UNC | GILLINGS SCHOOL OF GLOBAL PUBLIC HEALTH

PHAIR

A SMART Clinical Trial for Chronic Lower Back Pain

John Sperger and Michael R. Kosorok

2024-04-08

Outline

Biomarkers for Evaluating Spine Treatments (BEST) Trial

Goal: inform a precision medicine to treating chronic low-back pain (cLBP).

- Chronic low-back pain (cLBP) lasting $3+$ months with pain occurring on most days affects 10-20% of adults in the US.
- cLBP treatment is challenging due to the diverse etiology of back pain and the varied phenotypes of back pain patients.

Study Objectives

Primary Objective: Estimate an algorithm for assigning sequences of two CLBP treatments based on phenotypic markers and patient response, optimizing effectiveness.

Secondary Objectives:

1. Estimate DTRs balancing multiple outcomes considering participant preferences.
2. Estimate DTRs incorporating additional "deep" phenotypic markers collected on a subset of participants.
3. Assess whether treatment effectiveness is sustained 24 weeks post-second treatment (36-week endpoint).

Study Interventions

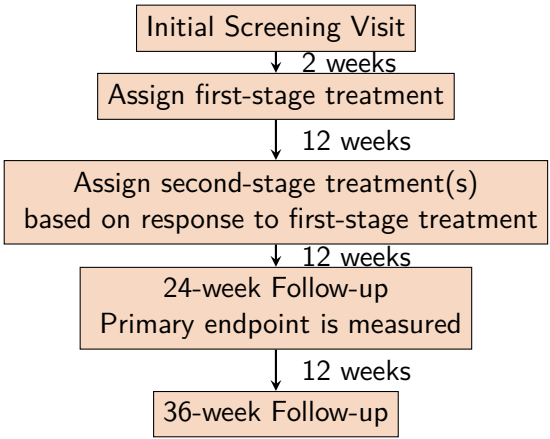
BEST is a two-stage SMART to evaluate four evidence-based interventions for cLBP that span a range of treatment modalities:

- Enhanced Self-Care (ESC)
- Acceptance & Commitment Therapy (ACT)
- Evidence-Based Exercise and Manual Therapy (EBEM)
- Duloxetine

These interventions are all well-established but only have moderate effect sizes.

Trying to find the intervention with the highest average treatment effect (ATE) is not a useful goal—differences are likely small, wouldn't help most patients.

Design Overview



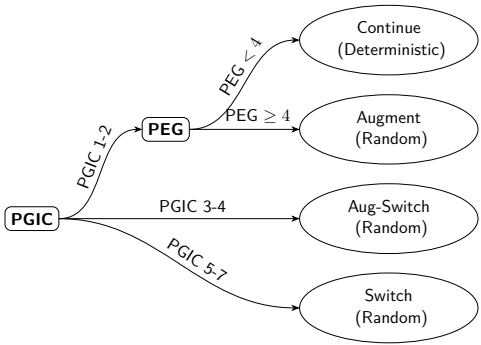


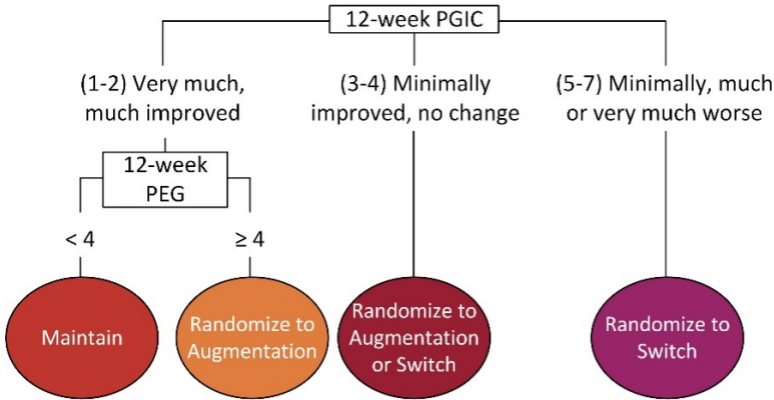
Table of Contents

Key Design Features

- **Two-stage SMART design** to mimic clinical decision-making, with treatment tailoring based on initial treatment response.
- **Pragmatic design** elements, including inclusion of participants contraindicated for one study intervention, and use of four responder categories to align with clinical judgment.
- **Consortium-wide protocol** elements for future data integration, enabling broader research applications.

