Utilizing Technology, Online Outreach, and Statistical Methods to Amplify Patient-Centric Research in Personalized N-of-1 Trials for Behavioral Health Interventions

2024
BOSTON

SCT | 45TH ANNUAL MEETING

Personalized N-of-1 Trials:

A Very Brief Introduction

Mark Butler, PhD
Institute of Health System Science
Feinstein Institutes for Medical Research
Northwell Health

Personalized Trials

Customize medical decisions and treatments according to the unique properties and responses of the individual

A Paradigm Shift in Treatment

Transitioning from One-Size-Fits-All to a Personalized Trial model

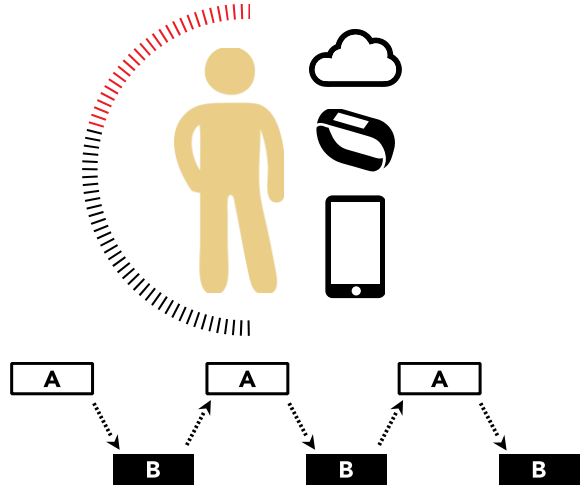
One-Size-Fits-All

One treatment based on large studies

Stratified

Different treatments based on sub-types

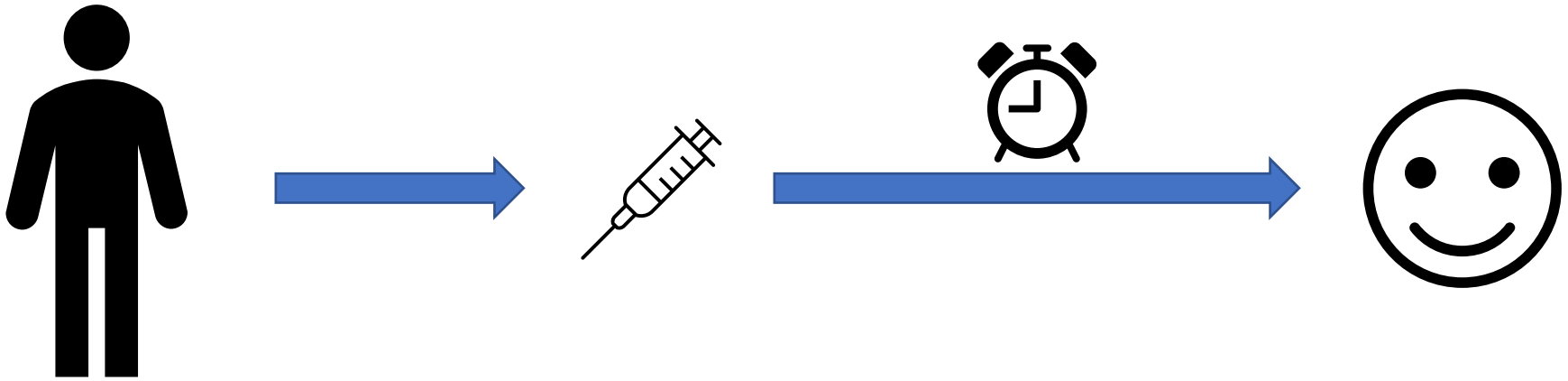
Personalized Trials



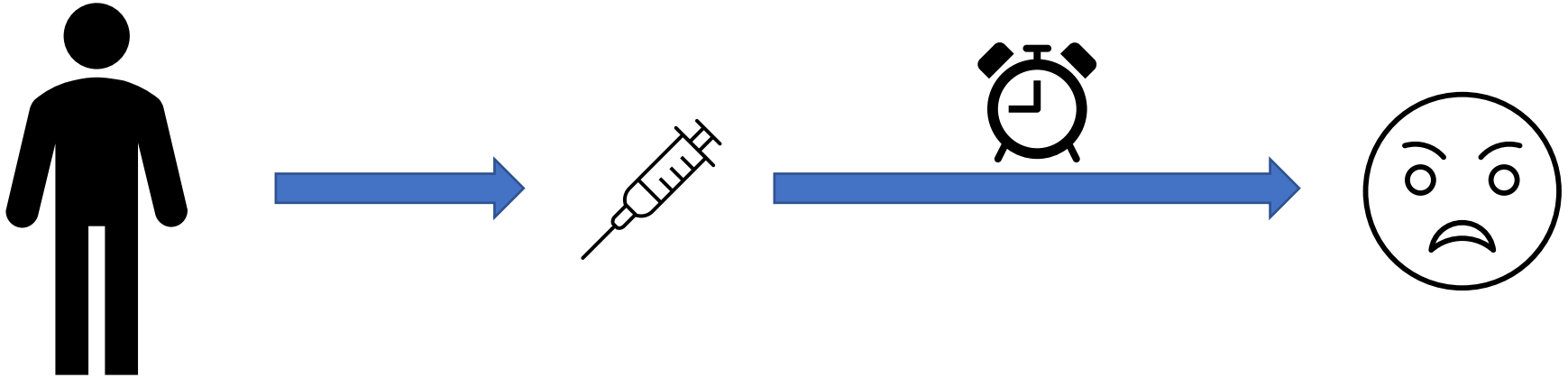
Treatment based on the individual needs of that patient

Common Clinical Practice

Patient 1



Patient 2



?

Why does a treatment work for one patient and not another?

PERSPECTIVE
https://doi.org/10.1038/s41562-021-01143-3

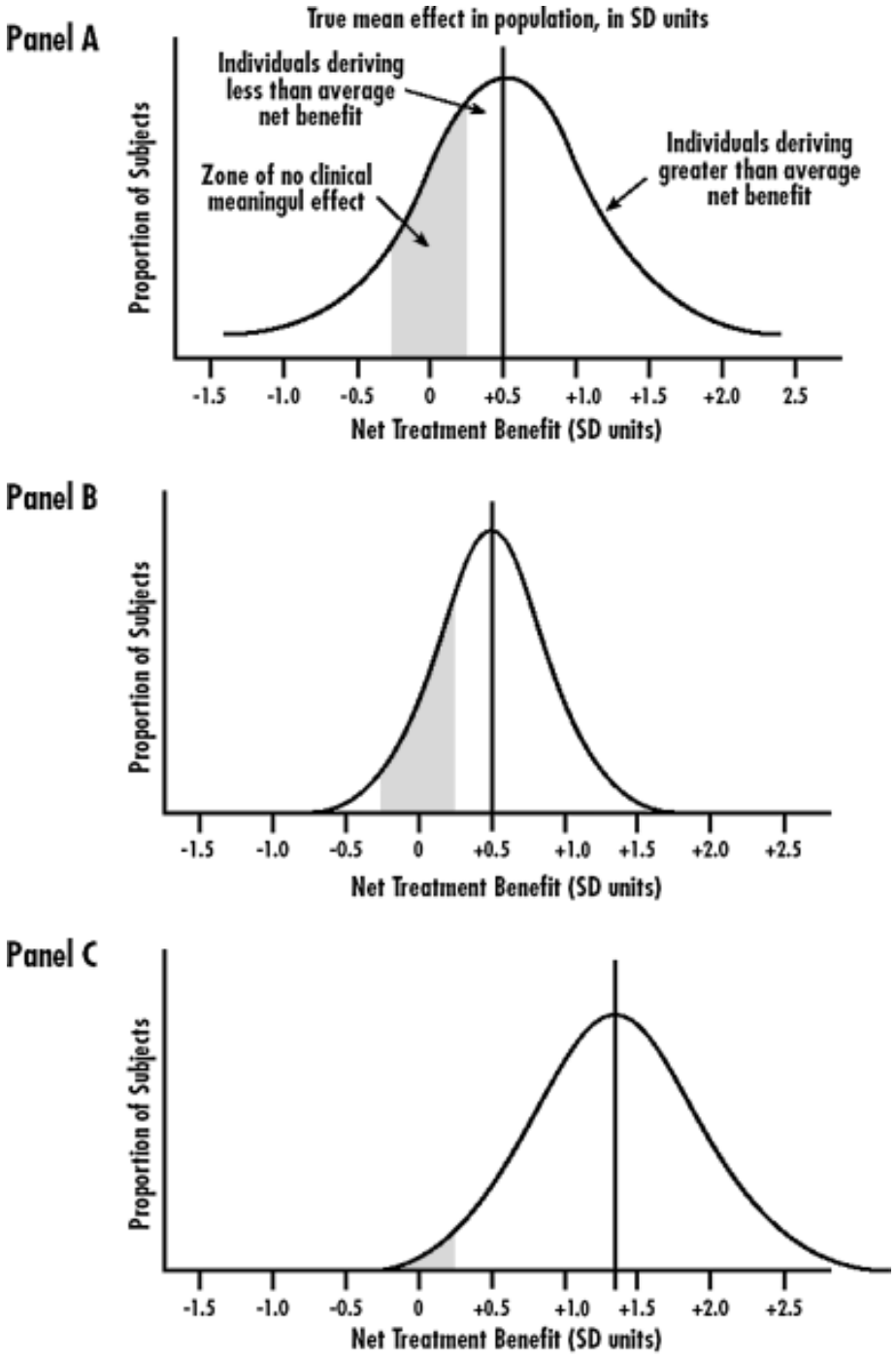
nature human behaviour

Behavioural science is unlikely to change the world without a heterogeneity revolution

Christopher J. Bryan[1]✉, Elizabeth Tipton[2]✉ and David S. Yeager[1]✉

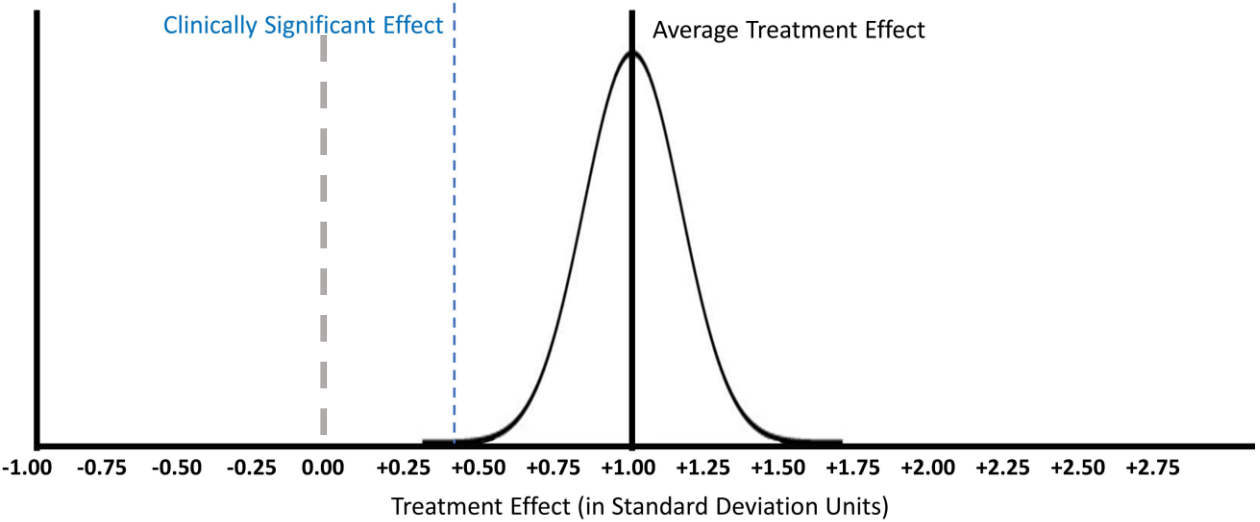
In the past decade, behavioural science has gained influence in policymaking but suffered a crisis of confidence in the replicability of its findings. Here, we describe a nascent heterogeneity revolution that we believe these twin historical trends have triggered. This revolution will be defined by the recognition that most treatment effects are heterogeneous, so the variation in effect estimates across studies that defines the replication crisis is to be expected as long as heterogeneous effects are studied without a systematic approach to sampling and moderation. When studied systematically, heterogeneity can be leveraged to build more complete theories of causal mechanism that could inform nuanced and dependable guidance to policymakers. We recommend investment in shared research infrastructure to make it feasible to study behavioural interventions in heterogeneous and generalizable samples, and suggest low-cost steps researchers can take immediately to avoid being misled by heterogeneity and begin to learn from it instead.

Heterogeneity of Treatment Effects (HTE)

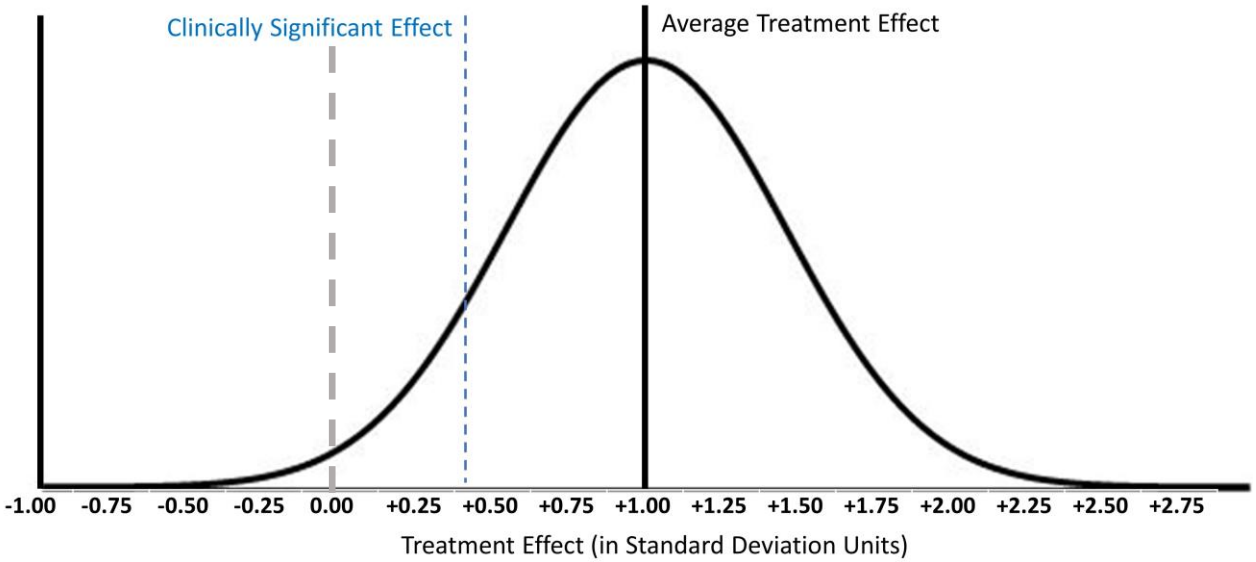


Source: Richard L. Kravitz, Naihua Duan, and Joel Braslow. "Evidence-based medicine, heterogeneity of treatment effects, and the trouble with averages." *The Milbank Quarterly* 82.4 (2004): 661-687.

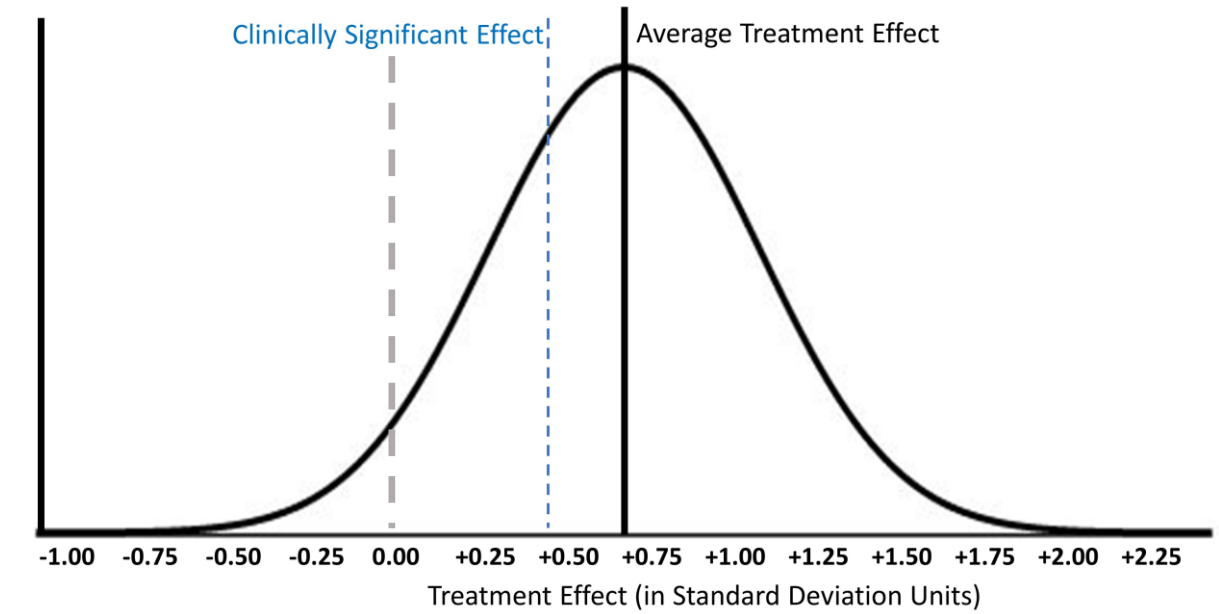
Panel 1: Low Variation, Large Magnitude Effect



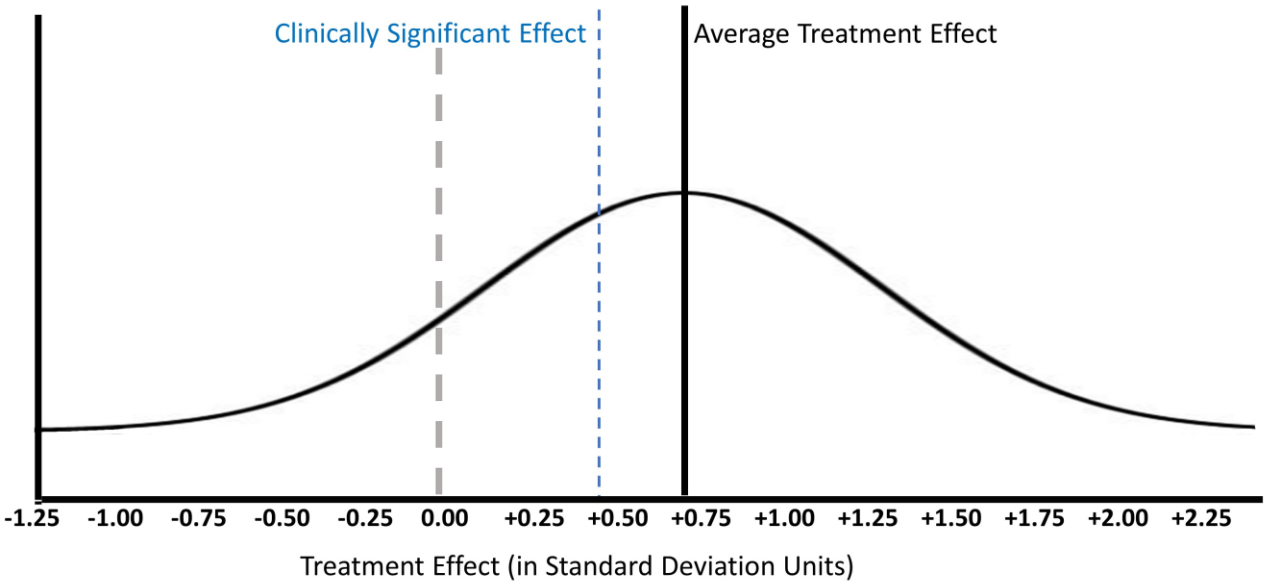
Panel 2: Moderate Variation, Large Magnitude Effect



Panel 3: Moderate Variation, Medium Magnitude Effect



Panel 4: Large Variation, Medium Magnitude Effect



What factors influence heterogeneity?

