Mar 10, 2022

Statistical Analysis Plan for Validation of the STarTBack Tool for Management of Low Back Pain in the Military Health System

Daniel I Rhon[1], Steve George[2], Emily Poehlein[3], Cindy Green[3]

[1]Department of Rehabilitation Medicine, School of Medicine, Uniformed Services University of the Health Sciences;

[2]Vice Chair of Research, Department of Orthopaedic Surgery and Therapeutic Area Lead, Musculoskeletal and Surgical Sciences, Duke Clinical Research Institute;

[3]Department of Biostatistics and Bioinformatics, Duke University School of Medicine

dx.doi.org/10.17504/protocols.io.b55pq85n

Daniel Rhon

DISCLAIMER: The view(s) expressed herein are those of the author(s) and do not reflect the official policy or position of Brooke Army Medical Center, the U.S. Army Medical Department, the U.S. Army Office of the Surgeon General, the Department of the Army, the Department of Defense, the Uniformed Services University of the Health Sciences, the Defense Health Agency or the U.S. Government.

Citation: Daniel I Rhon, Steve George, Emily Poehlein, Cindy Green Statistical Analysis Plan for Validation of the STarTBack Tool for Management of Low Back Pain in the Military Health System https://dx.doi.org/10.17504/protocols.io.b55pq85n

DOI

dx.doi.org/10.17504/protocols.io.b55pq85n

Daniel I Rhon, Steve George, Emily Poehlein, Cindy Green 2022. Statistical Analysis Plan for Validation of the STarTBack Tool for Management of Low Back Pain in the Military Health System. **protocols.io** https://dx.doi.org/10.17504/protocols.io.b55pq85n

Clinical Rehabilitation Medicine Research Program (CRMRP) - Formerly JPC-8
Grant ID: W81XWH-18-1-0475

document ,

Mar 09, 2022

Mar 10, 2022

59279

DISCLAIMER:

DISCLAIMER: The view(s) expressed herein are those of the author(s) and do not reflect the official policy

or position of Brooke Army Medical Center, the U.S. Army Medical Department, the U.S. Army Office of the Surgeon General, the Department of the Army, the Department of Defense, the Uniformed Services University of the Health Sciences, the Defense Health Agency or the U.S. Government.

Administrative Information

The SAP below is written following the Guidelines for Content of Statistical Analysis Plans in Clinical Trials (adapted for the purposes of this cohort study) JAMA Reference: http://dx.doi.org/10.1001/jama.2017.18556

A	B
TRIAL FULL TITLE	Validation of the STarT Back Screening Tool in the Primary Care Management of Low Back Pain in the Military Health System: A Randomized Trial of Risk-stratified vs. Usual Care
Trial Registration	ClinicalTrials.gov NCT03127826
SAP VERSION	Version 1
SAP VERSION DATE	1/7/2021
TRIAL STATISTICIAN	Emily Poehlein, MB; Cindy Green, PhD
TRIAL PRINCIPAL INVESTIGATOR	Dan Rhon, DPT, DSc, PhD
SAP AUTHOR(s)	Emily Poehlein, MB; Cindy Green, PhD
ROLES AND RESPONSIBILITIES	Dan Rhon and Steven George: Investigators Emily Poehlein and Cindy Green: Biostatisticians – responsible for finalizing SAP, agreement on best practices and statistical methods for analysis

1Abbreviations and Definitions

Citation: Daniel I Rhon, Steve George, Emily Poehlein, Cindy Green Statistical Analysis Plan for Validation of the STarTBack Tool for Management of Low Back Pain in the Military Health System https://dx.doi.org/10.17504/protocols.io.b55pq85n

A	B
CEQ	Credibility Expectancy Questionnaire
COVID-19	Coronavirus Pandemic
EQ-5D	EuroQoL
LBP	Low Back Pain
MHS	Military Health System
MTF	Military Teatment Facility
NPQ	Revised Neurophysiology of Pain Questionnaire
OSPRO-YF	Optimal Screening for Prediction of Referral and Outcome Yellow Flags
PASS	Patient Acceptable Symptom Scale
PROMIS	Patient Reported Outcomes Measurement Information Systems
SBST	STarT Back Screening Tool
SOPA-35	Survey of Pain Attitudes-35
QoL	Quality of Life