Treatment Response and Decision Making

Figure: BEST Trial Responder Statuses and Treatment Eligibility



** Patients initially randomized to ESC with 12-week PGIC scores of 3-7 will be randomized to augmentation and not switching.*

Table of Contents

Treatment Rule Estimation

- Employ Q-learning with functional estimation to estimate the optimal DTR, using 24-week Pain, Enjoyment of Life, and General activity (PEG) score as the response.
- Incorporate baseline PEG as a predictor to optimize efficacy for the primary endpoint.
- Utilize both "black-box" and interpretable methods for regression functions to balance interpretability and discovery.

Q-Learning Overview

Q-learning applies a general framework reducing the multi-stage DTR optimization to a series of regressions via backwards induction.[1]

- Q-learning is agnostic to the regression function estimation method, accommodating both "black-box" and interpretable methods.[2]
- Primary Outcome: 24-week PEG score.
- Predictors: Baseline PEG score, Patient Reported Outcomes, Biomechanical Assessment,

[1] Christopher J. C. H. Watkins and Peter Dayan. "Q-Learning". In: *Machine Learning* 8.3 (May 1, 1992), pp. 279–292; Bibhas Chakraborty and Susan A. Murphy. "Dynamic Treatment Regimes". In: *Annual Review of Statistics and Its Application* 1.1 (2014), pp. 447–464.

[2] Anastasios A. Tsiatis et al. *Dynamic Treatment Regimes: Statistical Methods for Precision Medicine*. New York: Chapman and Hall/CRC, Dec. 19, 2019. 618 pp.

Interpretable DTR Estimation

We will estimate the DTR as a decision list of "if-then" statements for clarity and ease of use in clinical settings.

- Ensures the DTR is understandable for both clinicians and patients so the DTR can be audited for scientific validity and ethical concerns.
- Performance cost of interpretability is negligible in many real-world applications.[3]
- Makes complex treatment guidelines accessible and actionable, facilitating future research directions and clinical adoption.

[3] Cynthia Rudin. "Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead". In: *Nature Machine Intelligence* 1.5 (5 May 2019), pp. 206–215.

Ensemble DTR Estimation

We will estimate a DTR using a weighted ensemble of regression methods.[4]

- Potential regression methods include elastic net, random forests, gradient boosting, kernel regression, and the interpretable modeling method.
- The ensemble can serve as a a performance benchmark by showing the trade-off, if any, between interpretability and performance.
- Variable importance metrics can be used with the ensemble to suggest biomarkers for further study beyond those included in the interpretable DTR.

[4] Mark J. Van der Laan, Eric C. Polley, and Alan E. Hubbard. "Super Learner". In: *Statistical applications in genetics and molecular biology* 6.1 (2007).

Biomarker Discovery

Goal: Guide future research towards understanding CLBP mechanisms and intervention effects.

- DTRs describe the "best treatment for each patient." We are also interested in the "best patients for each treatment."
- Purpose is to inform future research rather than clinical care (compared to DTR).
- Use variable importance measures on fitted regression models to identify key biomarkers for treatments.

Current Approaches to Powering SMARTs

Currently the three most common approaches are:[5]

1. Comparative effectiveness of the initial treatments in terms of their ATEs. This can be done using only the first stage data and outcomes, or by using the outcome after all stages of treatment and averaging over later treatments.
2. Comparing the efficacy of treatment options for nonresponders to the initial treatment.
3. Comparing two or more embedded DTRs (eDTRs) in the study. An eDTR is a DTR that involves only paths that are directly assigned in the SMART.

None of these are appropriate given the study's primary objective of DTR estimation involving covariates.

[5] Giulia Lorenzoni et al. "Use of Sequential Multiple Assignment Randomized Trials (SMARTs) in Oncology: Systematic Review of Published Studies". In: *British Journal of Cancer* 128.7 (7 Mar. 2023), pp. 1177–1188.