Sample Characteristics

Clinical trials seldom use representative samples:

- Lack of representation in terms of race/ethnicity, sex, gender, etc.
- Disease/condition severity and etiology
- Willingness to participate or engage in an intervention
- Environmental and social context

Patient preferences

- A conjoint survey in 500 respondents
- Elicit preference about the design parameters of N-of-1 trials, including treatment types, patient commitment level, cost, study duration, etc.

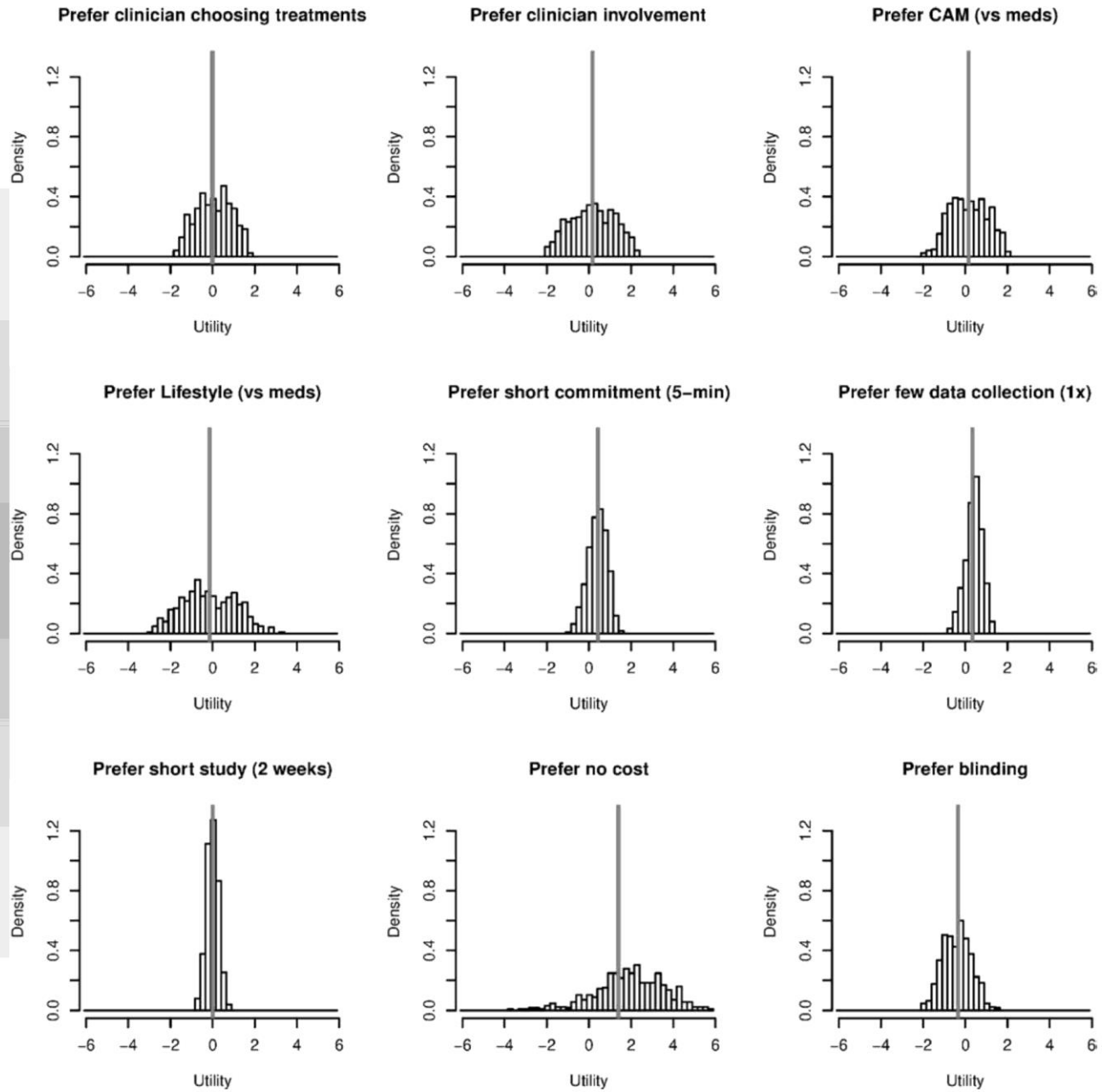


Figure 3. Cheung et al. (BMJ Open 2020)

How do we address heterogeneity?

JAMA Pediatrics | Special Communication

Experimental Designs to Optimize Treatments for Individuals
Personalized N-of-1 Trials

Karina W. Davidson, PhD, MASc; Michael Silverstein, MD, MPH; Ken Cheung, PhD; Rocco A. Paluch, MA; Leonard H. Epstein, PhD

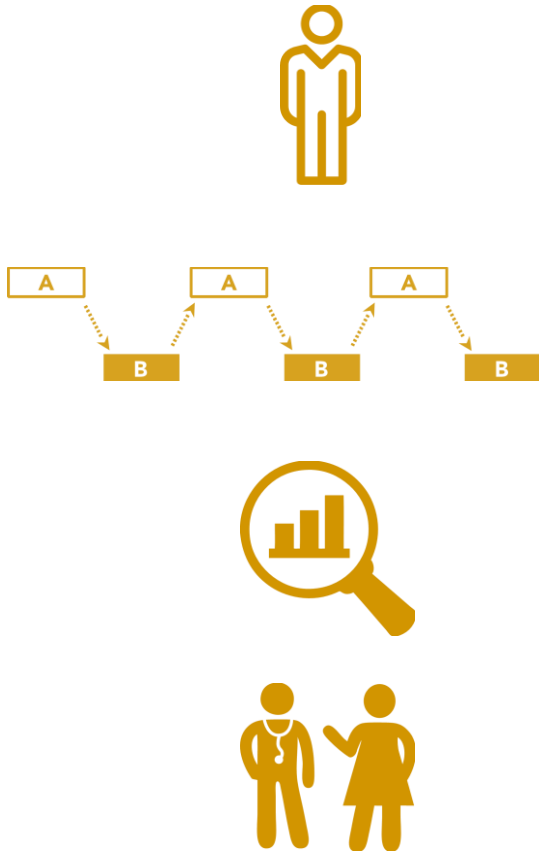
Conventional randomized clinical trials (RCTs) compare treatment effectiveness to provide support for evidence-based treatments that can be generalized to the average patient. However, the information obtained from RCTs may not always be useful for selecting the best treatment for individual patients. This article presents a complementary approach to identifying optimized treatments using experimental designs that focus on individuals. Personalized, or N-of-1, designs provide both a comparative analysis of treatments and a functional analysis demonstrating that changes in patient symptoms are likely because of the treatment implemented. This approach contributes to the zeitgeist of personalized medicine and provides clinicians with a paradigm for investigating optimal treatments for rare diseases for which RCTs are not always feasible, identifying personally effective treatments for patients with comorbidities who have historically been excluded from most RCTs, handling clinical situations in which patients respond idiosyncratically (either positively or negatively) to treatment, and shortening the time lag between identification and implementation of an evidence-based treatment. These designs merge experimental analysis of behavior methods used for decades in psychology with new methodological and statistical advances to assess significance levels of changes in individual patients, and they can be generalized to larger populations for meta-analytic purposes. This article presents a case for why these models are needed, an overview of how to apply personalized designs for different types of clinical scenarios, and a brief discussion of challenges associated with interpretation and implementation of personalized designs. The goal is to empower pediatricians to take personalized trial designs into clinical practice to identify optimal treatments for their patients.

JAMA Pediatr. 2021;175(4):404-409. doi:10.1001/jamapediatrics.2020.5801
Published online February 15, 2021.

Author Affiliations: Center for Personalized Health, Northwell Health, New York, New York (Davidson); Donald and Barbara Zucker School of Medicine at Hofstra University/Northwell Health, Hempstead, New York (Davidson); Boston University, Boston, Massachusetts (Silverstein); Mailman School of Public Health, Columbia University, New York, New York (Cheung); Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, Buffalo, New York (Paluch, Epstein).

Corresponding Author: Karina W. Davidson, PhD, MASc, Center for Personalized Health, Northwell Health, 130 E 59th St, Ste 14C, New York, NY 10022 (kdavidson2@northwell.edu).

What are Personalized, N-of-1 Trials?



Single patient, multiple crossover design, often randomized

Systematic collection of data on treatment effects

Rigorous statistical analysis

Shared decision-making

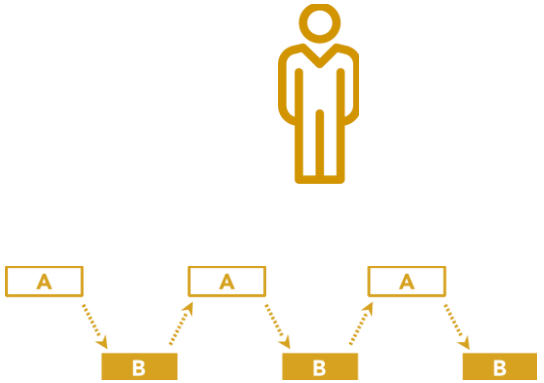
History of Personalized N-of-1 Trials?

- Single-case research is not new
- Many disciplines including medicine, psychology, and education use single case observational and case-study methods to explore topics
- Personalized N-of-1 Trials really began to be refined in 1986 by Guyatt and colleagues
- As time went on, personalized N-of-1 trials were adopted for lots of outcomes in both medicine and psychology, with over 1,000 personalized N-of-1 trials being conducted in the domain of health psychology alone

Guyatt, G., Sackett, D., Taylor, D. W., Ghong, J., Roberts, R., & Pugsley, S. (1986). Determining optimal therapy—randomized trials in individual patients. New England journal of medicine, 314(14), 889-892.

Shaffer JA, Kronish IM, Falzon L, Cheung YK, Davidson KW. N-of-1 randomized intervention trials in health psychology: a systematic review and methodology critique. Annals of Behavioral Medicine. 2018 Sep;52(9):731-42.

Personalized N-of-1 Trials of Massage and Yoga for Chronic Lower Back Pain



Single patient, multiple crossover design, often randomized

Systematic collection of data on treatment effects

Rigorous statistical analysis

Shared decision-making

Personalized N-of-1 Trials of Massage and Yoga for Chronic Lower Back Pain

Baseline Assessment

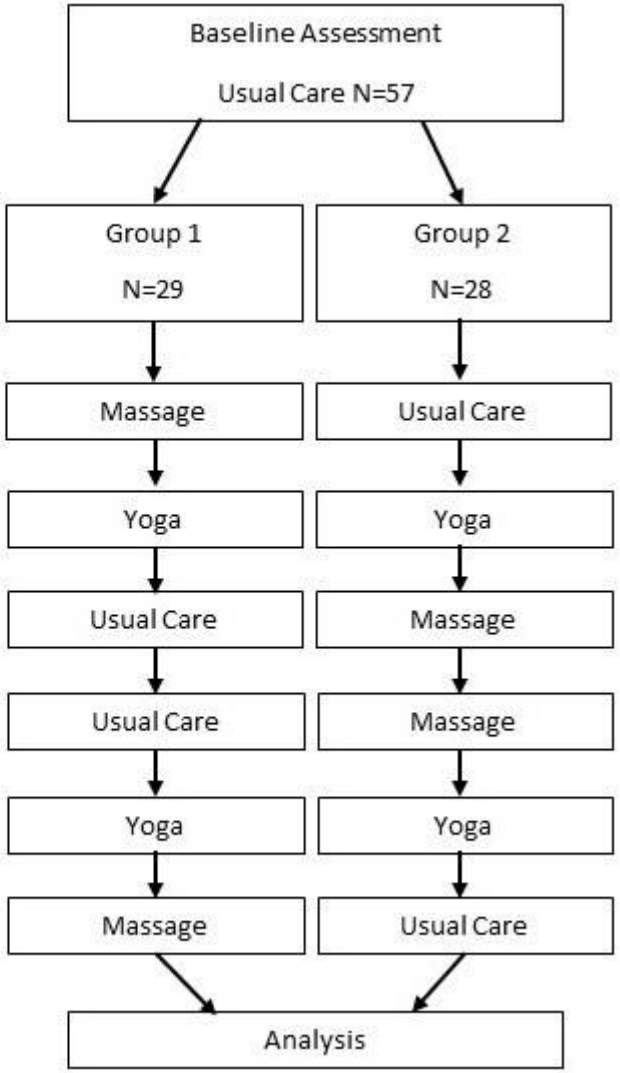
2 weeks in length

Randomization

Treatment Phase

6 Periods, 2 weeks in length

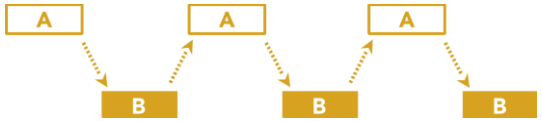
12 weeks total



- Continuous measurement of outcomes throughout the trial duration including:
  - PROMIS Pain Intensity
  - PROMIS Pain Interference
  - Ecological Momentary Assessment 3 time daily of pain, fatigue, and stress
  - Self-reported use of over-the-counter pain medication
  - Self-reported treatment side effects
  - Fitbit Physical Activity

Personalized N-of-1 Trials of Massage and Yoga for Chronic Lower Back Pain

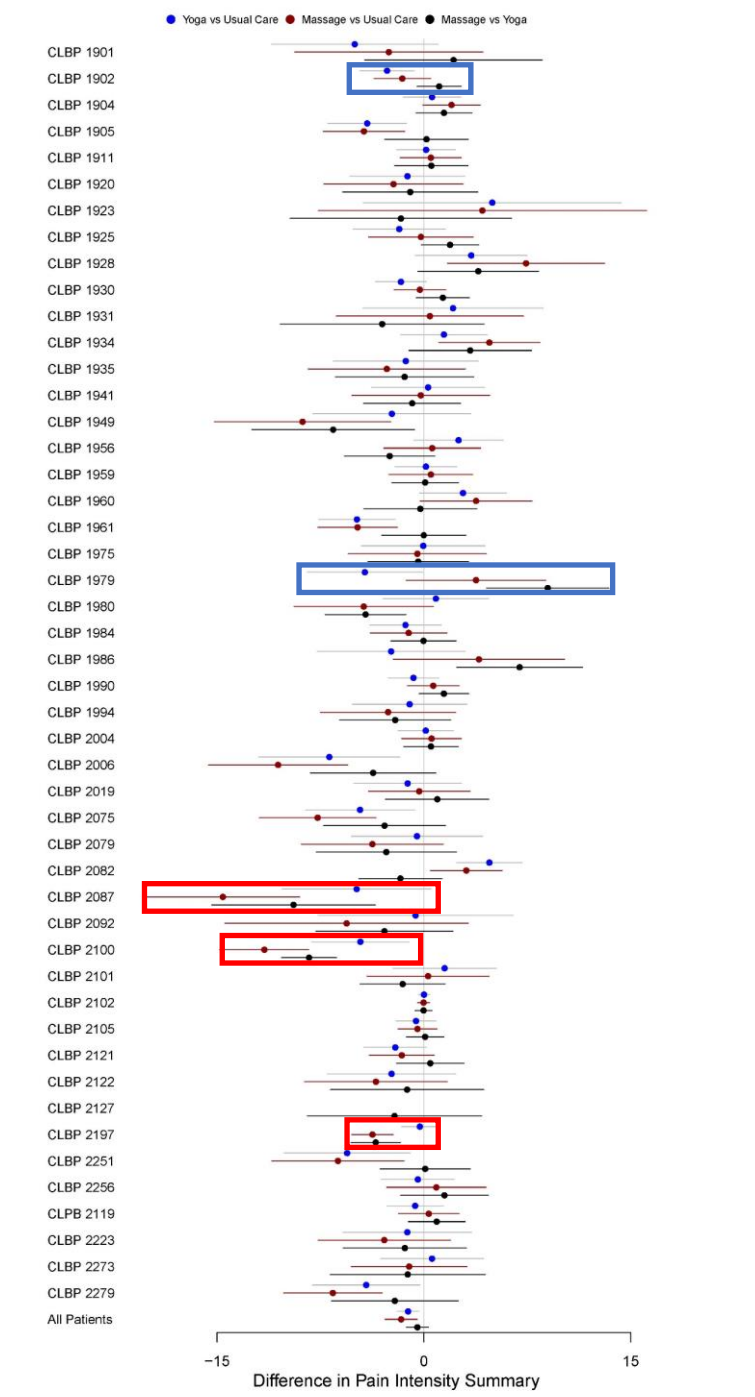
Single patient, multiple crossover design, often randomized

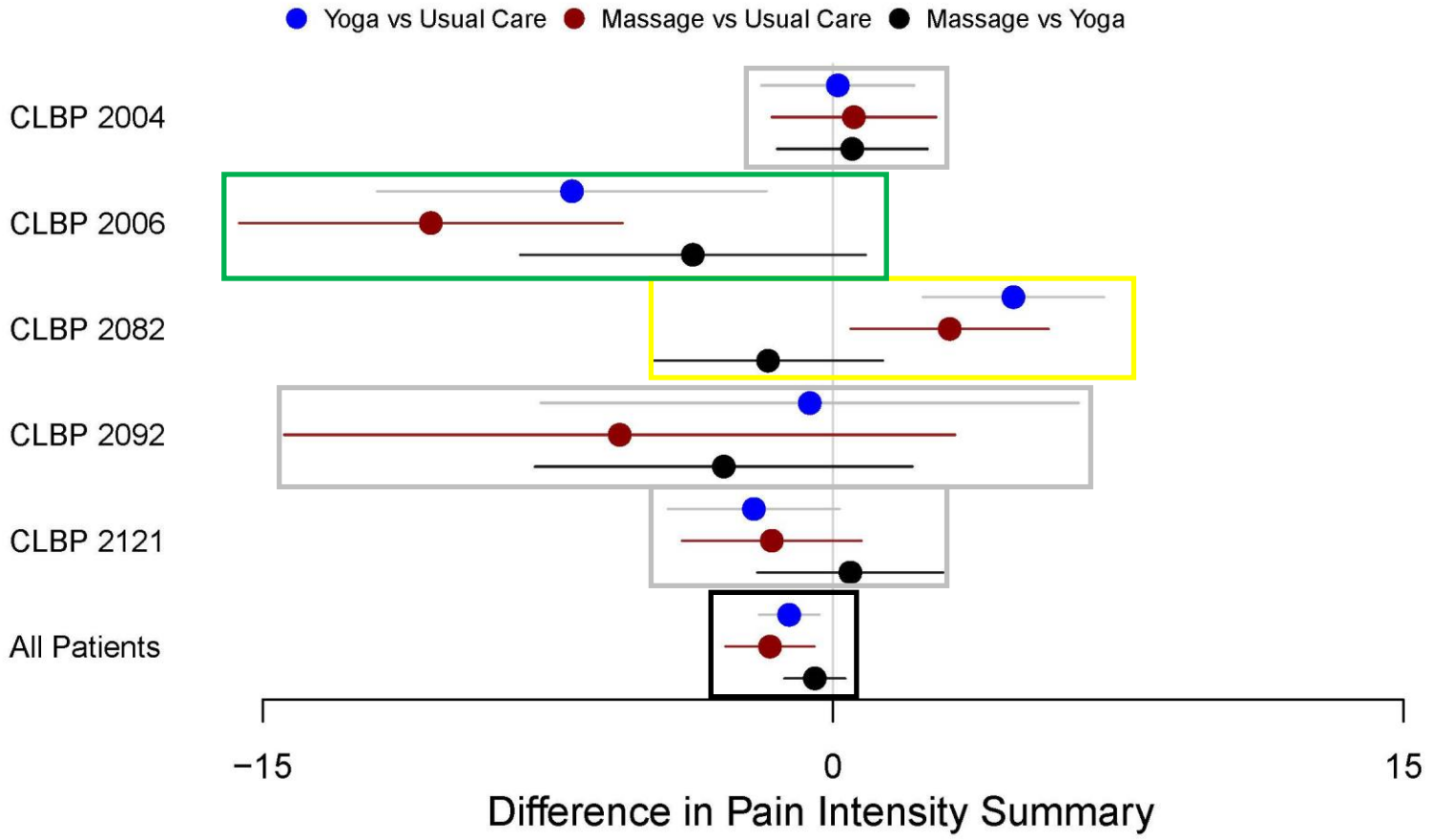


Systematic collection of data on treatment effects

Rigorous statistical analysis

Shared decision-making





Yoga worked better

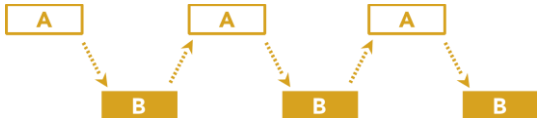
Massage Worked Better

Both Worked

Neither Worked

Both Hurt

Personalized N-of-1 Trials of Massage and Yoga for Chronic Lower Back Pain



Single patient, multiple crossover design, often randomized

Systematic collection of data on treatment effects

Rigorous statistical analysis

Shared decision-making

Personalized Feedback

1

Visual Summary

YOGA

Yoga helped your back pain!

