014

Table 1: Descriptive analysis of initial variables for whole cohort

Continuous Variables	Mean (SD)
Mean Age (SD)	47.8 (13.9)
Mean Pain (SD)	6.4 (2.0)
Mean BQ Total (SD)	34.3 (16.4)

Categorical Variables	Proportion (%)
Female	56.5%
Seen Practitioner before	24.5%
New patient	69.1%
Leg Pain	
Above the knee	33.0%
Below the knee	12.4%
>30 days pain in year	55.2%
Recurring	66.5%
Duration	
< 1 month	43.2%
1-3 months	10.0%
>3 months	46.6%
SBT Baseline	
Low	39.1%
Medium	36.8%
High	23.7%
SBT 2 Days Post Initial Treatment	
Low	39.0%
Medium	36.8%
High	24.2%

Table 2: Numbers of visits at outcome points across SBT groups as defined post visit

	14 days*	30 days*	90 days
	Mean (SD)	**Mean (SD)**	**Mean (SD)**
SBT Group			
LOW	2.6 (1.2)	3.7 (1.9)	4.2 (2.4)
MEDIUM	3.0 (1.3)	4.0 (1.9)	4.6 (3.0)
HIGH	3.1 (1.3)	4.4 (2.1)	5.1 (2.7)

*= $p<0.001$ for Kruskal Wallace between SBT categories at each follow up point

Table 3: Baseline variables across SBT categories measured at initial (IT) and 2 days post-initial treatment (PT)

Variables	**SBT Category**

	Low		Medium		High
			p value		
	IT (n=282)	PT (n=285)	IT (n=268)	PT (n=264)	IT
(n=166)	**PT (n=167)**		**IT** **PT**		
Mean (SD)					
Age	47.1(15.1)	46.2 (15.0)	46.8 (12.5)	47.6 (12.7)	50.5
(13.8)	50.4 (15.5)		** **		
Pain	5.3 (2.0)	5.4 (2.1)	6.8 (1.8)	6.8 (1.7)	7.5 (1.5)
	7.4 (1.5)		** **		
BQ Total	25.1(13.4)	26.0 (13.5)	37.7(13.7)	37.5 (14.9)	44.4
(16.6)	43.0 (16.3)		** **		
Proportion (%)					
Female	56.8	54.9	57.4	59.6	54.5
	54.4				
Seen Practitioner before	28.9	27.6	24.2	21.7	17.4
	23.6		*		
Is new patient	66.0	65.2	69.0	73.3	74.7
	69.2				
Leg Pain					
Above the knee	23.5	23.9	36.8	40.1	43.3
	36.8		*	*	
Below the knee	8.8	9.6	10.8	14.1	20.8
	14.3		*		
>30 days pain in year	58.8	58.0	52.0	55.2	55.1
	50.5				
Recurring	69.4	69.6	67.5	66.8	61.2
	61.0				
Duration					
< 1 month	36.7	36.5	50.2	45.8	43.8
	50.0		*	*	
1-3 months	13.6	12.3	5.8	9.4	10.7
	7.1				
>3 months	49.7	51.2	44.0	44.4	45.5
	41.8				

*p<0.05 (Chi² test for trend), **p<0.01 (Kruskal-Wallis), SD = Standard Deviation, BQ = Bournemouth Questionnaire

Table 4: Predicting improvement using SBT at initial (IT) and 2 days post initial treatment (PT)