Abbreviations and Definitions

1. INTRODUCTION

1.1Background/Rationale

Low back pain (LBP) is among the most frequent causes of medical visits and lost-duty time in the Military Health System (MHS). Most clinical practice guidelines recommend only advice and education for all patients with non-specific LBP during the initial weeks of management, with consideration of psychosocial factors and referral to physical therapy recommended only when recovery is delayed. Recent research has demonstrated that the STarT Back Screening Tool (SBST) is useful for classifying patients as being at low, medium, or high risk for experiencing chronic, disabling LBP. Targeted interventions for patients in each risk subgroup have also been developed to address the specific modifiable prognostic indicators identified by the tool. Patients who received stratified care utilizing the SBST demonstrated greater changes in disability, increased quality of life, and lower healthcare costs compared to patients in the control group at 12 months.

One of the limitations of the STarT Back trial's study design was that there was no standardization of the physical therapy interventions that were delivered. Therefore, it's difficult to ascertain whether the favorable outcomes in the risk stratified group are attributable to superior physical therapy intervention or the overall effectiveness of the stratification process in directing the right patients to physical therapy with or without psychological augmentation. It is also unknown whether a similar stratified care approach will achieve similar results in the primary care management of patients with LBP in the MHS. If these results could be validated, this simple-to-use screening strategy could be implemented practically and efficiently across the MHS, with the expectation that the MHS would realize substantial cost savings and lower disability among MHS beneficiaries with LBP. Therefore, the purpose of this study is to validate the clinical and cost effectiveness of the SBST in the primary care management of patients with LBP in the MHS.

1.2 Scope of the analyses

Citation: Daniel I Rhon, Steve George, Emily Poehlein, Cindy Green Statistical Analysis Plan for Validation of the STarTBack Tool for Management of Low Back Pain in the Military Health System https://dx.doi.org/10.17504/protocols.io.b55pq85n

These analyses will compare efficacy and costs resulting from a risk-stratified approach to a usual care approach in the management of LBP in primary care using the recently developed SBST.

2. STUDY OBJECTIVES AND ENDPOINTS

2.1 Study Objectives

PRIMARY AIM:

Specific Aim 1: Compare clinical outcomes of care between risk-stratified care according to the STarT Back Screening Tool and usual care approach in the management of patients with LBP in the primary care setting.

Hypothesis Specific Aim 1: Greater improvements in both short- and long-term clinical outcomes will be observed among patients treated with a risk-stratified versus a usual care approach.

SECONDARY AIMS:

Specific Aim 2: Compare direct costs associated with risk-stratified versus usual care in the management of patients with LBP in the primary care setting.

Hypothesis Aim 2: Direct costs will be lower at 12 months among patients treated with risk-stratified care versus usual care.

.

2.2 Endpoints

PRIMARY ENDPOINT:

The primary endpoint is self-reported disability (Roland Morris Disability Index) at 12 months.

SECONDARY ENDPOINTS:

- Pain and Quality of Life (QoL)
- Patient Reported Outcomes Measurement Information Systems (PROMIS)-57 pain interference subdomain T Score at 12 months
- PROMIS-57 physical function subdomain T Score at 12 months Healthcare Utilization
- Total number of LBP related visits over 12 months of follow-up
- Requiring surgery over 12 months of follow-up
- Requiring another invasive procedure over 12 months of follow-up
- Requiring opioid use over 12 months of follow-up
- Referral to pain clinic over 12 months of follow-up
- Change in SOPA from baseline to 6 weeks
- Change in NPQ from baseline to 6 weeks

3. STUDY METHODS

3.1 General Study Design and Plan

This study is a parallel, two-arm randomized clinical trial with a 1-year follow-up period. The trial will assess superiority of risk-stratified care compared to usual care. The study coordinator will remain blind to the patient's treatment group assignment throughout the study. Patients will be randomized to either

Citation: Daniel I Rhon, Steve George, Emily Poehlein, Cindy Green Statistical Analysis Plan for Validation of the STarTBack Tool for Management of Low Back Pain in the Military Health System https://dx.doi.org/10.17504/protocols.io.b55pq85n

risk-stratified or usual care, where patients will receive various levels of physical therapy and other management for LBP according to the following table.