During yoga weeks, you reported *less* pain intensity* and pain interference* than during usual care. You reported *decreased* momentary pain one day after treatments compared to usual care. We also observed some positive health benefits during yoga weeks that may be of interest to you

For you, yoga helped manage your chronic lower back pain *best*.

MASSAGE

Massage helped some aspects of your back pain.

During massage weeks, you reported *less* pain intensity* and pain interference* than during usual care. However, you reported *increased* momentary pain one day after treatments compared to usual care. We also observed some positive health benefits during massage weeks that may be of interest to you

2

Visual Summary

MASSAGE

Massage helped *some* aspects of your back pain, but not all.

During massage weeks you reported a *decrease* in momentary pain* one day after treatments. However, you also reported *higher* pain intensity* and *higher* pain interference compared to usual care. We observed a few positive health benefits during massage weeks that may be of interest to you.

For you, massage helped manage your chronic lower back pain *best*.

YOGA

We did not observe any trends suggesting yoga may have helped your back pain.

During yoga weeks you reported *higher* pain intensity* and *higher* pain interference compared to usual care, as well as *slightly increased* momentary pain one day after treatments. We observed a few positive health benefits during yoga weeks that may be of interest to you.

3

Visual Summary

MASSAGE

Massage helped some aspects of your back pain!

During massage weeks you reported *lower* pain intensity and *lower* pain interference* compared to usual care. However, you reported a *slight increase* in momentary pain one day after treatments. We also observed several positive health benefits during massage weeks that may be of interest to you.

For you, both massage and yoga helped manage your chronic lower back pain.

YOGA

Yoga helped some aspects of your back pain, too.

During yoga weeks you reported *slightly lower* pain intensity and *lower* pain interference* compared to usual care. You reported *minimal change* in momentary pain one day after treatments. We also observed several positive health benefits during massage weeks that may be of interest to you.

4

Visual Summary

MASSAGE

Massage helped your back pain!

During massage weeks you reported *less* pain intensity and *less* pain interference than during usual care.

We also observed several positive health benefits during massage weeks that may be of interest to you.

For you, both massage and yoga helped certain aspects of your back pain.

YOGA

Yoga helped your back pain, too!

During yoga weeks you also reported *less* pain intensity and *less* pain interference than during usual care.

We also observed some positive health benefits during yoga weeks that may be of interest to you.

5

Visual Summary

YOGA

We did not observe trends suggesting yoga may have helped your back pain.

During yoga weeks, you reported *slightly higher* pain intensity and *higher* pain interference* than during usual care. However, we did observe several positive health benefits during yoga weeks that may be of interest to you.

Neither yoga nor massage helped manage your chronic lower back pain.

MASSAGE

We did not observe trends suggesting massage may have helped your back pain.

During massage weeks, you reported *slightly higher* pain intensity and *higher* pain interference* than during usual care. However, we did observe several positive health benefits during massage weeks that may be of interest to you.

6

Visual Summary

YOGA

MASSAGE

There was insufficient data due to treatment adherence to make a conclusion on whether yoga or massage helped manage your chronic lower back pain.

Example Pages of Pain Patterns:
Pain Intensity, Pain Interference, Side Effects, Medication

Pain Patterns – Daily Recap

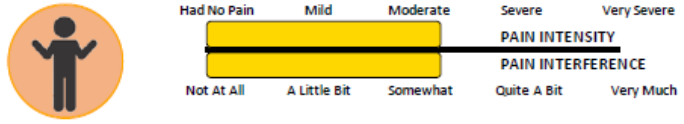
Pain intensity is how strong you perceive your pain to be. ***Pain interference*** is how much your pain gets in the way of your daily life. The treatment that gives you the **lowest** average pain intensity and **lowest** average pain interference is ideal.

During yoga treatment weeks, you reported *lower* pain intensity* and *lower* pain interference* compared to usual care.

During massage treatment weeks, you reported *lower* pain intensity* and *lower* pain interference* compared to usual care.

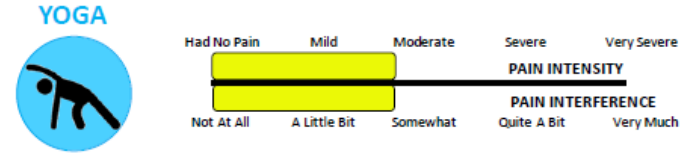
USUAL CARE

100% Completed



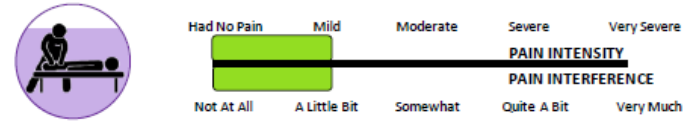
YOGA

100% Treatment Completed



MASSAGE

87.5% Treatment Completed

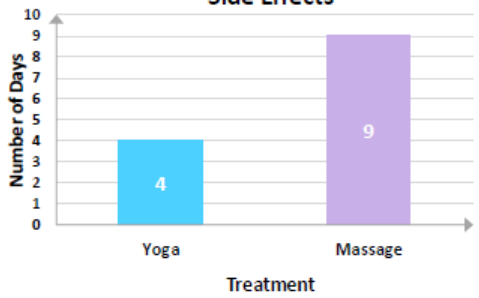


Please see the *Pain Patterns* video for an explanation of this page.

*This result was statistically significant.

Pain Patterns – Daily Recap, cont'd.

Side Effects



You reported **13 days** with side effects during treatment blocks. The treatment that gives you the ***least*** number of days with side effects is ideal.

During yoga treatment blocks, you reported the *least* days with side effects.

Medication

You took the following medications for your lower back pain throughout the study:

Medication Name	*Amount of Times Taken*	
	Over-the-Counter	Prescription-Strength
Acetaminophen (e.g. Tylenol)	11 10 8	
Naproxen (e.g. Aleve, Naprosyn)		
Ibuprofen (e.g. Motrin, Advil)		
Salicylic Acid (e.g. Aspirin)		
Other		

Total:
11 times during usual care weeks
10 times during yoga weeks
8 times during massage weeks

Please see the *Pain Patterns* video for an explanation of this page.

Example Pages of Study Adherence:

High Adherence

Adherence

Adherence is based on how closely you followed study protocol. You are adherent when you wear your Fitbit, complete yoga and massage treatments, and answer your surveys. The more adherent you are, the more accurate your results are.

97%

Total Adherence Score

		Fitbit	Treatment	Surveys
MASSAGE	Weeks 3-4 2/17 - 2/23 2/24 - 3/1			Evening: 100% Check-Ins: 95.2% Weekly: 100%
YOGA	Weeks 5-6 3/2 - 3/8 3/9 - 3/15			Evening: 100% Check-Ins: 97.6% Weekly: 100%
USUAL CARE	Weeks 7-8 8/10 - 8/16 8/17 - 8/23			Evening: 100% Check-Ins: 88.1% Weekly: 100%
USUAL CARE	Weeks 9-10 8/24 – 8/30 8/31 – 9/6			Evening: 100% Check-Ins: 81.0% Weekly: 100%
YOGA	Weeks 11-12 9/7 – 9/13 9/14 – 9/20			Evening: 100% Check-Ins: 81.0% Weekly: 100%
MASSAGE	Weeks 13-14 9/21 – 9/27 9/28 – 10/4			Evening: 100% Check-Ins: 85.7% Weekly: 100%

Total Fitbit Adherence:	*Total Treatment Adherence:*	*Total Survey Adherence:*
100%	100%	95.8%

Please see the *Adherence* video for an explanation of this page.

Low Adherence

Adherence

Adherence is based on how closely you followed study protocol. You are adherent when you wear your Fitbit, complete yoga and massage treatments, and answer your surveys. The more adherent you are, the more accurate your results are.

77%

Total Adherence Score

		Fitbit	Treatment	Surveys
MASSAGE	Weeks 3-4 8/10 - 8/16 8/17 - 8/23			Evening: 66.7% Check-Ins: 73.8% Weekly: 50.0%
YOGA	Weeks 5-6 8/24 - 8/30 8/31 - 9/6			Evening: 83.3% Check-Ins: 83.3% Weekly: 100%
USUAL CARE	Weeks 7-8 9/7 - 9/13 9/14 - 9/20			Evening: 83.3% Check-Ins: 66.7% Weekly: 100%
USUAL CARE	Weeks 9-10 9/21 - 9/27 9/28 - 10/4			Evening: 75% Check-Ins: 76.2% Weekly: 100%
YOGA	Weeks 11-12 10/5 – 10/11 10/12 – 10/18			Evening: 66.7% Check-Ins: 64.3% Weekly: 0%
MASSAGE	Weeks 13-14 10/19 – 10/25 10/26 – 11/1			Evening: 41.7% Check-Ins: 35.7% Weekly: 50%

Total Fitbit Adherence:	*Total Treatment Adherence:*	*Total Survey Adherence:*
63.1%	100%	67.6%

Please see the *Adherence* video for an explanation of this page.

Example Page of Momentary Pain, Stress & Fatigue Patterns

Momentary Pain, Stress, & Fatigue Patterns

PAIN

Your **highest average pain levels** were from **6 PM to 9 PM** during **yoga** weeks

STRESS

Your **highest average stress levels** were from **6 PM to 9 PM** during **usual care** weeks

FATIGUE

Your **highest average fatigue levels** were from **12 PM to 3 PM** during **usual care** weeks

Compared to Usual Care...

ONE DAY AFTER YOGA

- ✓ Pain levels **decreased** by 10.2%
- ✓ Stress levels **decreased** by 19.6%
- ✗ Fatigue levels **increased** by 51.2%

ONE DAY AFTER MASSAGE

- ✓ Pain levels **decreased** by 11.0%
- ✓ Stress levels **decreased** by 31.2%
- ⊘ Fatigue levels **had minimal change**

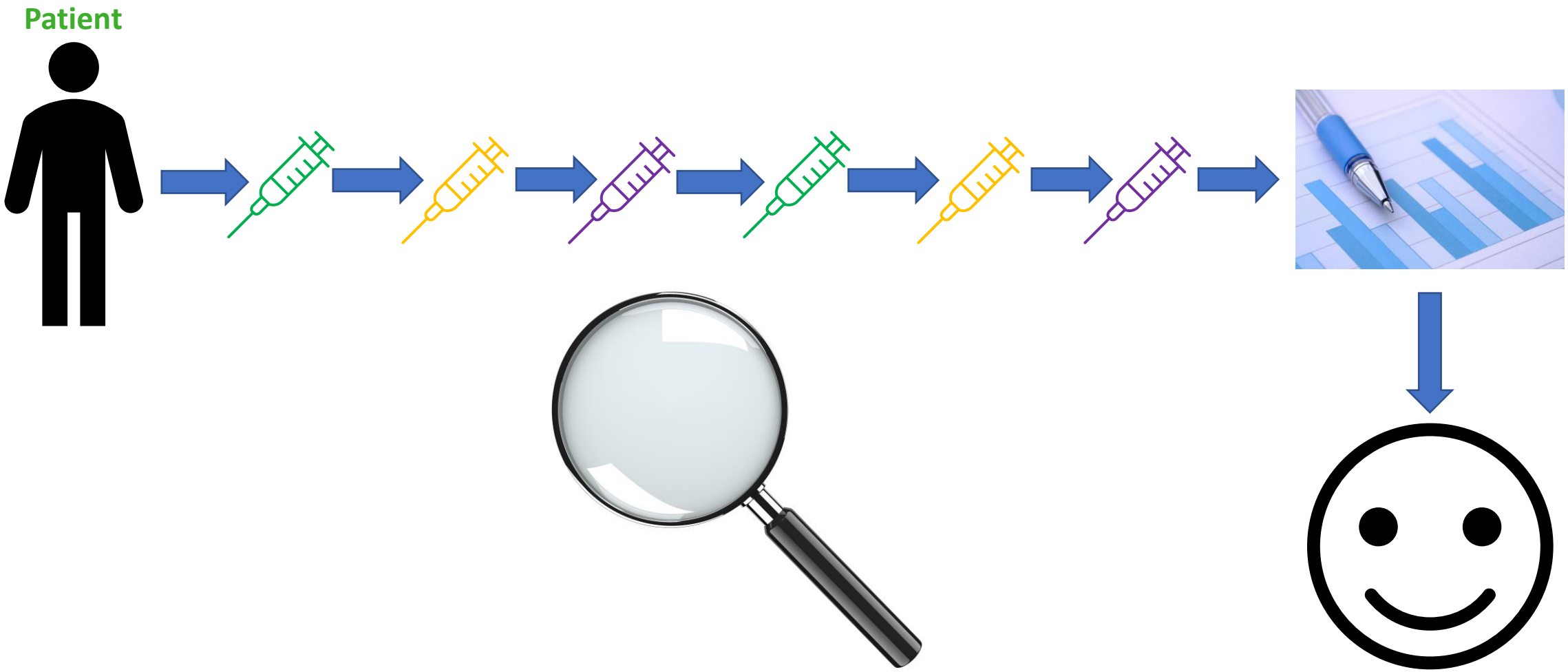
One day after **yoga** you reported *decreased* momentary pain and stress, and *increased* momentary fatigue* compared to **usual care**.

One day after **massage** you reported *decreased* momentary pain and stress, and *minimal change* in momentary fatigue compared to **usual care**.

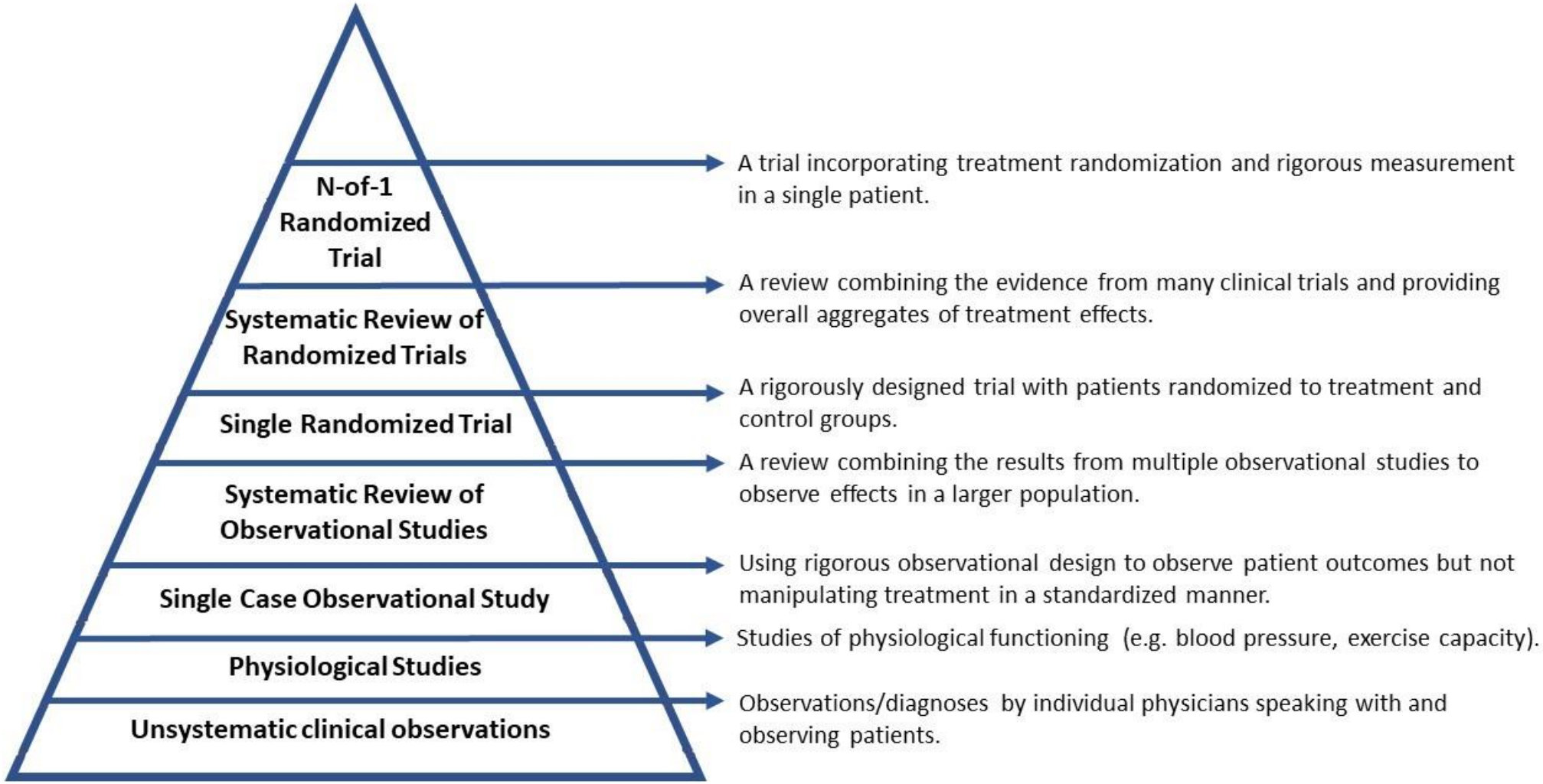
Please see the *Pain, Stress, and Fatigue Patterns* video for an explanation of this page.

*This result was statistically significant.

Personalized N-of-1 Clinical Practice



Hierarchy of Evidence-Based Research

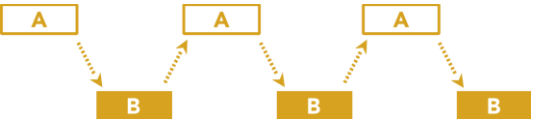


Guyatt, G., & Buse, W. (2006). The philosophy of evidence-based medicine. In Montori VM (Ed.), *Evidence-based endocrinology*, (pp. 25-33). Humana Press. https://doi.org/10.1007/978-1-59745-008-9_3.

If personalized N-of-1 trials are so great, why don't we use them all the time?