Challenges Powering for DTR Estimation

- **Covariates:** DTR estimation involves covariates which are not often incorporated in standard sample size calculation methods.
  - eDTRs may include covariates, but to do so they must be incorporated in the study design. This requires that they be specified ahead of time and not data-derived.
- **Post-selection Inference:** DTR estimation often involves variable selection, i.e. inference after model selection, which require inference methods that account for the additional uncertainty from model selection.
- **Published estimates** of effect sizes for treatment interactions and sequence effects rarely exist.

Sample Size for DTR Estimation

1. Evaluate performance in terms of the value of the estimated policy $\hat{\pi}_n$ and the value of the optimal policy π^*

2. Define a power analogue: the probability that the estimated DTR is within a set tolerance of the optimal DTR

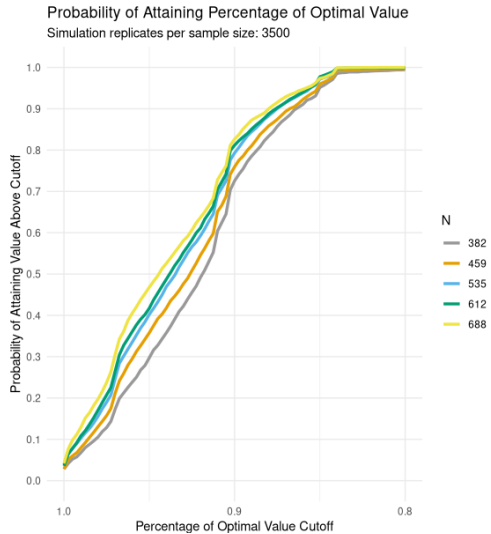
$$P\left(\frac{\mathrm{V}(\hat{\pi}_n)}{\mathrm{V}(\pi^*)} \geq \delta\right), \quad \delta \in [0, 1]$$

3. Ensures the estimated treatment policy is close to optimal, and the optimal policy and value are known in simulation.

4. For the study we chose $\delta = .9$ and 80% power.

BEST Sample Size Determination

Enrolling 740 participants yields at least an 80% probability to attain a 90% of the optimal value in our simulation scenario, assuming a dropout rate of not greater than 15% at 24 weeks.



Minimization[6] with Contraindications

- Stringent exclusion criteria can reduce the external validity of RCT results.
- Many cLBP patients have comorbidities and have previously attempted treatments.
- **Compromise:** BEST participants may be contraindicated for one of the four study treatments.

Minimization is used to maintain balance in important prognostic covariates across treatments, but must be modified to allow for contraindications.

[6] Stuart J. Pocock and Richard Simon. "Sequential Treatment Assignment with Balancing for Prognostic Factors in the Controlled Clinical Trial". In: *Biometrics* 31.1 (1975), pp. 103–115; Donald R. Taves. "Minimization: A New Method of Assigning Patients to Treatment and Control Groups". In: *Clinical Pharmacology & Therapeutics* 15.5 (1974), pp. 443–453.

Minimization with Contraindications

Input: Number of Arms K, History H_t, Imbalance function g, Imbalance Weights ω, Randomization Bias ρ

Output: Treatment Assignment for the New Patient

1. Observe covariate vector x_t and contraindicated treatment c
2. Calculate the imbalance score for each treatment arm k if the patient were to be assigned to that arm: $S_k = g(H_t, x, \omega, k)$
3. Define the feasible arm set $\mathcal{A}_t = [K]\ c$
4. Find the treatment arm with the minimum imbalance score from the set of feasible arms: $k^* = \operatorname{argmin}_{\mathcal{A}_t} S'_k$
5. Randomly assign the patient to treatment arm k^* with probability ρ or to another feasible arm with probability $(1-\rho)\frac{\alpha_k}{\sum_{k' \in \{\mathcal{A}_t k^*\}} \alpha_{k'}}$
6. Update the history H_{t+1} with the participant's covariates and assigned treatment arm $H_{t+1} = H_t \cup (x, a)$

Current Status of BEST

1. The BEST trial is underway and the statistical design paper is close to submission for peer review

2. The BEST trial has met and exceeded targeted enrollment

3. The statistical analysis plan (SAP) for the primary aim has been drafted and is being finalized

4. Acknowledgement: BEST (besttrial.org) is part of the Back Pain Consortium (BACPAC) Research Program which is funded through the NIH HEAL Initiative. Many centers and individuals have contributed to this project.