3.2 Randomization and Blinding
After completion of baseline assessments, subjects will be randomized into one of two arms (Group I = Usual Care, Group II= Risk Stratified Care based on the SBST). The method of group assignment will be sequentially numbered opaque sealed envelopes (SNOSE). To minimize the risk of predicting the treatment assignment of the next eligible patient, randomization will be performed in permuted blocks of two or four with random variation of the blocking number. Of the 290 subjects expected for enrollment at SAMMC, ~240 (83%) are expected to be enrolled at BAMC, and ~50 (17%) to be enrolled at WHASC. Randomization envelopes will be managed at the BAMC location, and due to its proximity to WHASC, the research coordinator will be able to provide the randomization number to the research team at WHASC as necessary. Treatment allocation will be tracked and coded locally at each site. Because we are using permuted block randomization by site, there will be a balanced number of patients in each test group for each site. Statisticians responsible for the primary and secondary analyses will be blinded to the treatment assignment until the SAP is signed and models/tables are programmed.

Participants are briefed on the purpose of the study prior to consenting; however, they were not made aware of which group or risk stratification category they may have been assigned to. Physical therapists trained in the risk stratification approach were referred the patients in the risk stratification arm and untrained physical therapists received the usual care participants. The investigators were blinded to group allocation throughout the entire period of the study. Folllow-up data collection is automated through data collection portals, with a reminder and link sent to the participant at a pre-determined time point. Those who do not respond are contacted by study coordinators blinded by treatment allocation. Data provided to the statisticians for analysis was stripped of all identifiers. Treatment allocation was coded so that the statisticians would be unaware of treatment allocation.

3.3 Endpoint definitions and variable formats
Primary endpoint: Roland Morris Disability Index
The Roland Morris Disability Index is calculated from the 24 items from the assessment as the number of times the patient has ticked “today” with greater scores indicating a greater level of disability. It will be analyzed as a count variable.

Pain and Quality of life
PROMIS-57: The PROMIS-57 pain interference and physical function subdomain scores are continuous measures of quality of life with greater scores indicating a lower quality of life. They will be analyzed using sub-component T scores as continuous variables

Health care utilization
LBP-related visits: The number of LBP-related visits including physical therapy, specialist care, etc. will be counted. It will be analyzed as a count variable.

Surgery: The need for LBP-related surgery over 12 months of follow-up will be recorded for each patient. It will be analyzed as a binary variable.

Invasive procedures: the need for an invasive procedure other than surgery such as epidural steroid

Citation: Daniel I Rhon, Steve George, Emily Poehlein, Cindy Green Statistical Analysis Plan for Validation of the STarTBack Tool for Management of Low Back Pain in the Military Health System https://dx.doi.org/10.17504/protocols.io.b55pq85n

injection, facet joint injection, trigger point injection, or sacroiliac injection over 12 months of follow-up will be recorded for each patient. It will be analyzed as a binary variable.

Opioid use: the use of opioids at any point over 12 months of follow-up will be recorded for each patient. It will be analyzed as a binary variable. Additional opioid use endpoints include count of unique fills, total days' supply, morphine milligram equivalents, and chronic opioid use.

Referral to pain clinic: Referral to a pain clinic over the 12 month follow-up period will be recorded for each patient. It will be analyzed as a binary variable.

4. SAMPLE SIZE

We will enroll 290 consecutive patients seen in primary care with a primary complaint of LBP who consent to participation. Patients will be enrolled in the primary care clinics at four military treatment facilities (MTFs).

Sample Size Estimation

The specific aims will be tested by use of pretreatment randomization allocation to low-risk, medium-risk, and high-risk groups in the intervention and control groups. The sample size calculation is based on the ability to detect a between-group effect size of 0.3 at the 12-month primary endpoint (Roland-Morris Disability Index). Based on a two-tailed significance level at an alpha level equal to 0.05, 80% power, and allowing for a 25% loss to follow-up, we will aim to recruit 290 participants, 145 in each group overall (Risk Stratified and Usual Care).