Single patient, multiple crossover design, often randomized

- Requires effort to design and implement
- Careful consideration of outcome/disease/condition
- Careful consideration of how treatments will affect the patient over time, interact with each other, etc.
- Needs to fit with an N-of-1 design



Systematic collection of data on treatment effects

- Requires a system for collecting and storing data
- Requires patients to actually complete the measures
- Requires measures which are rapidly reactive to the intervention

Rigorous statistical analysis

- Requires expertise
- Requires software
- Requires thought

Shared decision-making

- Requires synthesizing data in a form which is easily interpretable and actionable
- Requires communication and potentially clinical expertise

1. **Dr. Codruta Chiuzan –** *Design and Implementation of Personalized N-of-1 trials*

2. **Dr. Thevaa Chandereng –** *Analysis of Personalized N-of-1 Trials*

3. **Dr. Ziwei Liao –** *Modeling EMA Data Using Bayesian Methods*

Conclusion and Discussion - Dr. Ying Kuen Cheung

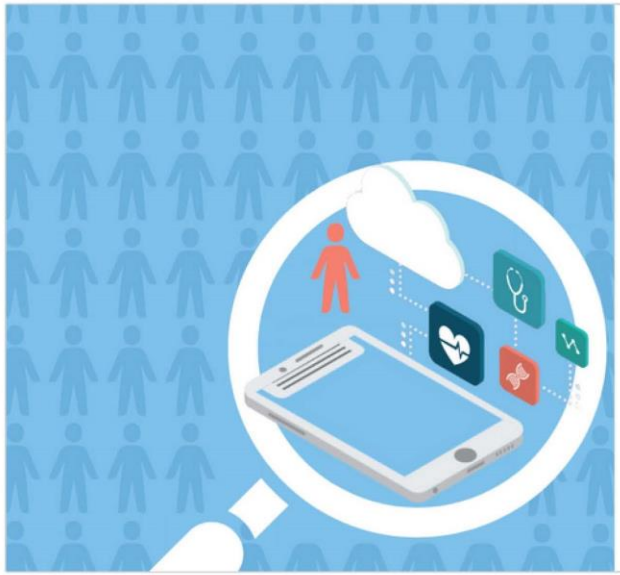
https://hdsr.mitpress.mit.edu/specialissue3

HDSR Search Dashboard Login or Signup

Personalized (N-of-1) Trials: Methods, Applications, and Impact

Special Issue 3

FROM THE EDITORS

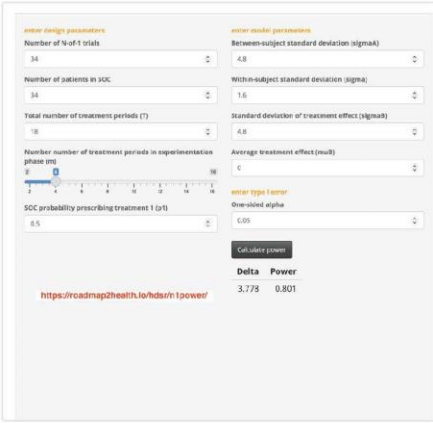


Introducing Data Sciences to N-of-1 Designs, Statistics, Use-Cases, the Future, and the Moniker 'N-of-1' Trial

by Karina Davidson, Ken Cheung, Ciaran Friel, and Jerry Suls

Published: Sep 08, 2022

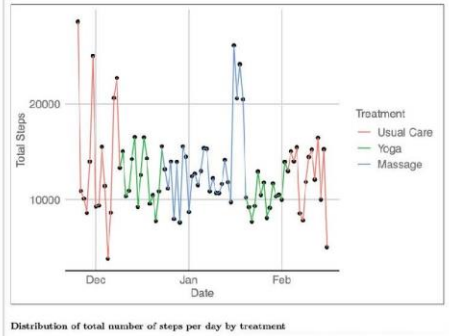
Special Issue 3: Personalized (N-of-1) Trials: Methods, Applications, and



Evaluating Personalized (N-of-1) Trials in Rare Diseases: How Much Experimentation Is Enough?

by Ken Cheung and Hiroshi Mitsumoto

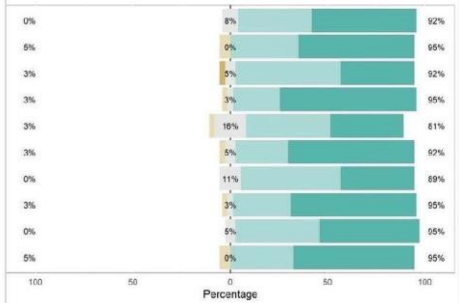
Published: Sep 08, 2022



An R Shiny App for a Chronic Lower Back Pain Study, Personalized N-of-1 Trial

by Thevaa Chandereng

Published: Sep 08, 2022



A Series of Virtual Interventions for Chronic Lower Back Pain: A Feasibility Pilot Study for a Series of Personalized (N-of-1) Trials

by Mark Butler, Stefani D'Angelo, Melissa Kaplan, Zarrin Tashnim, Danielle Miller, Heejoon Ahn, Louise Falzon, Andrew J. Dominello, and 4 more

Published: Sep 08, 2022

Acknowledgements

Karina W. Davidson, Phd, MASc
Stefani D'Angelo, MS
Joan Duer-Hefele, MA
Danielle Miller, MA
Heejoon Ahn, MS
Alexandra Perrin, BS
Elizabeth A. Vrany, PhD
Melissa Kaplan, MA
Jordyn Rodillas, BA
Anne-Marie Romain, BA
Ruby Weiner, BA
Theodoros P. Zanos, PhD






Society of Clinical Trials Annual Conference

May 20, 2024

Personalized N-of-1 Trials: Design, Implementation and Practical Considerations

Cody Chiuzan, PhD
Associate Professor
Institute of Health System Science

Feinstein Institutes for Medical Research
Northwell Health®

Outline

- Best-case uses and evolution over time
- Desing Considerations
- Implementation Aspects
- Example of an ongoing N-of-1 trial testing mind-body interventions
  - Challenges
  - Lessons learned
  - Next steps

N-of-1 Trials: Making It Personal

Goal: Finding the treatment with the greatest efficacy for that individual

- These multi-crossover trials use the participant as his/her self-control and the outcome of interest is evaluated across different treatment periods
  - Possible to aggregate for estimating the population average treatment effect (PATE)
- Because a participant experiences the intervention and the control over different times, some interventions are not appropriate for assessing in an N-of-1 trial

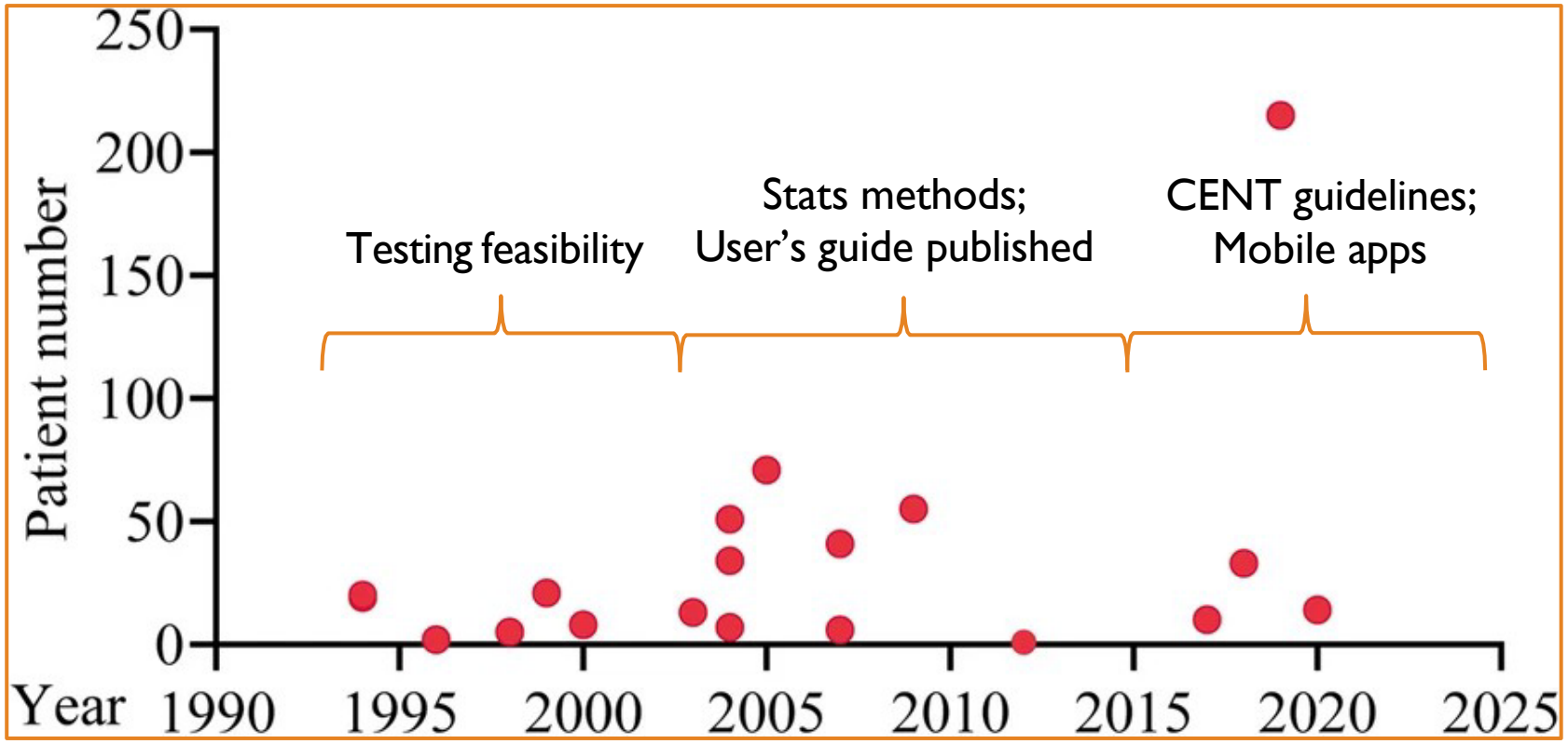
N-of-1 Trials: Making It Personal

N-of-1 trials are suitable for chronic conditions, which require long-term maintenance therapy, but **NOT** for:

- **Therapies** that have delayed effects
- **Rapidly progressive conditions**
- **Interventions with carry-over effects**, i.e., produce a change that remains for some time after the intervention stops
- **Highly invasive therapies**, typically involving multiple treatment cycles
- **Experimental treatments without established safety** profiles and/or unpredictable side effects

N-of-1 Trials in Various Clinical Settings

Recruitment patterns in N-of-1 trials in Chronic Pain Management



He,W., Cui, Z., Chen,Y. et al. Status of N-of-1 Trials in Chronic Pain Management: A Narrative Review. Pain Ther 10, 1013–1028 (2021).

N-of-1 Trials in Various Clinical Settings

Chronic Pain Management

- Majority of the trials were searching for treatments (e.g., analgesics) for osteoarthritis, chronic musculoskeletal pain and neuropathic pain
- N-of-1 trials required higher costs for conducting procedures, follow-up, analysis, but saved on the long-term (fewer medical interventions)
- Withdrawal rate, comparable to standard RCT (~30%)
- Wireless devices and smartphone apps essential for recording personal information

N-of-1 Trials in Various Clinical Settings

Clinical question	Outcome assessment
Evaluate individual responses to stimulants among children with ADHD	Parent, teacher, and self-rating of ADHD symptoms: assessed during and at the end of each treatment period
Verify perceived side effects of statin tablets	Self-report of side effects, symptoms severity assessed daily
Evaluate pharmacologic, electric stimulants and symptoms of neuromuscular and neurodegenerative disorders	Patient-reported outcomes (Likert scales); radiologic or neurophysiologic measures
Intra-patient pharmacokinetic-guided dose escalation in cancer patients with RET alteration	Grade 3 and 4 toxicities and drug plasma levels (AUC_{0-24})

Samuel JP, Wootton SH, Tyson JE. N-of-1 trials: The epitome of personalized medicine? J Clin Transl Sci. 2023 Jul 3;7(1)PMID: 37528940
Gouda, MA Buschhorn L, Schneeweiss A, et al; N-of-1 Trials in Cancer Drug Development. Cancer Discov 1 June 2023; 13 (6): 1301–1309.

N-of-1 Trials: Design Considerations

- **Intervention/Treatment**
  - Appropriate for repeated administration and reversible outcomes
  - The number of times a treatment is repeated can impact the precision of the relative effects of each treatment
  - Trial design is most feasible when the treatment effects have a rapid onset and offset of action
- **Washout Periods**
  - Mitigate carryover effects; treatment periods should be long enough to capture a difference in outcomes, but short to be practical for patient
- **Outcome Measures:** Measurable, Reliable and Clinically Relevant
  - Primary: condition-specific clinical assessments and/or validated instruments (baseline and at pre-defined timepoints)
    - Secondary: symptoms, side-effects
  - Convenient enough that patients are willing to assess repeatedly

N-of-1 Trials: Design Considerations

- **Randomization**
  - Randomization and counterbalancing attempt to balance treatments both withing and across blocks
  - **ABBA vs ~~AABA~~ or ~~ABAB~~**
- **Blinding**
  - If feasible, double blinding (patients and clinicians) is important, especially with self-reported outcomes

Kravitz, R. L., Duan, N., Vohra, S., Li, J., & The DEcIDE Methods Center N-of-1 Guidance Panel. (2014). Chapter 1. Introduction to N-of-1 trials: Indications and barriers. In R. L. Kravitz & N. Duan (Eds.), Design and implementation of N-of-1 trials: A user's guide: AHRQ publication no. 13(14)-EHC122-EF (pp. 1–11). AHRQ.

Implementation of N-of-1 Trials

- **Recruitment**
  - Identify eligible participants based on specific inclusion/exclusion criteria
- **Informed Consent**
  - Obtain informed consent from participants outlining the purpose, procedures, potential risks, and the opportunity to receive individualized treatment recommendations based on personal responses
- **Data Collection and Monitoring**
  - Rigorous data collection and monitoring procedures (adherence and response to interventions) are essential to ensure data accuracy and completeness
  - Electronic diaries and wearable devices for real-time data capture
- **Reporting Standards**
  - CONSORT Extension for N-of-1 Trials (CENT) guidelines

Personalized N-of-1 Trial of Mind-Body Interventions to Improve Sleep Duration

Try **yoga, guided walking,** and **mindfulness** in this **virtual, at-home** trial!

To participate in this study, you must:

- Be a female Northwell health-care employee between 40 and 60 years of age.
- Have a history of short sleep duration
- Self-report experiencing stress
- Own a smart phone that can receive text messages
- Be able to receive e-mails
- Be able to wear a Fitbit device

NIH U.S. National Library of Medicine

ClinicalTrials.gov

Personalized N-of-1 Trial of Mind-Body Interventions to Improve Sleep Duration

During COVID-19 pandemic, poor sleep and stress were highly prevalent among healthcare workers, especially women.

NIH Stage II Model for behavioral intervention development

Primary Objective: Determine if a personalized (N-of-1) trial employing a Mind-Body Intervention (MBI) can produce a meaningful increase in sleep duration among women aged 40-60 years working at Northwell Health, a major healthcare system in New York.

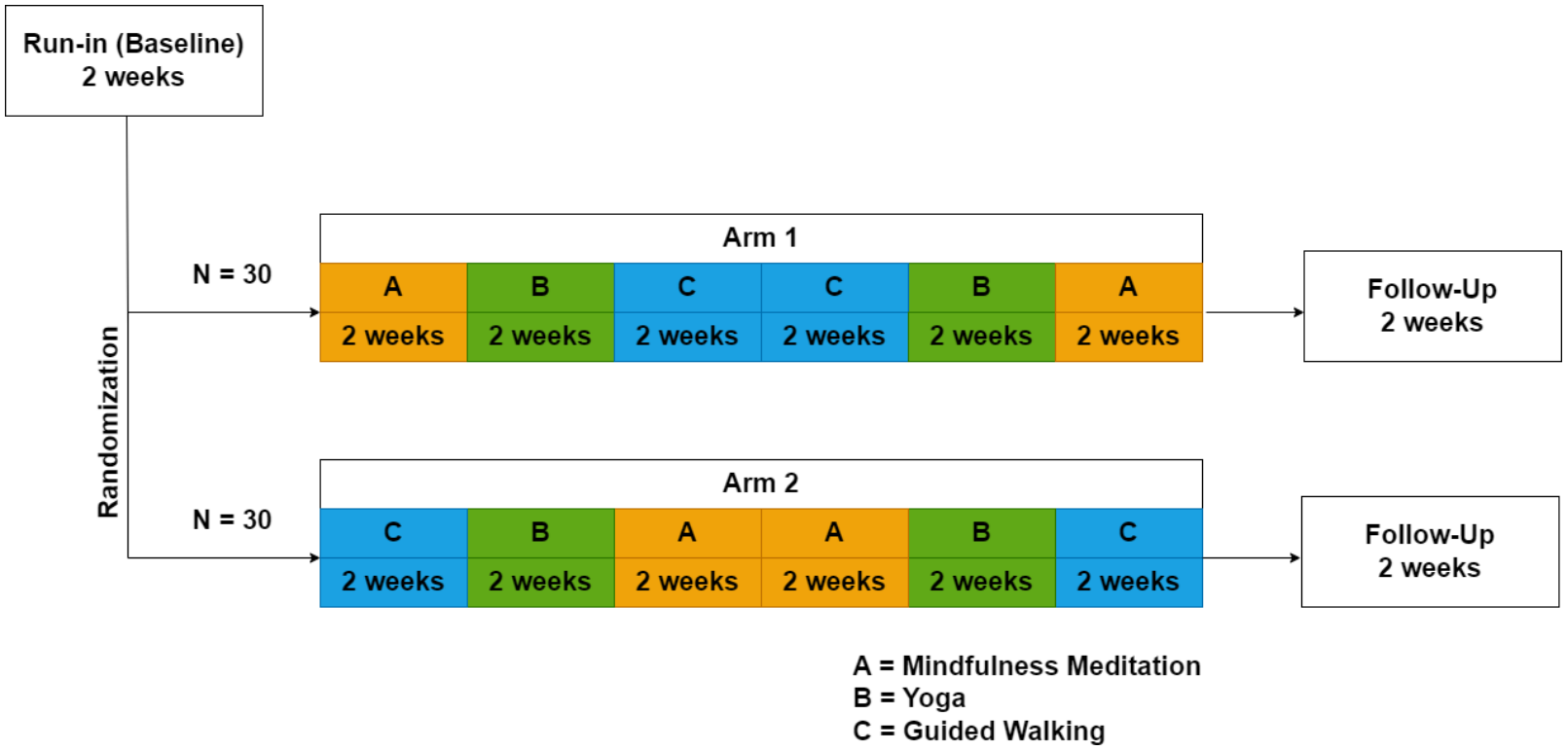
Secondary Objective(s): Evaluate the effect of the MBI on sleep quality components (sleep latency, efficiency), physiological factors (resting heart rate, number of steps), perceived stress, and anxiety and depression scores.

Personalized N-of-1 Trial of Mind-Body Interventions to Improve Sleep Duration

Outcome Measures Mean within-subject difference between run-in and each treatment period	**Assessment**
Sleep Duration	Daily via Fitbit devices and self-reported questionnaires
Sleep Latency	Daily via Fitbit devices and self-reported questionnaires
Sleep Efficiency	Daily via Fitbit devices
Resting Heart Rate and Number of Steps	Daily via Fitbit devices
Perceived Stress Score Anxiety and Depression Score	Run-in, bi-weekly, post-intervention Perceived Stress Scale (PSS-10) Patient Health Questionnaire4 (PHQ-4)

Personalized N-of-1 Trial of Mind-Body Interventions to Improve Sleep Duration

A total of 60 participants will be randomized 1:1 into one of two **12-week intervention arms**, each comprising six 2-week blocks of guided (A) mindfulness meditation, (B) yoga, (C) or guided walking.



Personalized N-of-1 Trial of Mind-Body Interventions to Improve Sleep Duration

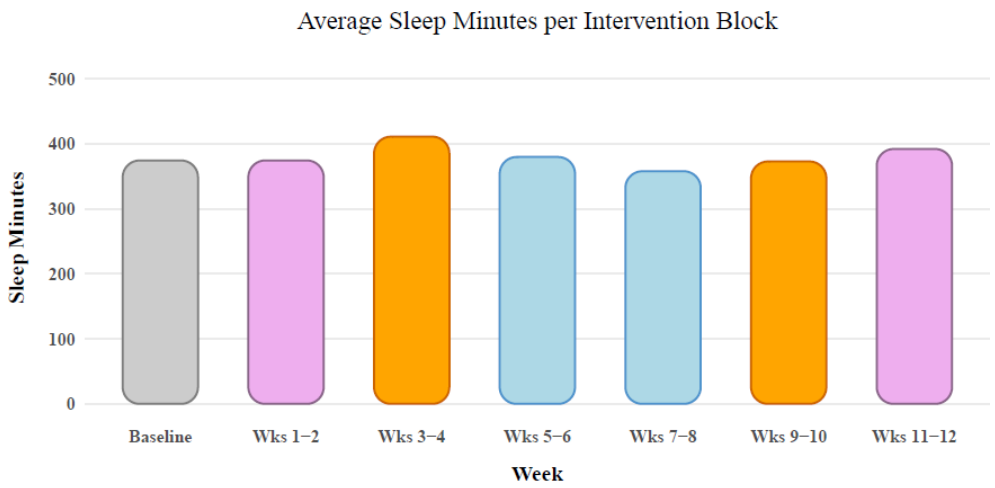
- **Run-in/Baseline (2 weeks)**
  - Participants are asked to wear the **Fitbit device 24/7**
  - Eligible for randomization if: 1) an average sleep duration <7 h per 24-h period and 2) achieved ≥80% adherence to Fitbit wear and survey completion
- **Intervention (12 weeks)**
  - The MBI consists of **30-min x 3 times per week** guided audiovisual recordings on mindfulness meditation, yoga, and walking created by Zeel
  - Study coordinators monitor the Fitbit sync (<5 days), survey responses (<50%), and intervention data (< 7 days)
- **Follow-up (2 weeks)**
  - No intervention
  - Send the personalized reports

Personalized Reports

Brief Summary

Our study team determines sleep duration by measuring the total number of sleep minutes within 24 hours, and has determined that the intervention with the highest sleep duration is the one which improved your sleep duration the most. This page of your report shares how each intervention **improved** your average daily **sleep duration**.