	14 days (n= 542) OR (95% CI)	**30 days (n=416) OR (95% CI)**	**90 days (n=318) OR (95% CI)**
SBT BASELINE			
Low	1.0	1.0	1.0
Medium	**1.6 (1.1 to 2.5)**	**1.6 (1.0 to 2.6)**	1.1 (0.6 to 2.1)
High	1.1 (0.7 to 1.8)	0.8 (0.5 to 1.3)	0.8 (0.4 to 1.6)
Nagelkerke	0.015	0.022	0.004
AUC	0.56 (0.51 to 0.61)	0.57 (0.51 to 0.63)	0.53 (0.45 to 0.61)
	14 days (n=545) OR (95% CI)	30 days (n=418) OR (95% CI)	90 days (n=318) OR (95% CI)
SBT POST VISIT			
Low	1.0	1.0	1.0
Medium	1.2 (0.8 to 1.8)	**1.8 (1.1 to 3.0)**	1.4 (0.7 to 2.7)
High	1.3 (0.8 to 2.0)	1.0 (0.6 to 1.6)	1.6 (0.7 to 3.4)
Nagelkerke	0.004	0.024	0.01
AUC	0.53 (0.48 to 0.58)	0.57 (0.51 to 0.63)	0.55 (0.47 to 0.63)

PGIC = Patient Global Impression of Change; SBT = STarT Back Tool; OR= Odds Ratio; AUC= Area under the curve, Bold= significant at $p<0.05$

Table 5: Results of multivariate* analysis for predicting improvement at 14, 30 and 90 days following initial visit

Follow up point	Variables in the equation	OR (95% CI)	ROC AUC (95% CI)
	Nagelkerke R^2		
14 days (n=545)			**0.70 (0.66 to 0.75)**
	0.16		
	Pain above the knee (NO)	1.7 (1.2 to 2.5)	
	>30 days pain in year (NO)	2.3 (1.4 to 3.6)	
	Duration		
	> 3 months	1.0	
	1-3 months	2.5 (1.3 to 4.9)	
	< 1 month	2.2 (1.4 to 3.6)	
30 days (n=367)			**0.82 (0.77 to 0.87)**
	0.37		
	Improved at 14 days (PGIC)	12.3 (7.3 to 21.0)	
	Gender (Female)	1.7 (1.0 to 3.0)	
	SBT Ranking Post Treatment		
	Low	1.0	
	Medium	2.0 (1.1 to 3.8)	
	High	0.9 (0.4 to 1.7)	
90 days (n=241)			**0.84 (0.78 to 0.91)**
	0.41		
	Improved at 14 days (PGIC)	4.4 (1.9 to 10.2)	
	Improved at 30 days (PGIC)	8.7 (3.7 to 20.4)	

* All baseline variables and SBT ranking categories at initial and post initial visit were included

Table 6: Proportion of patients changing SBT risk groups between initial and two days post initial visit and odds of subsequent improvement at follow up

	Proportion* (%) in each change group	**Proportion (%) improved at follow up**	
	Post-initial visit (N)	14 days (Total N)	30 days (Total N)
	90 days (Total N)		
Deteriorated			
Low-Medium	7.3 (55)	57 (37)	80 (20)
	94 (16)		
Low-High	3.2 (24)	71 (21)	87 (15)
	92 (13)		
Medium-High	6.2 (47)	76 (34)	77 (26)
	89 (18)		
Total deteriorated	16.7 (126)		
		Odds [OR (95% CI)] of improved (PGIC) if increased SBT risk group	
		1.2 (0.7 to 2.0)	**2.0 (1.1 to 4.0)§**

	Proportion* (%) in each change group	**Proportion (%) improved at follow up**	
	Post-initial visit (N)	14 days (Total N)	30 days (Total N)
	90 days (Total N)		
Improved			
Medium-Low	8.8 (66)	67 (43)	70 (40)
	75 (28)		
High-Low	1.6(12)	78 (9)	80 (5)
	80 (5)		
High-Medium	7.6 (57)	59 (44)	71 (35)
	74 (31)		
Total improved	18.0 (135)		
		Odds [OR (95% CI)] of improved (PGIC) if reduced SBT risk group	
		1.0 (0.7 to 1.7)	1.2 (0.7 to 2.0)
	0.6 (0.3 to 1.2)		
Total any change	34.7 (261)		
		Odds [OR (95% CI)] of improved (PGIC) if any change in SBT risk group	
		1.1 (0.8 to 1.7)	1.6 (1.0 to 2.6)
	1.1 (0.6 to 2.0)		

* Proportion of the patients that had changed SBT categories at post initial visit SBT categorisation: §= < 5 in one cell