5. GENERAL ANALYSIS CONSIDERATIONS

5.1Timing of Analyses

The final analysis will be performed when the last recruited patient has 12 months of follow-up and after the finalization and approval of this SAP document. The final analysis will be performed on blinded data transferred to statisticians following data cleaning conducted by a study research coordinator. Data cleaning will include scoring outcomes and checking for missing data/extreme values and outliers.

5.2 Analysis Populations

5.2.1 *Safety Populations:* All subjects who were randomized and participated in at least one post-baseline treatment session will be analyzed in the secondary analysis. 7

5.2.2 Modified intention-to-treat (MITT) popoulation

All subjects who were randomized participated in at least one post-baseline treatment session and had baseline and one other Roland-Morris Disability Index measures taken and recorded will be analyzed for the primary aim.

5.3 Covariates and subgroups

Models for the primary efficacy analysis will include baseline score as a covariate. Secondary models will include active duty status, age, sex, and back pain duration. The CEQ, a 6-item self-report evaluating treatment credibility and expectations for improvement, will be assessed at baseline after treatment

Citation: Daniel I Rhon, Steve George, Emily Poehlein, Cindy Green Statistical Analysis Plan for Validation of the STarTBack Tool for Management of Low Back Pain in the Military Health System https://dx.doi.org/10.17504/protocols.io.b55pq85n

group assignment is revealed to provide descriptive information about participants' perceptions of their treatment assignment and optimism for improvement. The Credibility Expectancy Questionnaire (CEQ) may be used as a covariate if perceptions differ meaningfully between groups as participants' initial perceptions of treatment credibility can impact outcomes.

Additionally, for a secondary efficacy analysis, subgroup analyses for, PASS, back pain duration, and active duty status will be conducted using interaction terms.

5.3.1 Multicenter Studies
We will enroll 290 consecutive patients seen in primary care with a primary complaint of LBP who consent to participation. Patients will be enrolled in primary care clinics at Brooke Army Medical Center (BAMC) and Wilford Hall Ambulatory Surgical Center (WHSAC).We will plan to enroll approximately 240 subjects at the BAMC clinic, as this will be the primary site, and plan for approximately 50 subjects at the WHASC clinic.If either of the other sites under-enrolls, then the other site can enroll a higher number, up to the total 290 subjects.We will ensure that total enrollment does not go over 290 between both sites. Since sites are in the same geographical area with similar procedures and patient population, site will not be considered as a covariate in the primary efficacy model but will be evaluated in secondary analyses.

5.4 Missing Data
Missing at random will be judged based on tabulating missing 12 month Roland Morris values by treatment group and demographic characteristics. Mixed effects modelling will be used to allow for missing data. If there is a substantial amount of missing data and missingness is considered to be not random, sensitivity analyses utilizing worst-case imputation will be used for descriptive and inferential assessments to address attrition bias.

5.5 Multiple Testing
Adjustment for multiple testing will be utilized for healthcare utilization endpoints using the Benjamini Hochberg procedure for controlling False Discovery Rate (FDR). The FDR will be controlled at

6. SUMMARY OF STUDY DATA

All continuous variables will be summarized using the following descriptive statistics: n (non-missing sample size), mean, standard deviation (SD), median, 25^{th} and 75^{th}percentiles (Q1, Q3), maximum and minimum. The frequency and percentages of observed levels will be reported for all categorical measures using non-missing data. Categorical score data may also be presented as mean (SD). All summary tables will be structured with a column for each treatment in the order (Risk-stratified, Usual) and will be annotated with the total population size relevant to that table/treatment, including any missing observations. SAS version 9.4 will be used for all analyses.

.
6.1 Subject Disposition
Some withdrawal from the study is expected for various reasons. A CONSORT diagram will be used to

Citation: Daniel I Rhon, Steve George, Emily Poehlein, Cindy Green Statistical Analysis Plan for Validation of the STarTBack Tool for Management of Low Back Pain in the Military Health System https://dx.doi.org/10.17504/protocols.io.b55pq85n

describe the number of patients that reached the various stages of the trial and the number of patients that dropped out and for what reasons.

6.2 Protocol Deviations
Protocol deviations are expected to be minimal and will be summarized in tables by study population and treatment.