For you, yoga increased your average daily sleep duration the most.



Your Fitbit reported *less* minutes of sleep during mindfulness meditation weeks than during your **baseline** weeks.

Your Fitbit reported *more* minutes of sleep during yoga weeks than during your **baseline** weeks.

Your Fitbit reported *more* minutes of sleep during guided walking weeks than during your **baseline** weeks.

Personalized Reports

Expanded Summary

The colored boxes on this page summarize how the interventions may have impacted your **sleep duration** as well as other health benefits. The ✓ shows a *positive change*, while the ✗ shows a *negative change*. A ⊘ shows *minimal change*. A blank box indicates that changes could not be computed due to missing information.

Compared to **baseline**, during **mindfulness meditation** weeks, you also had changes in:

✓ Steps	✓ Sleep Latency
✗ Sleep Duration	✓ Sleep Quality*
✓ Sleep Efficiency*	✓ Stress
✗ Resting Heart Rate	✓ Anxiety/Depression

Compared to **baseline**, during **yoga** weeks, you also had changes in:

✓ Steps	✓ Sleep Latency*
✓ Sleep Duration	✓ Sleep Quality*
✓ Sleep Efficiency	✓ Stress
✗ Resting Heart Rate	✓ Anxiety/Depression

Compared to **baseline**, during **guided walking** weeks, you also had changes in:

✓ Steps	✓ Sleep Latency*
✓ Sleep Duration	✓ Sleep Quality*
✓ Sleep Efficiency*	✓ Stress
✗ Resting Heart Rate	✓ Anxiety/Depression

NOTE: This report is not meant to offer medical advice. The goal of this report is to help you understand how sleep interventions may impact your self–reported symptoms. You should seek a medical professional's opinion if you have any questions or concerns.

Challenges and Current Status

- Anticipated completion date: June 30, 2024 (on target!)
- High enthusiasm and interest to participate in the trial

- **What happened during the run-in phase?**
  - Baseline failures due to "normal" sleep duration
  - Low stress score (PSS<18)
- **What happened during the intervention?**
  - Issues with Fitbit sync
  - <3 times/week of intervention (video watching)
  - Limited survey response (stress, anxiety/depression)
  - Halfway check-in – **"You're a partner in our research!"**

Conclusions

- N-of-1 trials offer a pathway to provide patient-centered care by generating evidence directly from the individual patient
- A major strength is the opportunity to investigate potential heterogeneity of effects of the interventions
  - Pooling N-of-1 data may yield deeper insights into the effectiveness of the intervention
- Treatment assessments can be overwhelming; carefully select preferred and "naturally attractive" outcomes for the patients
- Need a dedicated team to oversee the study execution
- Need regular checks to ensure accurate, timely data capture

Selker, H.P., Cohen,T., D'Agostino, R.B.,et al. (2022), A Useful and Sustainable Role for N-of-1 Trials in the Healthcare Ecosystem. Clin Pharmacol Ther, 112: 224-232

Special Thanks to the Sleep(less) Team!

- Danielle Miller, MPH
- Krisia Romano, BA
- Jordyn Rodillas, BA
- Joan Duer-Hefele, MA
- Heejoon Ahn, MS
- Frank Vicari, BS
- Ciaran Friel, EdD
- Ashley Goodwin, PhD
- Karina Davidson, PhD

This work is supported by the National Institute on Aging (NIA) P30AG063786-01

Feinstein Institutes for Medical Research
Northwell Health®

Personalized N-of-1 Trials:
Analyzing Data from an Individual Level and in Pooled Level

Thevaa Chandereng, PhD

Institute of Health System Science

Feinstein Institutes for Medical Research

Northwell Health

Statistical Analysis: Individual Level

1. Unlike traditional RCT, treatment arms are not independent, outcomes are collected continuously throughout the trial
2. Statistical analyses need to account for correlation structure

In a standard linear model,

$$Y = X\beta + \varepsilon,$$

where Y is the *n* x 1 response vector, X is an *n* x *p* model matrix, β is a *p* x 1 vector of estimated regression coefficients, and ε is an *n* x 1 vector of errors. The ordinary-least square (OLS) estimator of β assuming that $\varepsilon \sim N(0, \sigma^2 I_n)$ (i.e. the errors are uncorrelated),

$$\beta_{OLS} = (X^T X)^{-1} X^T Y,$$

with the covariance matrix

$$Var(\beta_{OLS}) = \sigma^2 (X^T X)^{-1}.$$

Statistical Analysis: Individual Level

Generalized least-squares (GLS) regression extends ordinary least-squares (OLS) estimation of the standard linear model by providing for possibly unequal error variances and for correlations between different errors. Let V = Var(ε | X).

$$V = Var(\varepsilon) = \begin{bmatrix} \sigma_1^2 & \rho_{1,2}\sigma_1\sigma_2 & \cdots & \rho_{1,n}\sigma_1\sigma_n \\ \rho_{2,1}\sigma_1\sigma_2 & \sigma_2^2 & \cdots & \rho_{2,n}\sigma_2\sigma_n \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n,1}\sigma_n\sigma_1 & \rho_{n,2}\sigma_n\sigma_2 & \cdots & \sigma_n^2 \end{bmatrix}$$

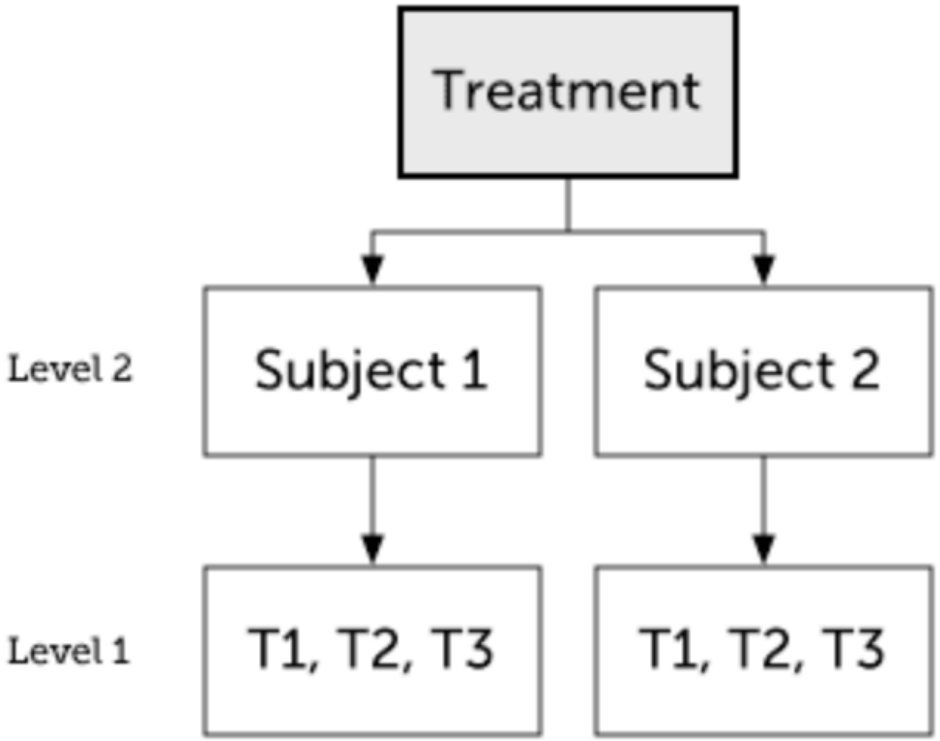
$$\beta_{GLS} = (X^T V^{-1} X)^{-1} X^T V^{-1} Y^*.$$

We used AR (1) model for the analyses. In an AR(1) model,

$$\sigma^2 = \sigma_1^2 = \cdots = \sigma_n^2.$$

$$\rho_{a,a+k} = \rho_{a+k,a} = \rho^k, \quad -1 \leq \rho \leq 1, \quad a, k > 0.$$

Statistical Analysis: Pooled Level



Two-level random slope model

$$Y_{ij} = \beta_{0j} + \beta_{1j} t_{ij} + R_{ij}$$

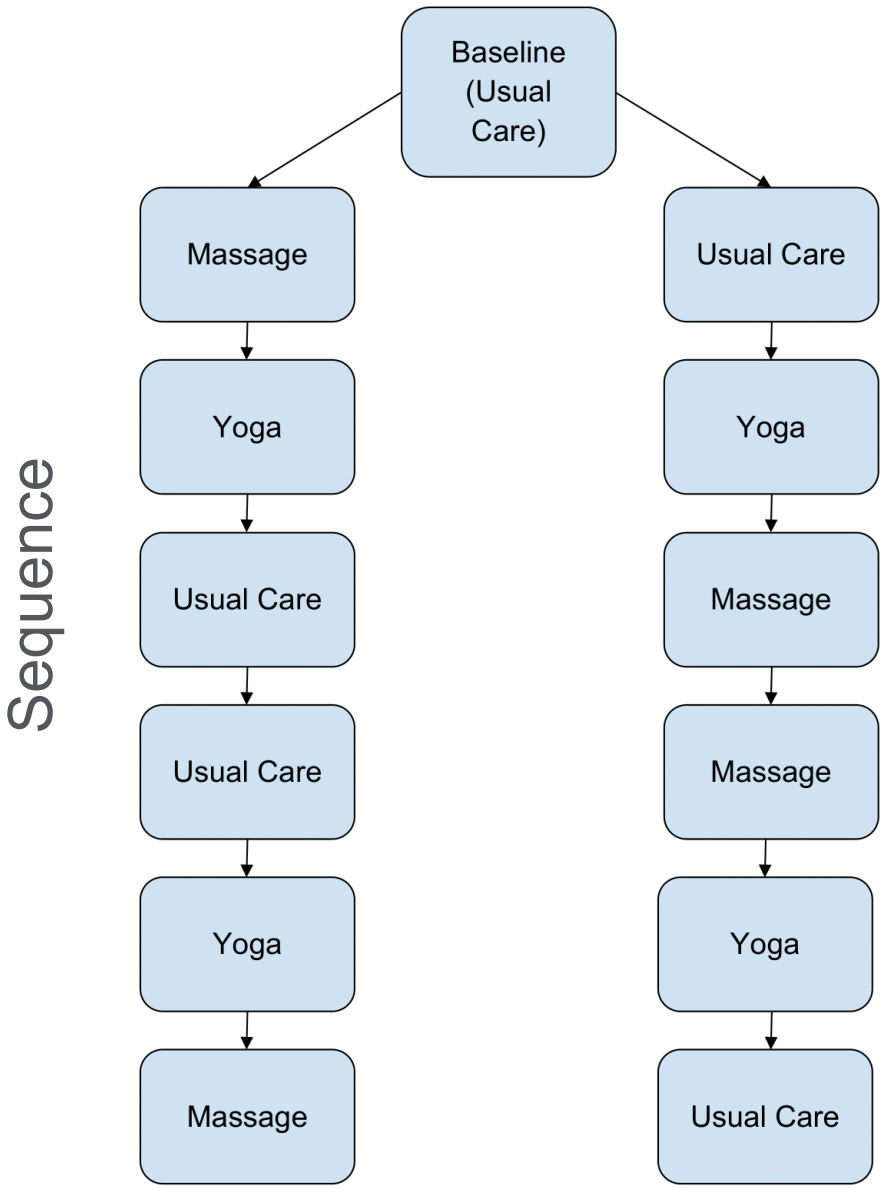
$$\beta_{0j} = \gamma_{00} + U_{0,j}$$

$$\beta_{1j} = \gamma_{10} + U_{1,j}$$

$$\begin{pmatrix} U_{0j} \\ U_{1j} \end{pmatrix} \sim \mathcal{N} \begin{pmatrix} 0, \tau^2_{0,0} & \tau_{0,1} \\ 0, \tau^2_{1,0} & \tau^2_{1,1} \end{pmatrix}$$

where Y is the *n* x 1 response vector, t is an *n* x *p* model matrix, β, γ are *p* x 1 vector of estimated regression coefficients, and r is an *n* x 1 vector of errors.

Personalized (N-of-1) Trial: Chronic Lower Back Pain (CLBP)



- Daily measurement of outcomes including:
  - PROMIS Pain Intensity
  - PROMIS Pain Interference
  - EMA 3 times daily of pain, fatigue, and stress (Average taken)
  - Self-reported use of over-the-counter pain medication
  - Self-reported treatment side effects
  - Fitbit Physical Activity (steps, sleep)

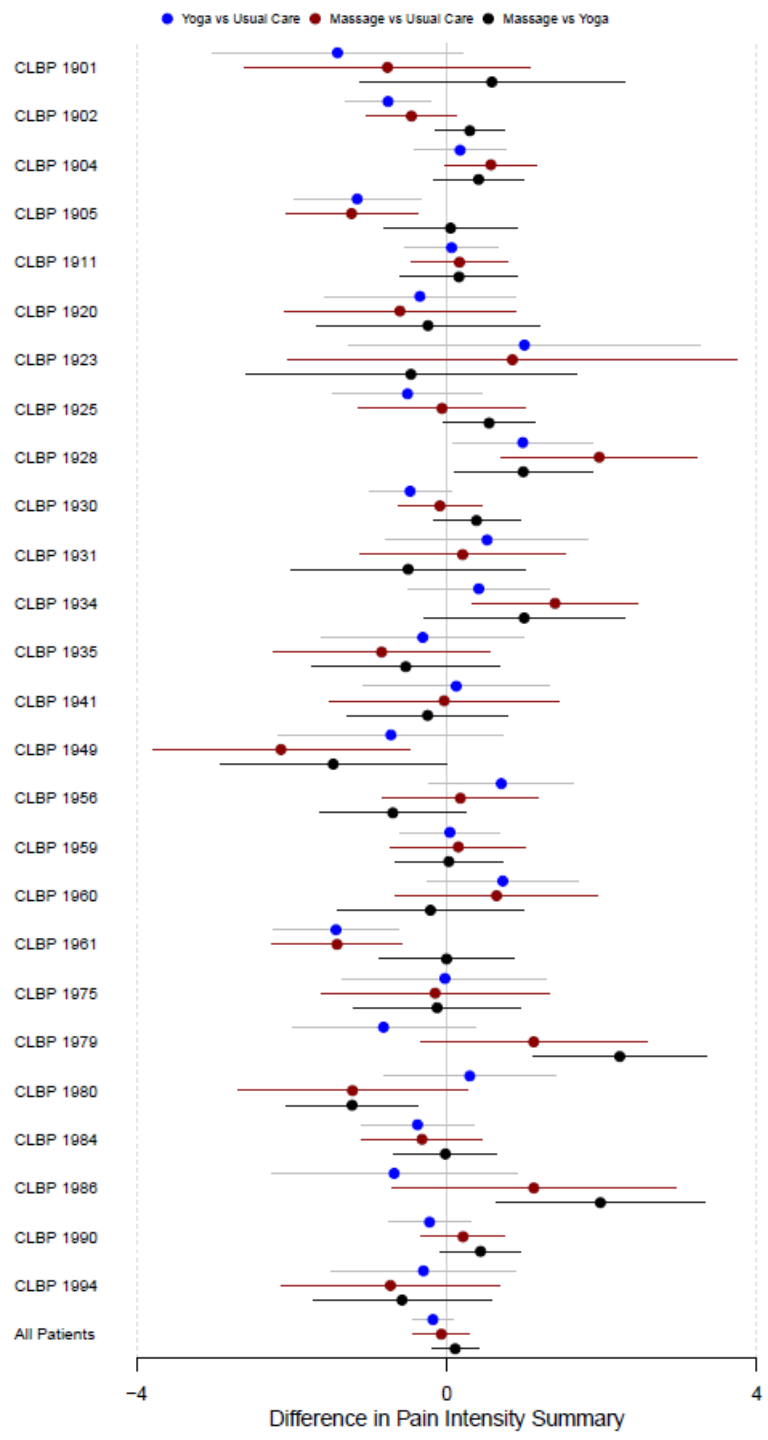
Trial Results: Pain Intensity Summary

Better

Worse

	Usual Care	Yoga	Massage
Usual Care	-	3	3
Yoga	1	-	1
Massage	2	3	-

Pain Intensity Summary

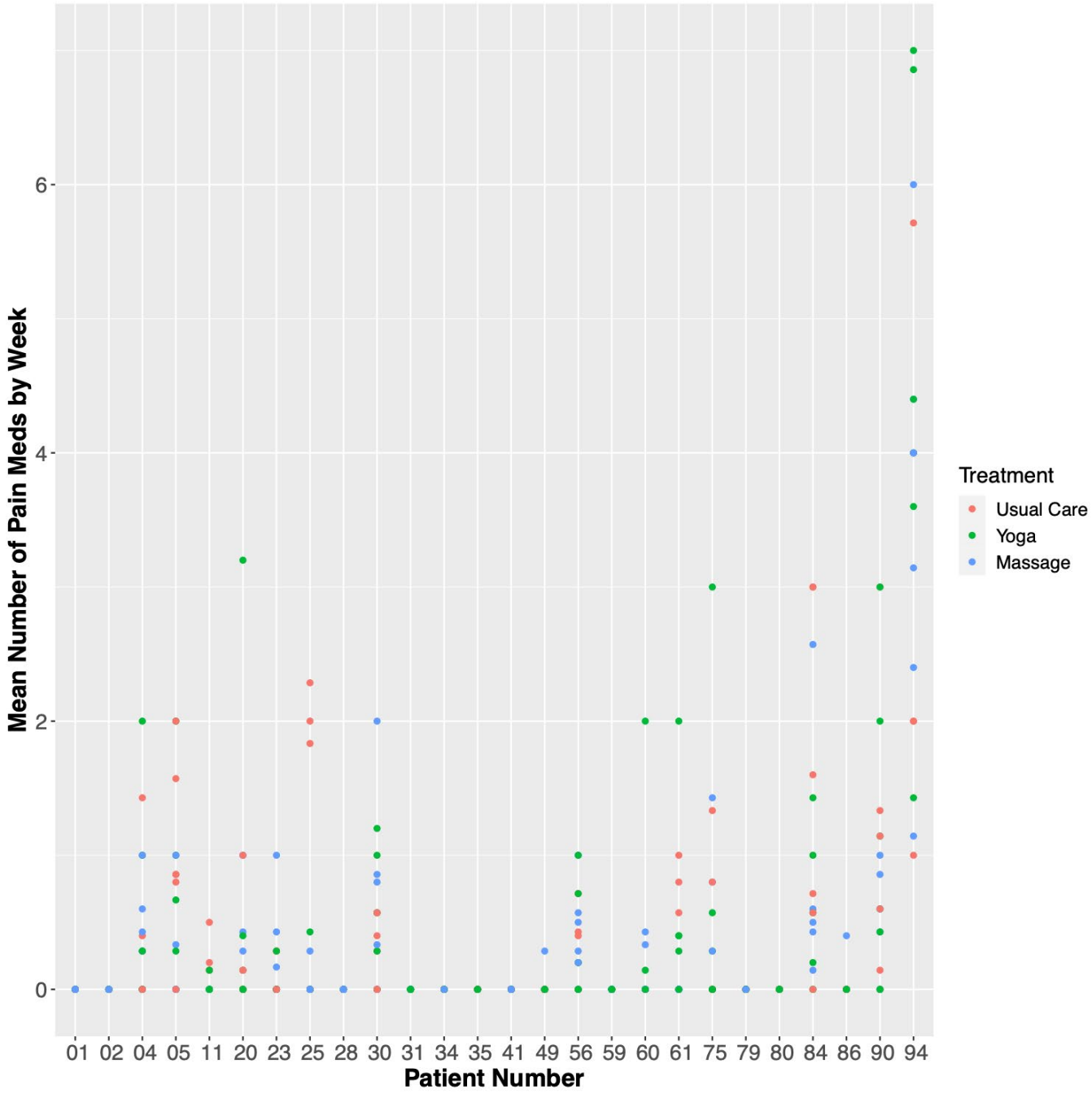


Ecological Momentary Assessments

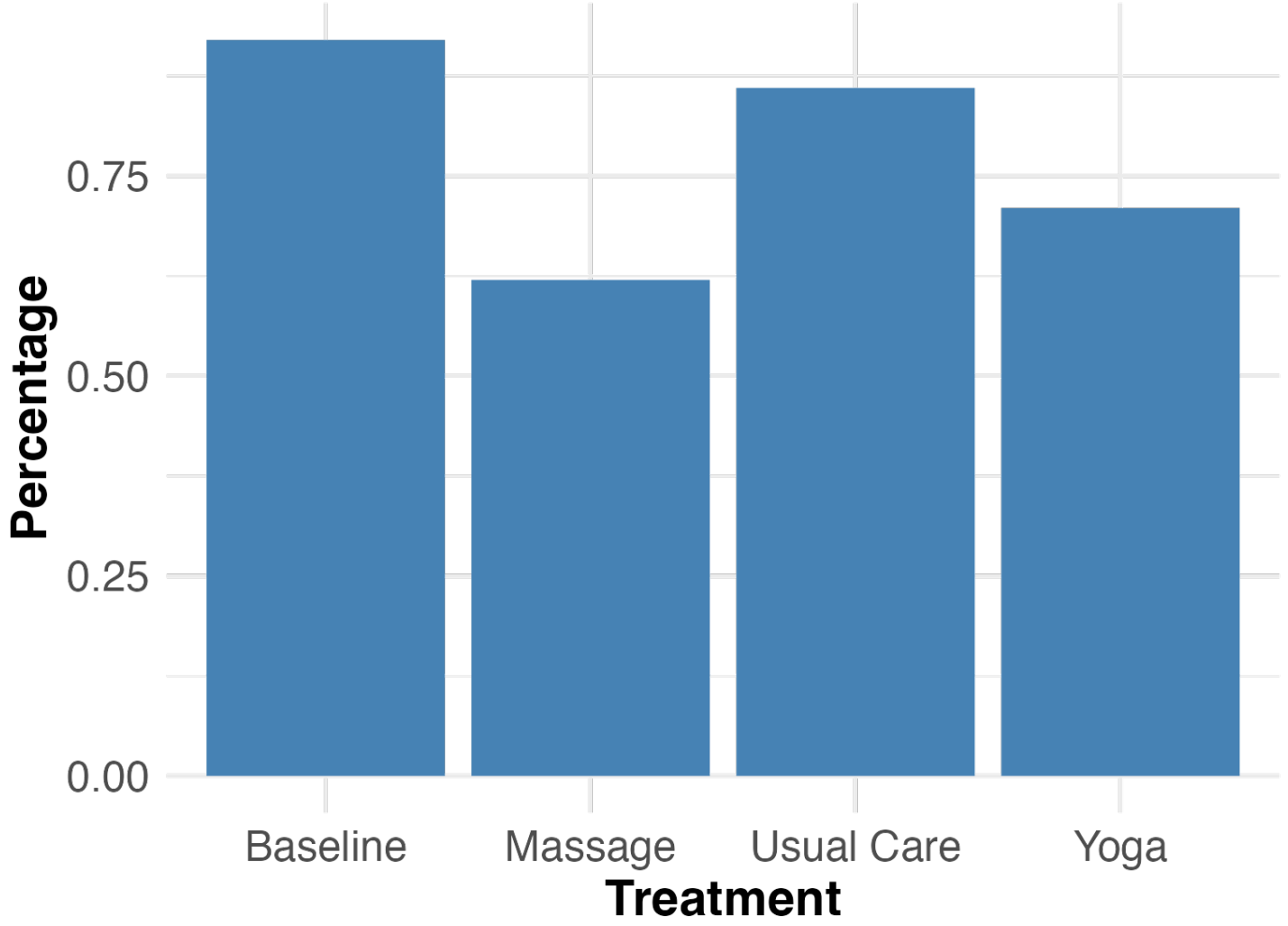
		Positive effect (lower EMA)	Negative effect (lower EMA)	No effect
EMA Pain	Yoga vs Usual Care	2	1	23
	Massage vs Usual Care	2	2	22
EMA Fatigue	Yoga vs Usual Care	3	1	22
	Massage vs Usual Care	2	3	21
EMA Stress	Yoga vs Usual Care	0	1	25
	Massage vs Usual Care	4	2	20

Number of participants with no, positive, and negative significant effects for all three daily average EMAs comparing both yoga and massage to usual care.

Trial Results



Pooled Percentage of Average Minutes Wearable Device (Fitbit) Worn During Each Treatment



WEARABLE (FITBIT) DATA

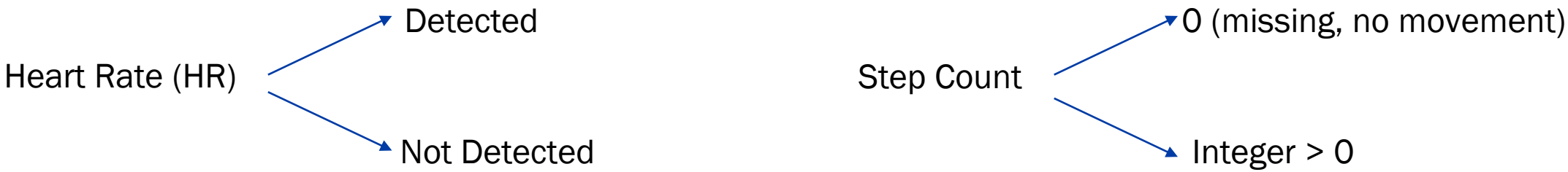
Data Source: Daily Step Count

Id	Activity Date	Date	Weekday	Repeat Measure	Total Steps
CLBP 1901	Day 1	11/11/19	Monday	No	14230
CLBP 1901	Day 2	11/12/19	Tuesday	No	13824
CLBP 1901	Day 3	11/13/19	Wednesday	No	11593
CLBP 1901	Day 4	11/14/19	Thursday	No	11485
CLBP 1901	Day 5	11/15/19	Friday	No	16001
CLBP 1901	Day 6	11/16/19	Saturday	No	9731
CLBP 1901	Day 7	11/17/19	Sunday	No	7147
CLBP 1901	Day 8	11/18/19	Monday	No	14060
CLBP 1901	Day 9	11/19/19	Tuesday	No	13207
CLBP 1901	Day 10	11/20/19	Wednesday	No	12221

WEARABLE (FITBIT) DATA

Data Source: Minute-by-Minute Step Count

Participant ID	Day	Steps	HR	Intensity
RBABCT000077	2/6/22 11:14	0		0
RBABCT000077	2/6/22 11:15	0	70	0
RBABCT000077	2/6/22 11:16	0	72	0
RBABCT000077	2/6/22 11:17	0	77	0
RBABCT000077	2/6/22 11:18	0	77	0
RBABCT000077	2/6/22 11:19	30	78	1
RBABCT000077	2/6/22 11:20	21	81	1
RBABCT000077	2/6/22 11:21	22	82	1
RBABCT000077	2/6/22 11:22	10	81	1



IMPUTATION

Daily level

- Typically 10 hours of daily active wear data required for a valid wear day

PROS	CONS
Many algorithms out there for imputation	Omitting useful information
Easier to deal	

Minute-by-minute level

PROS	CONS
Tedious	Not many algorithms published
Large scale computing but more accurate due to no data go unused	

IMPUTING STEP COUNT DATA (MINUTE-BY-MINUTE LEVEL)

Methods Proposed

Heart Rate	Wear /Non wear	Recorded Step Count	Real Step Count
Detected	Wear	Step Count Recorded	Recorded Step Count
Not Detected	Non-Wear	Step Count Recorded > 0	Recorded Step Count
Not Detected	Non-Wear	0	Impute

Impute

We assume Y_{ti} can be modeled using a parametric model,

$$Y_{ti} \sim Poisson(\lambda_{ti}),$$

$$log(\lambda_{ti}) = \alpha_t + \beta_t r_{ti} + \gamma_t z_{ti} + \epsilon_{ti}, \quad (1)$$

where λ_t is the Poisson rate at time t on day i, α_t, β_t, and γ_t are regression coefficients at time t, r_{ti} is the weekday indicator (1 - weekday/ 0 - weekend) at time t on day i, z_{ti} is the temperature (in Fahrenheit) at time t on day i, and ϵ_{ti} is the error term at time t on day i.

(Chandereng et al. 2022 HDSR)

Trial Results

N-of-1 Trial Fitbit Data

Choose a patient:

Patient 1

Use imputed steps data?

No

Document format

PDF HTML Word

Download Report

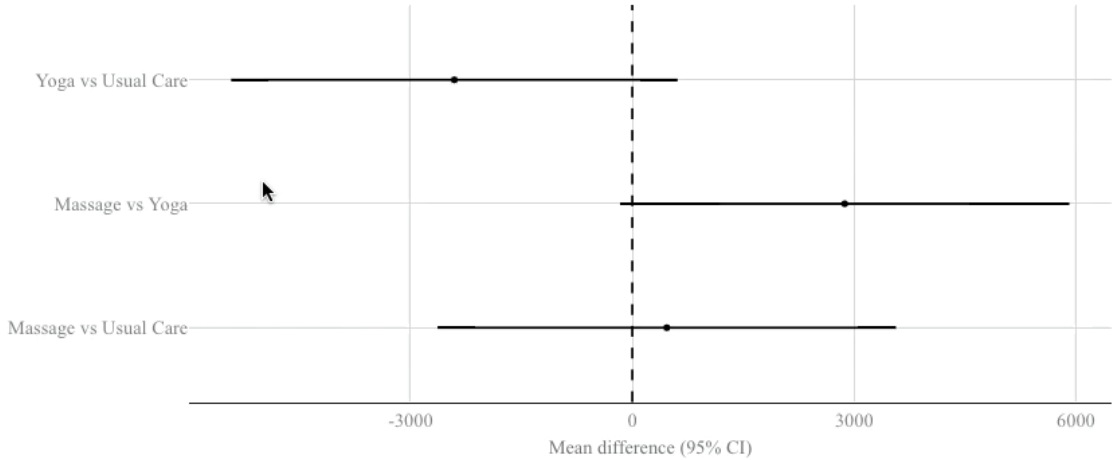
AR(1) model analysis with treatment effects

Copy CSV Excel PDF Print Search:

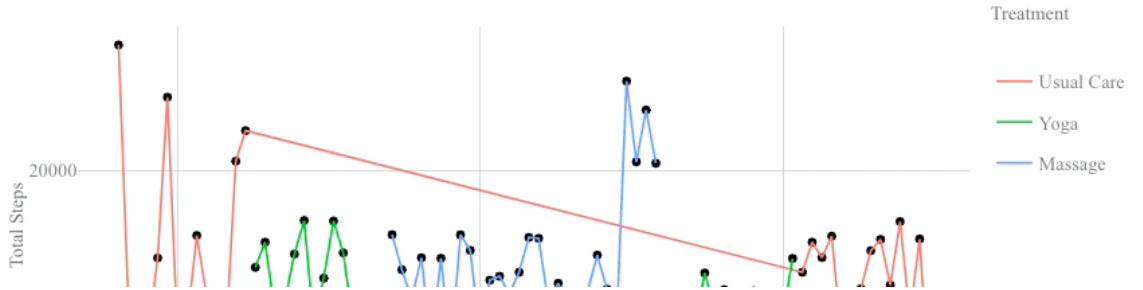
Output	Estimate	p.value
Total steps averaged during Usual Care	13480.68	-
Total steps averaged increased during Yoga	-2406.38	0.12
Total steps averaged increased during Massage	467.21	0.77

Showing 1 to 3 of 3 entries Previous 1 Next

Forest plot of treatment comparison

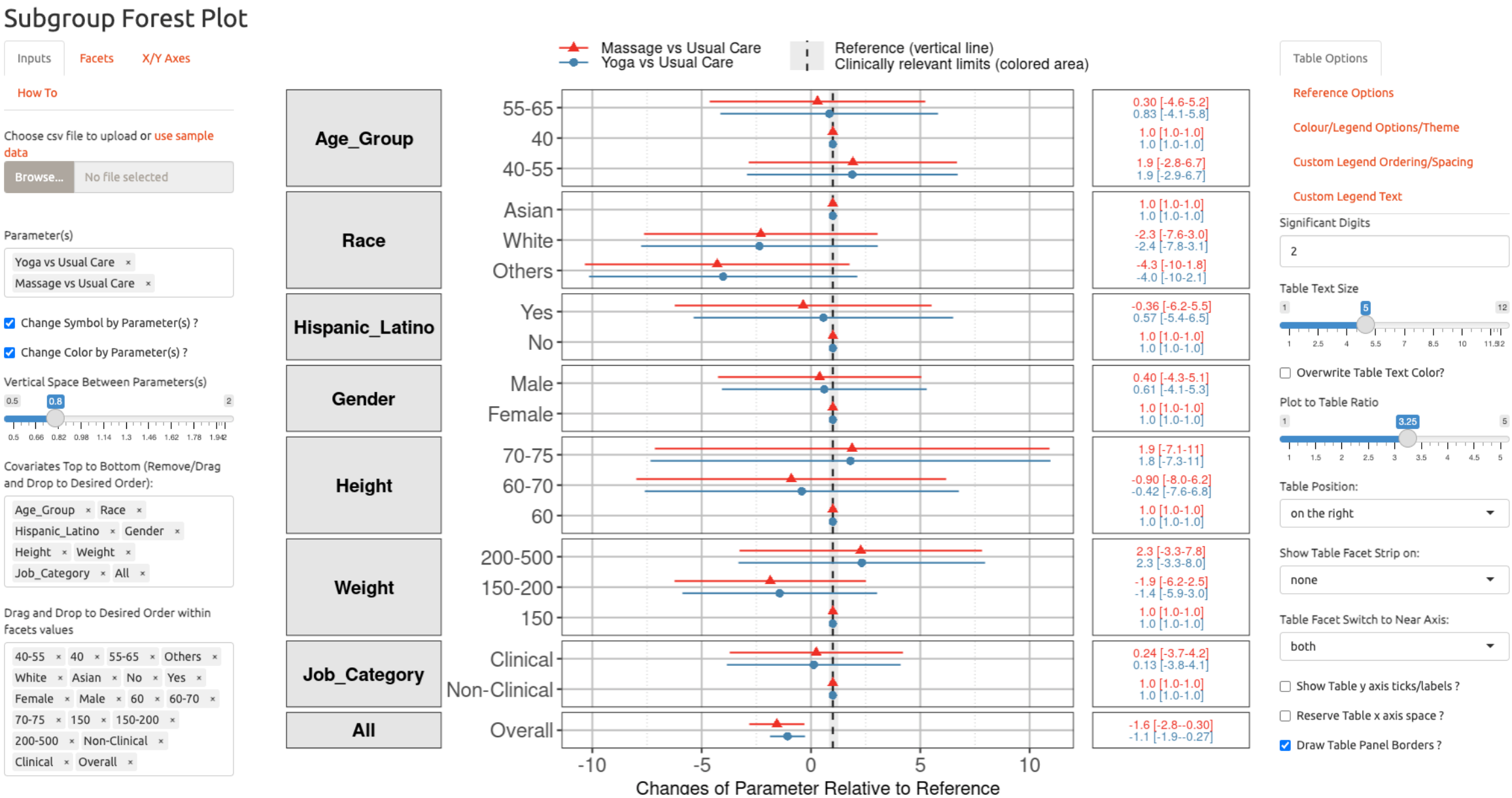


Number of steps per day



Trial Results: Subgroup Analysis

Shiny app: https://thevaa.shinyapps.io/subgroup/



Challenges

- Analyzing data pre- and post-pandemic data
- Summarizing patient reports with multiple outcomes
- Improving imputation algorithm for step counts

Novel Bayesian Model for N-of-1 Trials: Addressing Carryover Effects in Estimating Treatment Effect on Individual Responses

Ziwei Liao

Columbia University, Department of Biostatistics

May 13, 2024

Overview

Overview

Background

Traditional parallel-group randomized controlled trials (RCTs) haven long been the most prevalent study design to identify the population-level treatment effect in evidence-based medicine.

Some critiques on the current paradigm for producing clinical knowledge:

1. Practically impossible to conduct RCTs to address all clinically important questions.
2. Clinical evidence generated in those RCTs has poor generalizability.
3. Treatments shown to be safe and effective on average may deliver an uneven mix of risks and benefits to individual patients. (top ten highest-grossing drugs will only benefit between 1/25 and 1/4 of the patients who used them)

Background

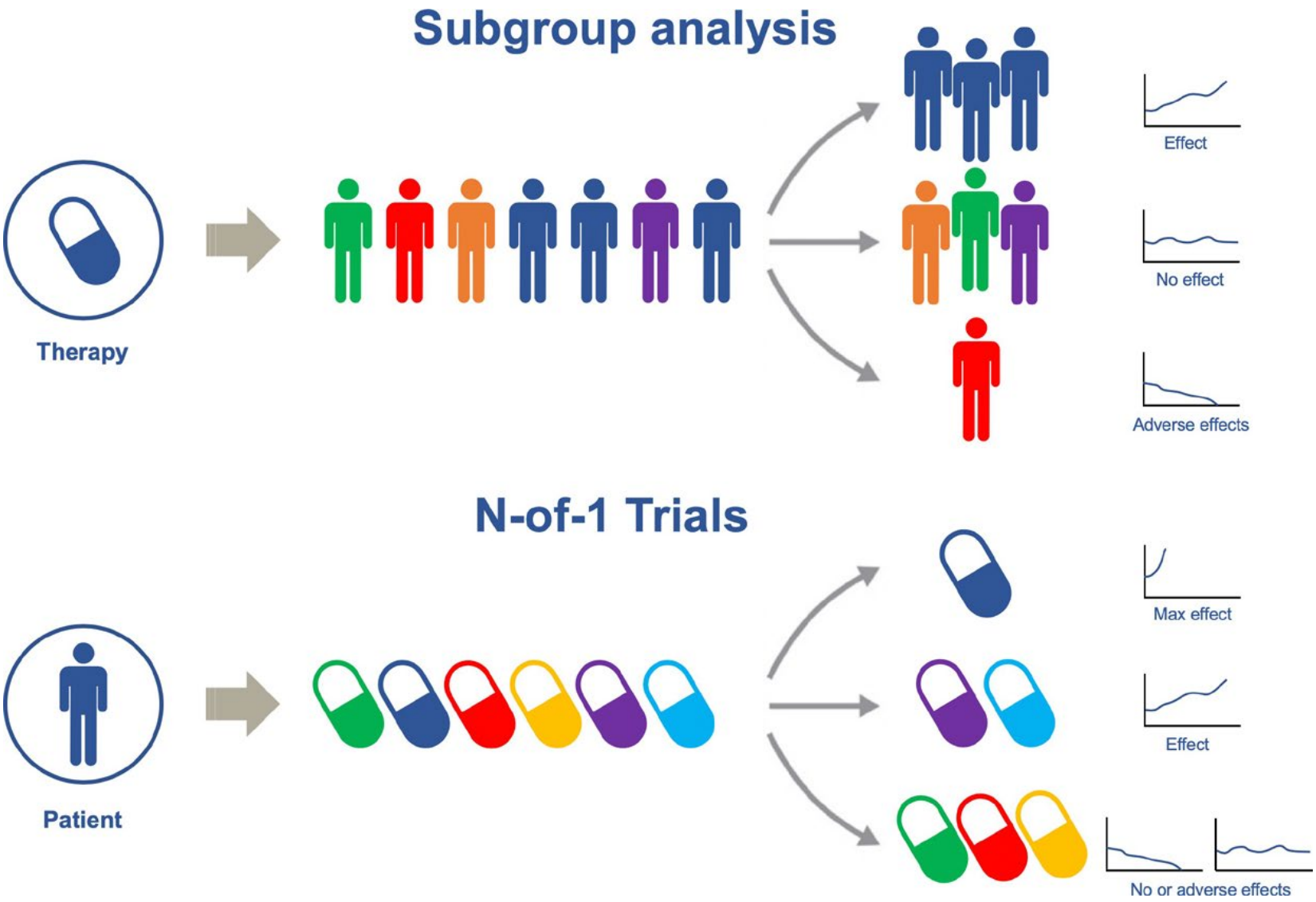
Heterogeneity of treatment effect (HTE):