6.3 Demographic and Baseline Variables
Baseline characteristics recorded at randomization or first treatment administration include demographic characteristics, medical history, current medications, and characterizations regarding LBP. The summary statistics will be produced in accordance with section 9.

6.4 Treatment Compliance
Treatment non-compliance is expected, but difficult to measure. The number of visits will be enumerated for each patient and summarized by treatment risk to determine whether non-compliance was extensive.

7 EFFICACY ANALYSIS

7.1Primary Efficacy Analysis
The primary analysis will compare the Roland Morris Disability Index at 12 months between the two treatment groups. First, Roland Morris Disability Index at each time point will be compared between the two treatment groups in a summary table in the Safety population. Then, using the Modified Intention-to-Treat (MITT) population, treatment effect estimates will be obtained using a repeated measures model including treatment, time (categorical), and an interaction between treatment and time with adjustment for baseline score using an unstructured covariance matrix and degrees of freedom by Kenward-Rodger. Potential distributions for count data include normal, lognormal, negative binomial, and Poisson and will be determined by assessing residuals and other graphical methods. Least-squares means (LS-mean) for each treatment group and the difference between the two will be reported with 95% confidence interval (CI) and two-sided p-value.

7.2 Secondary Efficacy Analyses
7.2.1 Secondary Analyses of Primary Efficacy Endpoint:
The analysis of the primary efficacy aim will be repeated with additional adjustment for age, sex, back pain duration, active duty status, and other covariates such as CEQ or PASS that are found to be unevenly distributed between the two treatment groups at baseline.

Additionally, if there is a substantial amount of missing data and missingness is considered to be not random, an additional analysis utilizing worst-case imputation will be implemented.

Lastly, an additional secondary analysis of the primary efficacy endpoint will include further adjustment for therapists' effects using a random effect for primary therapist in the model specified in section 10.1.

7.2.2 Exploratory Analyses for Primary Efficacy Endpoint
For exploratory analyses of the primary efficacy endpoint, subgroup analyses for, PASS, back pain duration, and active duty status will be conducted using interaction terms in the model specified in

Citation: Daniel I Rhon, Steve George, Emily Poehlein, Cindy Green Statistical Analysis Plan for Validation of the STarTBack Tool for Management of Low Back Pain in the Military Health System https://dx.doi.org/10.17504/protocols.io.b55pq85n

section 10.1.

7.2.3 Analyses of Secondary Endpoints
PROMIS-57 pain interference and physical function T-scores will be analyzed using the same methods as specified in section 10.1 using a normal likelihood function.

Health care utilization variables will be summarized using tables and compared using non-parametric methods for the number of visits and risk ratios for the other binary outcomes with accompanying p-values.

10 SAFETY ANALYSES

Adverse events (non-serious and serious) and deaths are expected to be minimal and will be summarized using listings. If there are a large number of adverse events, they will be tabulated by treatment arm.

11 REPORTING CONVENTIONS

P-values ≥0.001 will be reported to 3 decimal places; p-values less than 0.001 will be reported as "<0.001". The mean, standard deviation, and any other statistics other than quantiles, will be reported to two decimal places.

12 SUMMARY OF CHANGES TO THE PROTOCOL AND/OR SAP

Rationale for Adjustments of Statistical Analysis Plan from Protocol (Version 8, June 30, 2020)

1. Analysis of primary aim

PROTOCOL:"For the primary analyses, imputed datasets will be used for all descriptive and inferential assessments to address attrition bias, generated through multiple imputation (pooled estimates of five imputed datasets) by use of simulation based on a multivariate normal model (numerical variables) and a logistic regression model (categorical outcomes)"

SAP: For the primary analysis, we will plan to use a repeated measures model instead of imputation to address missing data since we do not know the extent of missingness in the dataset and imputation with repeated measures is complicated. In the case that missingness appears to not be random, worst-case imputation will be considered as a secondary sensitivity analysis. Repeated measures models allow all available data to be used and provide less bias than other missing data approaches such as complete case analysis or last observation forward.

Citation: Daniel I Rhon, Steve George, Emily Poehlein, Cindy Green Statistical Analysis Plan for Validation of the STarTBack Tool for Management of Low Back Pain in the Military Health System https://dx.doi.org/10.17504/protocols.io.b55pq85n