- Nonrandom, explainable variability in the direction and magnitude of treatment effects for individuals within a population.
- Variance of ITEs across patients

Individual treatment effect (ITE):

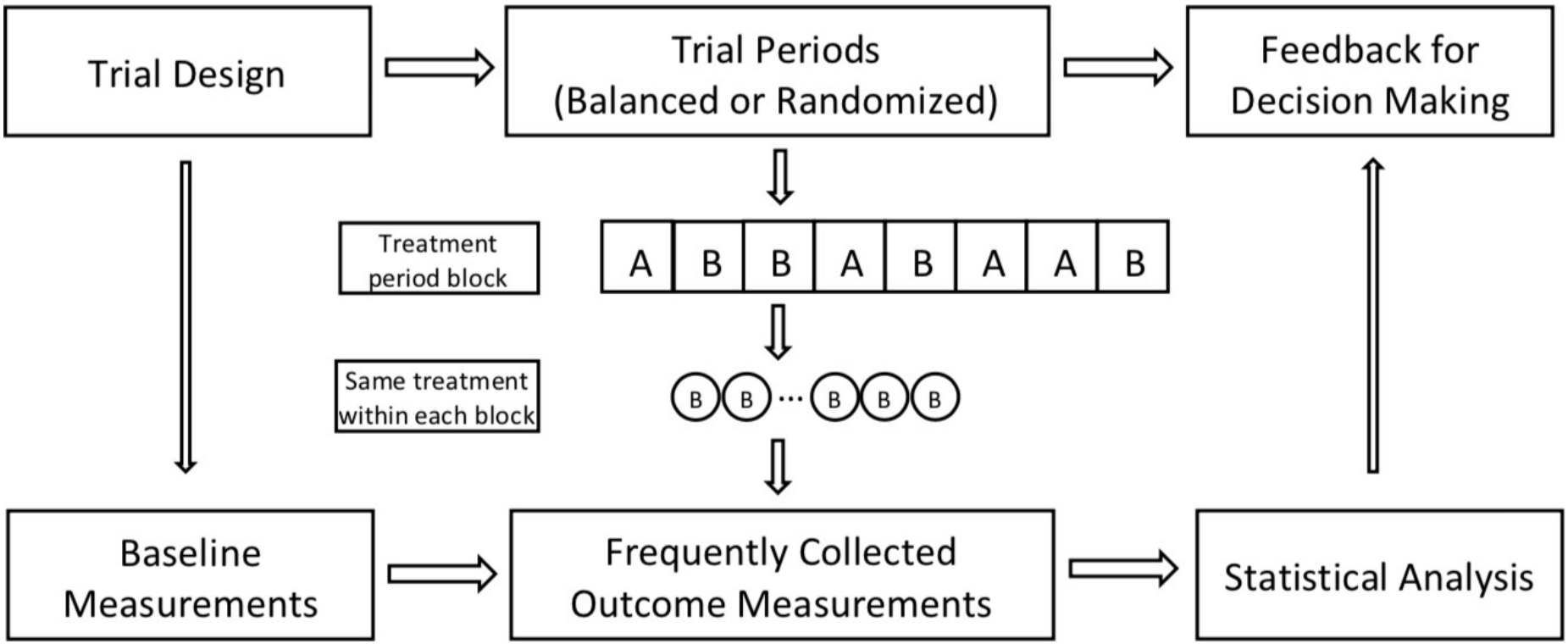
- The difference in effects (net benefits) between treatment A and treatment B for an individual patient

Background


Subgroup analysis
Therapy
Effect
No effect
Adverse effects
N-of-1 Trials
Patient
Max effect
Effect
No or adverse effects

Background : N-of-1 trials

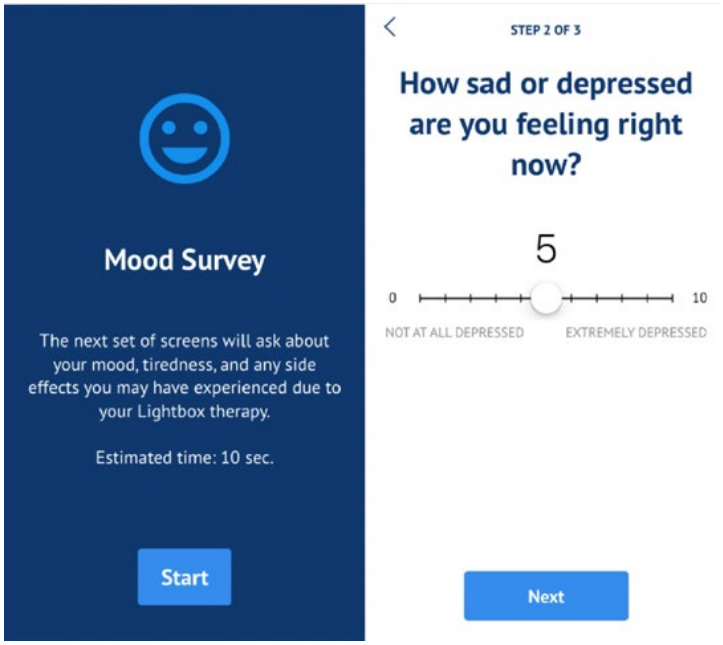
- N-of-1 trials are prospectively multiple-period crossover trials compar- ing two or more interventions within single participant.



Background : N-of-1 trials

- Application:
  - Chronic or slowly progressive conditions, e.g. diabetes, hypertension
  - Rare diseases
- Advantage:
  - Estimate ITE directly
  - Participants will immediately benefit from the study
  - Can be combined across patients to provide an estimate of either average treatment effect or of heterogeneity of treatment effects (Bayesian hierarchical model / meta analysis)
- Concern:
  - Costs and time involved in both patient and physician

Background : Light therapy study



- The data set is from Kronish et al., which studies the effectiveness of bright white light therapy for depressive symptoms within 9 cancer survivors.
  - Interventions: Bright white (10,000 lux), Dim red (50 lux)
  - Design: Single balanced sequence (ABBA or BAAB), no washout periods
  - Duration: 3 weeks' treatment block * 4 = 12 weeks
  - Outcome: Depressive score: 0-not at all depressed to 10-extremely depressed

Background : Light therapy study

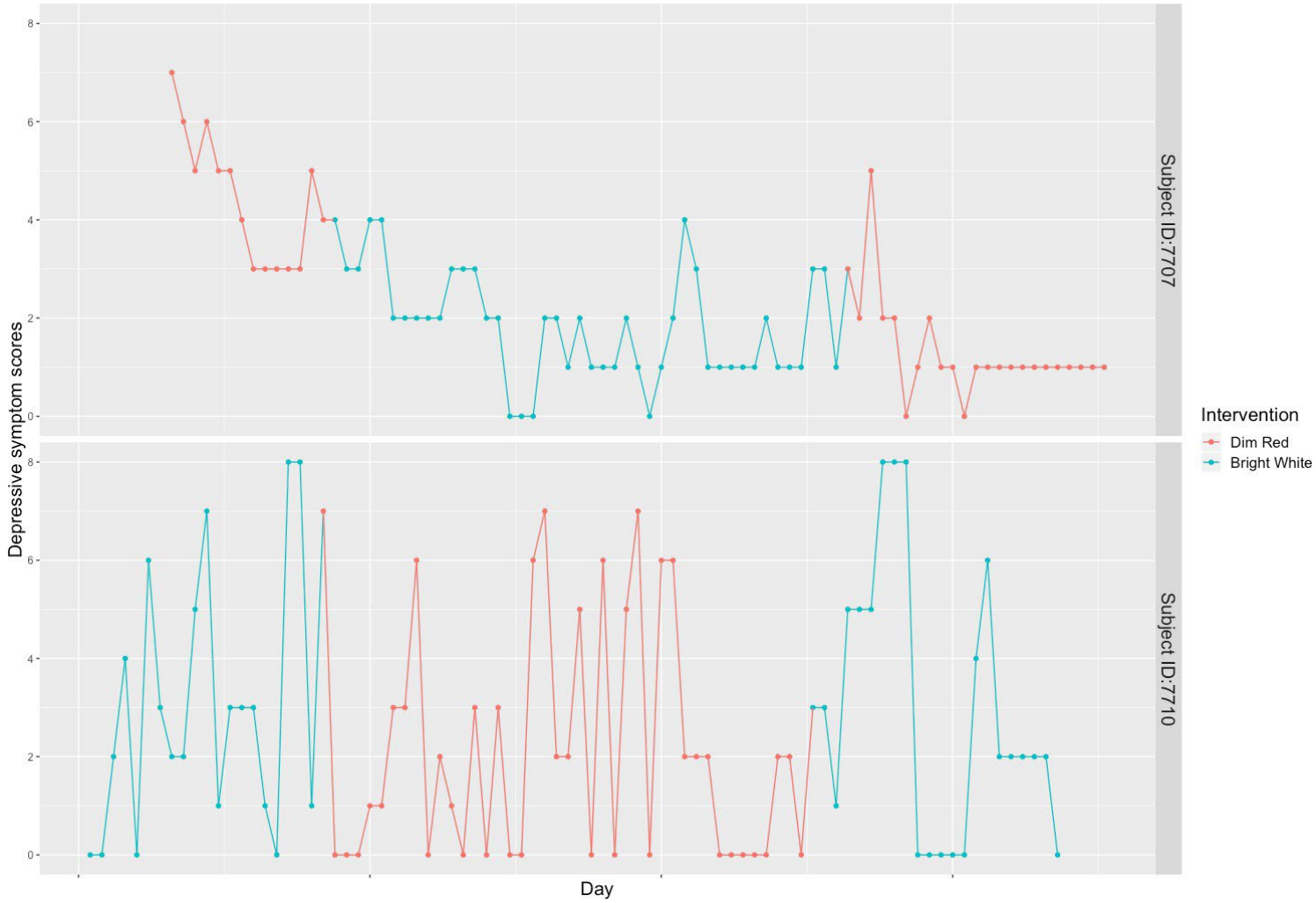


Figure: *Time series plot of two subjects. Blue curve represents bright white light intervention and red curve represents dim red light intervention.*

Overview

Current methods to analyze N-of-1 trial

1. **Nonparametric tests**
   - Sign test, Wilcoxcon signed rank test, et al.
   - Data gathered to treatment block/period level; Binary response for each block
   - Omit the size of difference in treatment effect

2. **ANOVA type tests**
   - Also use treatment block level information
   - Assume treatment blocks are mutually independent and outcomes are iid normal

Current methods to analyze N-of-1 trial (cont.)

3. **Regression based tests**
   - Use treatment block level information
   - Combination of several subject's data
     - Multiple N-of-1 trial can be regarded as grouped crossover design with fixed treatment sequence
     - Use linear/generalized mixed effect models
     - Assume normal distributed outcome

$$Y_{ij} = \beta_0 + \beta_1 X_{ij} + \beta_2 P_j + \lambda_A Z_{Aij} + \lambda_B Z_{Bij} + b_{0i} + b_{1i} X_{ij} + \varepsilon_{ij}$$

   - For single subject, model will be reduced to

$$Y_j = \beta_0 + \beta_1 X_j + \beta_2 P_j + \lambda_A Z_{Aj} + \lambda_B Z_{Bj} + \varepsilon_j$$

Important features of N-of-1 trial

- **Autocorrelation**
  - Time-series exhibit some form of serial dependence (adjacent time points)
  - Lead to underestimated/overestimated standard errors, increase the risk of Type I/II error
- **Carryover effects**
  - Washout periods does not guarantee a elimination of carryover effect
  - Lead to bias in the estimated treatment effects
- **Goal**: compare two different interventions
  - Estimate total effectiveness in treating the targeted outcome
  - Adjust for potential carryover effects (if no washout period or insufficient washout), and run-in period effects (slow onset of new treatment)
  - Account for autocorrelation in time series measurements

Overview

Distributed lag models (DLM)

- DLMs are used to model the current values of a dependent variable based on both the current values of an explanatory variable and the lagged values of this explanatory variable for time series data.

  The general form:

$$Y_t = \mu + \sum_{i=0}^{\infty} \beta_i x_{t-i} + u_t \tag{1}$$

  where Y_t denote a response variable of interest at time t and X_t denote a time-varying explanatory variable, which has some influence on all outcome after Y_t. u_t are $IN(0, \sigma^2)_u$, Lag length is usually truncated to some finite lag L.

- **Goal**: estimate (1). Distributed lag coefficients β_i, $(i = 1, 2, \ldots L)$; (2). Total effect $\sum \beta_i$

Distributed lag models (DLM)

- **Unconstrained DLM**
  - Estimate through unconstrained MLE
  - (1) Multicollinearity between of lagged explanatory variables
  - (2) Degree of freedom can be depleted quickly

- **Polynomial DLM**
  - Almon (1965) proposed to impose smoothness on the coefficients by restricting the lag coefficients to lie on a polynomial function
  - Coefficients β_i can be approximated by a d-th degree polynomial function of i, where d is usually much smaller than L, i.e.,

$$\beta_i = \sum_{j=0}^{d} \alpha_j i^j, \quad i = 0, \ldots L \text{ and } 0 < d \leq L \tag{2}$$

  - OLS estimator

Distributed lag models (DLM)

- **Geometric DLM**
  - Koyck (1954) proposed an infinite distributed lag model by adding the following constrain on lag coefficients in (1):

$$\beta_i = \beta_0 \lambda^i, \quad \text{and} \quad 0 < \lambda < 1 \tag{3}$$

where λ is the coefficient decaying rate, then we will have

$$Y_t = \mu + \beta_0 \sum_{i=0}^{\infty} \lambda^i x_{t-i} + u_t \tag{4}$$

  - Can be transformed from an infinite DLM into an autoregressive model with only three observable variables by subtracting λy_{t-1} from (4):

$$Y_t = (1-\lambda)\mu + \beta_0 X_t + \lambda y_{t-1} + (u_t - \lambda u_{t-1}) \tag{5}$$

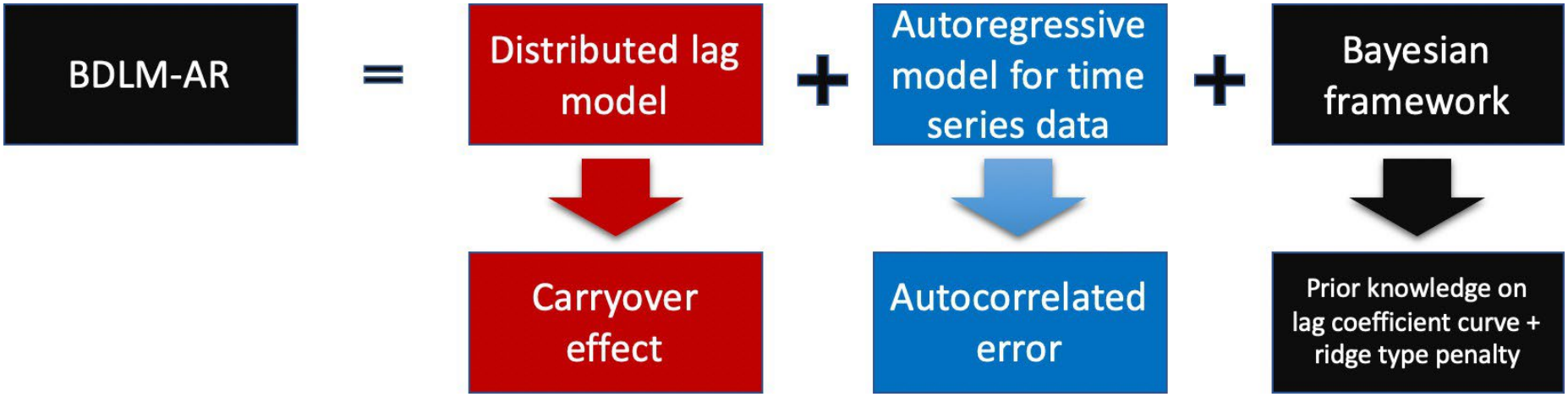
  - MLE of β_0 and λ can be obtained through iterative numeric methods.

Overview

Bayesian distributed lag model with autocorrelated errors (BDLM-AR)

Proposed model:

$$
\begin{aligned}
Y_t &= \mu + Z_t' b + \sum_{l=0}^{L} \beta_l X_{t-l} + \epsilon_t \\
\epsilon_t &= \phi_1 \epsilon_{t-1} + \phi_2 \epsilon_{t-2} + \ldots + \phi_p \epsilon_{t-p} + w_t \\
w_t &\sim N(0, \sigma^2)
\end{aligned}
\tag{6}
$$



BDLM-AR

- **Estimands of interest**:
  1. Total effect $\sum_{l=0}^{L} \beta_l$;
  2. Total carryover effect $\sum_{l=1}^{L} \beta_l$;
  3. Lag coefficients β_l, $(l = 0, 2, \ldots L)$.
     - β_0: Immediate treatment effect (treatment effect on day 0)
     - β_l, $(l/= 0)$: Carryover effect due to treatment on l days ago

- Data transformation:
  Let $\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p$, where B is the backshift operator, then the AR model for the errors is:

$$\Phi(B)\epsilon_t = w_t$$

Let $Y_t^* = \Phi(B)Y_t$, $X_t^* = \Phi(B)X_t$, $\mu^* = \Phi(B)\mu$, $Z_t^* = \Phi(B)Z_t$, $t = p+1, \ldots, n$. Then the model with transformed data is

$$(Y^* \mid X^*, \mu^*, \boldsymbol{\beta}) \sim \mathrm{N}(\mu^* \mathbf{1}_{n-p} + Z^* b + X^* \boldsymbol{\beta}, \sigma^2 \mathbf{I}_{n-p}) \qquad (7)$$

BDLM-AR : Prior distribution

1. Prior on distributed lag coefficient

- The prior on $\boldsymbol{\beta}$ is designed to be

$$(\boldsymbol{\beta}|\sigma^2, \gamma_1, \gamma_2) \sim N(\mathbf{0}, \sigma^2 \boldsymbol{\Omega}^{-1}(\gamma_1, \gamma_2)) \tag{8}$$

- Non-informative prior is used for μ and b

$$(\mu|\sigma^2) \sim N(0, \sigma^2 c_0^{-1}), \quad (b|\sigma^2) \sim N(0, \sigma^2 c_0^{-1} I) \tag{9}$$

BDLM-AR : Prior distribution (cont.)

- where $\mathbf{\Omega}(\gamma_1, \gamma_2)$ is the precision matrix:

$$\begin{bmatrix}
\lambda_0 + \lambda_0^* & -\lambda_0^* & 0 & \dots & \dots & 0 \\
-\lambda_0^* & \lambda_1 + \lambda_0^* + \lambda_1^* & -\lambda_1^* & \dots & & 0 \\
0 & -\lambda_1^* & \lambda_2 + \lambda_1^* + \lambda_2^* & \dots & & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & 0 \\
0 & 0 & 0 & -\lambda_{L-2}^* & \lambda_{L-2} + \lambda_{L-2}^* + \lambda_{L-1}^* & -\lambda_{L-1}^* \\
0 & 0 & 0 & \dots & -\lambda_{L-1}^* & \lambda_L + \lambda_{L-1}^* + \lambda_L^*
\end{bmatrix}$$

$\lambda_l = \exp[\gamma_1(l+1)] - 1, \lambda_l^* = \exp[\gamma_2(l+1)] - 1$ for $l = 0, 1, 2, \dots L$.

- γ_1: controls the rate at which the variance of β_l tapers to 0.
- γ_2: controls the increasing rate of smoothness of the curve.

Relationship with regularization methods

- The maximum *a posteriori* probability estimate of $\boldsymbol{\beta}$ minimizes a fused ridge-type penalty:

$$(\boldsymbol{Y}^* - \boldsymbol{X}^*\boldsymbol{\beta})^T(\boldsymbol{Y}^* - \boldsymbol{X}^*\boldsymbol{\beta}) + \alpha_0\mu^2 + \sum_{l=0}^{L} \lambda_l\beta_l^2 + \sum_{l=0}^{L} \lambda_l^*(\beta_l - \beta_{l+1})^2$$

- Regularize
  1. The ℓ_2-norm of the coefficients (ridge penalty);
  2. Successive differences of coefficients (local smoothness penalty);
  3. Monotone increasing penalty (Strong shrinkage effect to 0 at large lag).

BDLM-AR : Prior distribution (cont.)

1. Prior distribution on the mean model

- $(\boldsymbol{\beta}|\sigma^2, \gamma_1, \gamma_2) \sim N(\mathbf{0}, \sigma^2 \boldsymbol{\Omega}^{-1}(\gamma_1, \gamma_2))$
- $(\mu|\sigma^2) \sim N(0, \sigma^2 \epsilon_0^{-1})$
- $\pi(\gamma_1, \gamma_2) \propto \exp(-\gamma_1 - \gamma_2)$

2. Prior distribution on the error model

- $\pi(\sigma^2) \propto 1/\sigma^2$
- $\boldsymbol{\phi} \sim N_p(0_p, 200 \times \mathbf{I}_p) \mathbf{1}_{S_{\boldsymbol{\phi}}}(\boldsymbol{\phi})$

Proposed hybrid Metropolis-Hastings/Gibbs algorithm

Step 1. Set initial values for $\boldsymbol{\beta}$, σ^2, $\boldsymbol{\phi}$ and $\boldsymbol{\gamma}$;
for $i \leftarrow 1$ *to* $n_{iteration}$ **do**

Step 2. Given current value of $\boldsymbol{\phi}$, transform $\mathcal{Y}$, $\mathcal{X}$ to $\mathcal{Y}^*$, $\mathcal{X}^*$; Also construct precision matrix $\boldsymbol{\Omega}(\boldsymbol{\gamma})$ based on $\boldsymbol{\gamma}$;

Step 3. Update $\boldsymbol{\beta}$ based on $\pi(\boldsymbol{\beta} \mid \mathcal{Y}^*, \mathcal{X}^*, \sigma^2, \boldsymbol{\phi}, \boldsymbol{\gamma})$;

Step 4. Update σ^2 based on $\pi(\sigma^2 \mid \mathcal{Y}^*, \mathcal{X}^*, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\gamma})$;

Step 5. Update $\boldsymbol{\epsilon}$ conditional on current value of $\boldsymbol{\beta}$ and $\mathcal{Y}$, $\mathcal{X}$. Then update $\boldsymbol{\phi}$ based on $\pi(\boldsymbol{\phi} \mid \mathcal{Y}, \mathcal{X}, \boldsymbol{\beta}, \sigma^2, \boldsymbol{\gamma})$;

Step 6. Update (γ_1, γ_2) based on $\pi(\boldsymbol{\gamma} \mid \boldsymbol{\beta}, \sigma^2)$. Sample a proposal γ_i^* by $\gamma_i^* = \gamma_i + a * U(-1, 1)$ for $i = 1, 2$. a is an adjustable step size. Compute ratio

$$R_\gamma = \frac{\pi(\boldsymbol{\gamma}^* \mid \boldsymbol{\beta}, \sigma^2)}{\pi(\boldsymbol{\gamma} \mid \boldsymbol{\beta}, \sigma^2)}$$

update $\boldsymbol{\gamma} = \boldsymbol{\gamma}^*$ with probability $\min(1, R_\gamma)$;

end

Overview

Simulation settings

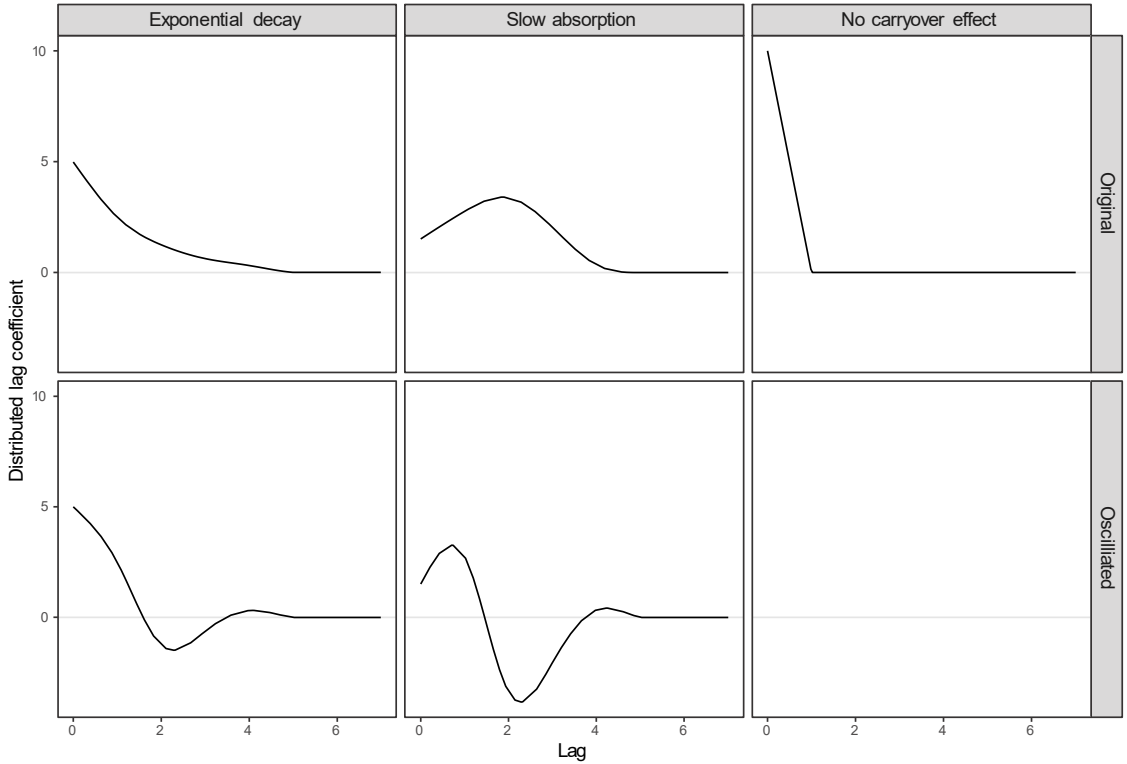
- Settings:

$$Y_t = \mu + \sum_{s=0}^{L} \beta_s X_{t-s} + \epsilon_t$$

  with $\epsilon_t = \phi_1 \epsilon_{t-1} + w_t$, and $w_t \sim N(0, \sigma^2)$.
- Total observations $N = 120$ with two treatments (A and B)
- $\sigma^2 = 100$; order of AR error is
- 1 Number of simulations: 100
- Other simulation parameters:
  1. Two block paired design : ABBA (sequence 1) and ABBABAAB (sequence 2);
  2. Two model standard deviation $\sigma = 10, 20$, then signal to noise ratio (SNR $:= \sum_{i=0}^{7} |\beta_i|/\sigma$) : SNR = 1 and 0.5;
  3. Autocorrelation between observations: $\phi_1 = 0.5$ and 0.2.

Simulation settings

- Five different coefficient curves of $\boldsymbol{\beta}$:



- Total effect = 10;
- Total carryover effects (δ) are 4.69, 0.94, 8.48, -2.30 and 0 respectively.

Simulation models summary

- One frequentist model (Koyck model) and four Bayesian models were used in simulation comparison

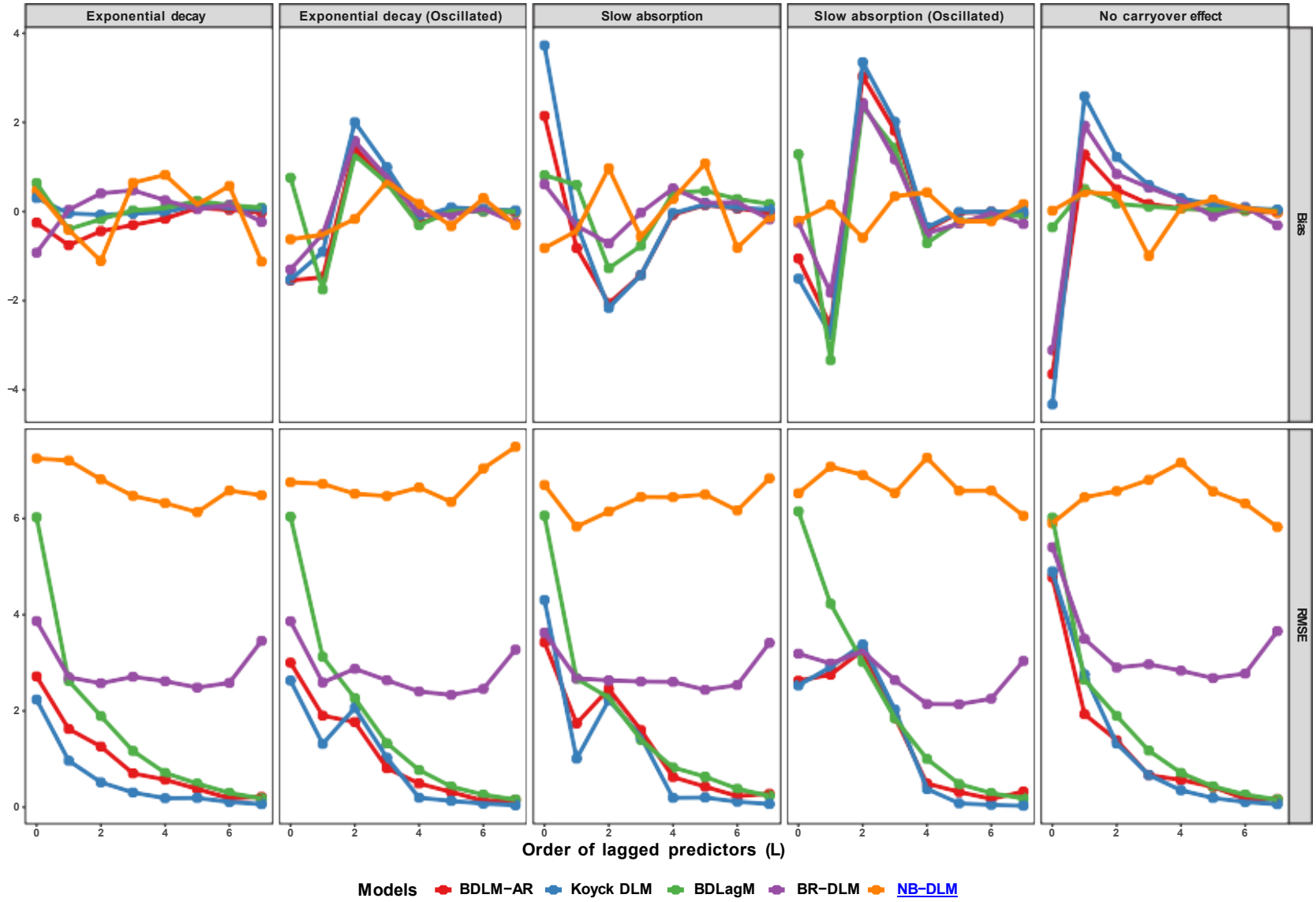
Method	BDLM-AR	Koyck DLM	BDLagM	Bayesian ridge DLM (BR-DLM)	Non-informative prior Bayesian DLM (NB-DLM)
Prior/constrain on lag coefficients	Normal with mean zero and designed prior on covarinace matrix, with increasing penalty on L_2 norm of coefficients and smoothness of adjancent coefficients	Lag coefficients are assumed to decrease geometrically	Normal with mean zero and designed prior on variance matrix, with decreasing variance of coefficients and increasing correlation between adjancent coefficients	Normal with mean zero and identity correlation matrix	Diffuse normal prior
Ridge penalty	Yes	-	No	Yes	No
Smoothness penalty	Yes	-	Yes	No	No
Prior knowledge on lag coefficient curve	Yes	-	Yes	No	No
Autoregressive coefficients	Yes	Predetermined	No	No	No

- Evaluation metrics for single coefficient: coefficients bias, coefficients root mean square error (RMSE).
- Evaluation metrics for whole coefficient curve:

$$\text{Euclidean Distance} = ||\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}||_2$$

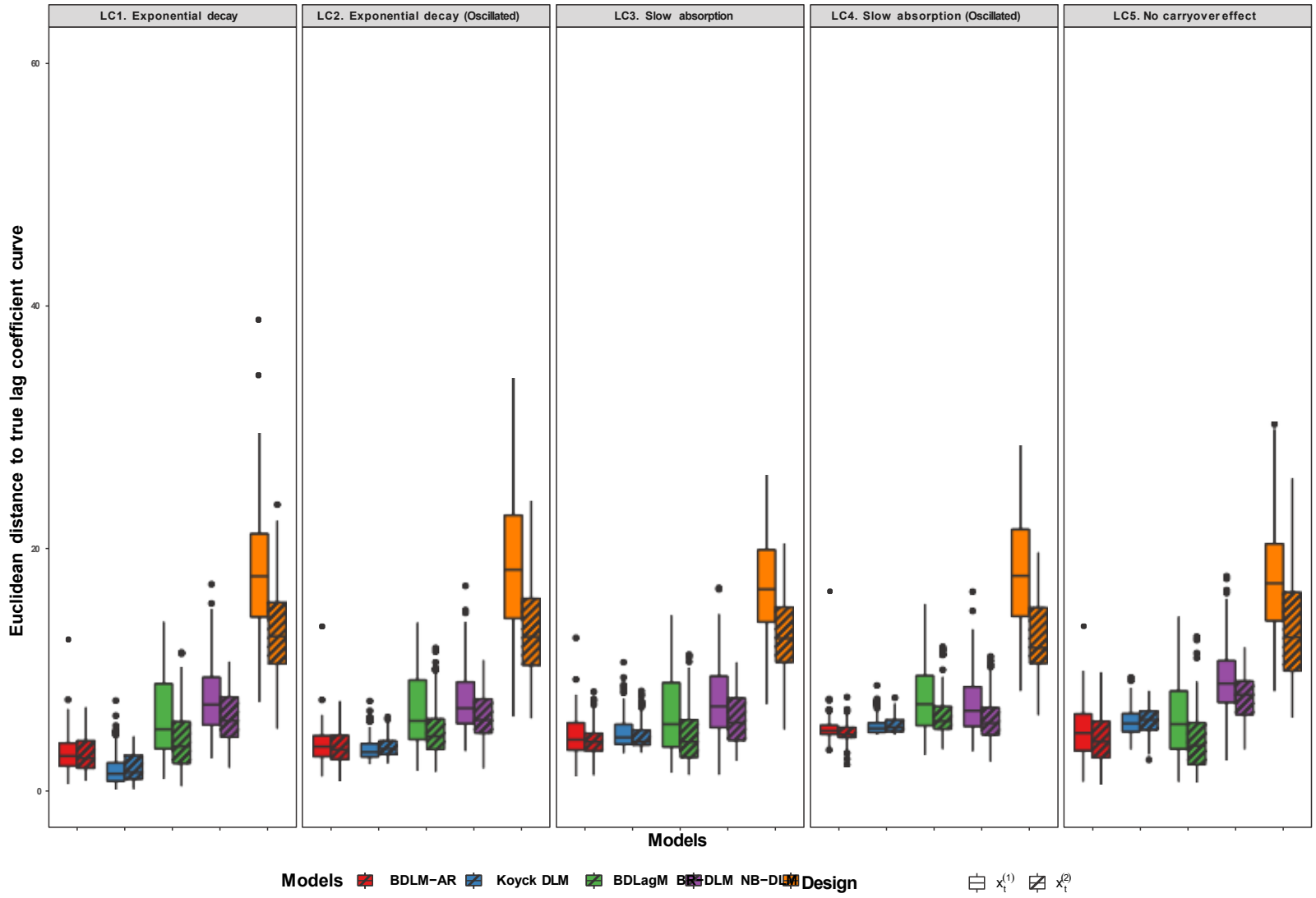
Simulation results

- Bias and RMSE of estimated lag coefficients



Distance under strength of autocorrelation

- Boxplot of distance to true lag coefficient curves



Overview

Application to light therapy study

- Use two subjects' data (ID:7706 and 7708), received treatment sequence BAAB and ABBA respectively (A:bright white lightbox; B:dim red lightbox)
- Each treatment block lasts 21 days
- Fitted with:
  1. Classical linear regression model with autoregressive errror (RegAR), with fixed AR order 1 and 7
  2. Proposed BDLM-AR model with AR(1) and AR(7) error
- Maximum number of lags L in DLMs is set to 7
- Ljung-Box test is used to check the fitness of each model

Application to light therapy study: Result

- MLE and posterior mean estimates of distributed lag coefficients and autoregressive coefficients

	Subject ID: 7708					
	BDLM-AR(1)	BDLM-AR(7)	BDLM-AR(1) with weekend effect	BDLM-AR(7) with weekend effect	RegAR(1)	RegAR(7)
μ	2.90 (1.67,4.12)	3.19 (0.31,6.88)	2.89 (1.51,4.22)	3.23 (0.16,7.25)	2.76 (1.91,3.62)	4.64(2.38,6.89)
b_1	-	-	0.02 (-1.16,1.18)	-0.01 (-1.36,1.34)	-	-
Total effect	-1.20 (-2.75,0.17)	-1.39 (-3.47,0.28)	-1.19 (-2.78,0.20)	-1.41 (-3.58,0.27)	-1.42 (-2.64,-0.19)	-1.72 (-3.14,-0.31)
Total carryover effect	-0.45 (-1.97,0.55)	-0.57 (-2.46,0.67)	-0.45 (-1.96,0.50)	-0.58 (-2.49,0.64)	-	-
β_0	-0.75 (-2.14,0.55)	-0.82 (-2.50,0.70)	-0.74 (-2.16,0.54)	-0.83 (-2.59,0.69)	-1.42 (-2.64,-0.19)	-1.72 (-3.14,-0.31)
β_1	-0.27 (-1.31,0.53)	-0.35 (-1.63,0.62)	-0.27 (-1.27,0.50)	-0.36 (-1.67,0.59)	-	-
β_2	-0.09 (-0.78,0.45)	-0.13 (-0.99,0.53)	-0.09 (-0.76,0.42)	-0.13 (-1.00,0.55)	-	-
β_3	-0.05 (-0.54,0.30)	-0.08 (-0.72,0.36)	-0.06 (-0.53,0.27)	-0.09 (-0.74,0.35)	-	-
β_4	-0.03 (-0.35,0.21)	-0.05 (-0.49,0.27)	-0.03 (-0.34,0.20)	-0.04 (-0.49,0.28)	-	-
β_5	-0.01 (-0.21,0.17)	0.01 (-0.24,0.30)	-0.01 (-0.21,0.16)	0.02 (-0.23,0.32)	-	-
β_6	0 (-0.12,0.13)	0.02 (-0.15,0.25)	0 (-0.12,0.12)	0.02 (-0.15,0.26)	-	-
β_7	0 (-0.09,0.08)	0.01 (-0.12,0.15)	0 (-0.08,0.08)	0.01 (-0.12,0.15)	-	-
ϕ_1	0.44 (0.26,0.62)	0.34 (0.10,0.57)	0.30 (0.08,0.52)	0.33 (0.09,0.57)	0.43 (0.24,0.59)	0.36 (0.17,0.54)
ϕ_2	-	-0.07 (-0.34,0.20)	-	-0.07 (-0.35,0.19)	-	-0.01 (-0.21,0.19)
ϕ_3	-	0.03 (-0.23,0.30)	-	0.03 (-0.23,0.30)	-	0 (-0.20,0.21)
ϕ_4	-	0.06 (-0.21,0.32)	-	0.06 (-0.21,0.33)	-	0.13 (-0.07,0.33)
ϕ_5	-	0.08 (-0.18,0.33)	-	0.07 (-0.18,0.32)	-	0.11 (-0.10,0.31)
ϕ_6	-	0.02 (-0.25,0.32)	-	0.03 (-0.25,0.31)	-	0.06 (-0.16,0.28)
ϕ_7	-	0.03 (-0.26,0.32)	-	0.04 (-0.25,0.33)	-	-0.09 (-0.29,0.12)
p-value	0.36	0.63	0.35	0.67	<0.001	0.91

References


Liao, Ziwei and Qian, Min and Kronish, Ian M and Cheung, Ying Kuen (2023)
Analysis of N-of-1 trials using Bayesian distributed lag model with autocorrelated errors
Statistics in Medicine 42(13), 2044 – 2060.

R functions: https://github.com/williammomo/BDLM-AR

Thank you!