Adaptive trial designs in rare disorders

Gary R Cutter MS PHD

Emeritus Professor of Biostatistics

UAB School of Public Health

Senior Scientist

The Institute for Human & Machine Cognition (IHMC)

THE UNIVERSITY OF ALABAMA AT BIRMINGHAM

ihmc
FLORIDA INSTITUTE FOR HUMAN & MACHINE COGNITION

Disclosures

- **Data and Safety Monitoring Boards:** Applied Therapeutics, AI therapeutics, AMO Pharma, Argenx, Astra-Zeneca, Avexis Pharmaceuticals, Bristol Meyers Squibb/Celgene, CSL Behring, Cynata Therapeutics, DiameticaTherapeutics, Horizon Pharmaceuticals, Immunic, Inhibrix, Karuna Therapeutics, Kezar Life Sciences, Medtronic, Merck, Mitsubishi Tanabe Pharma Holdings, Prothena Biosciences, Novartis, Pipeline Therapeutics (Contineum), Regeneron, Sanofi-Aventis, Reata Pharmaceuticals, Teva Pharmaceuticals, United BioSource LLC, University of Texas Southwestern, Visioneering Technologies, Inc.

- **Consulting or Advisory Boards:** Alexion, Antisense Therapeutics/Percheron, Avotres, Biogen, Clene Nanomedicine, Clinical Trial Solutions LLC, Endra Life Sciences, Cognito Therapeutics, Genzyme, Genentech, Hoya Corporation, Immunic, Immunosis Pty Ltd, Klein-Buendel Incorporated, Linical, Merck/Serono, Perception Neurosciences, Protalix Biotherapeutics, Regeneron, Roche, SAB Biotherapeutics, Sapience Therapeutics, Scott&Scott LLP.

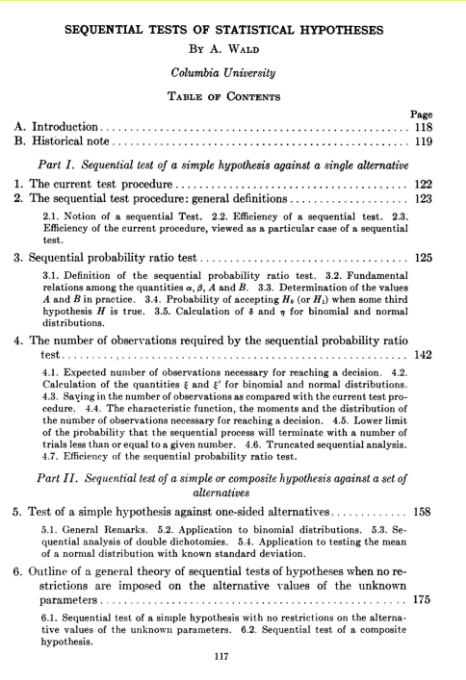
Adaptive Designs
Little Word - Many Meanings

- Requires the trial to be conducted in incremental stages with access to the accumulated data and predefined decisions
- Adaptive design may adapt using:
  - Allocation Rule:
    - how subjects are allocated to treatments
  - Sampling Rule:
    - how many subjects are used in the next stage
  - Stopping Rule:
    - when to stop the trial or a treatment dose (for efficacy, for harm, for futility)
  - Decision Rule:
    - how the next steps move forward

Sequential Designs
Test between two hypotheses H1 and H2

Sequential Probability Ratio Test SPRT

SPRT: Continue testing until outcome parameter crosses an upper or lower threshold



SEQUENTIAL TESTS OF STATISTICAL HYPOTHESES

By A. Wald

Columbia University

Table of Contents

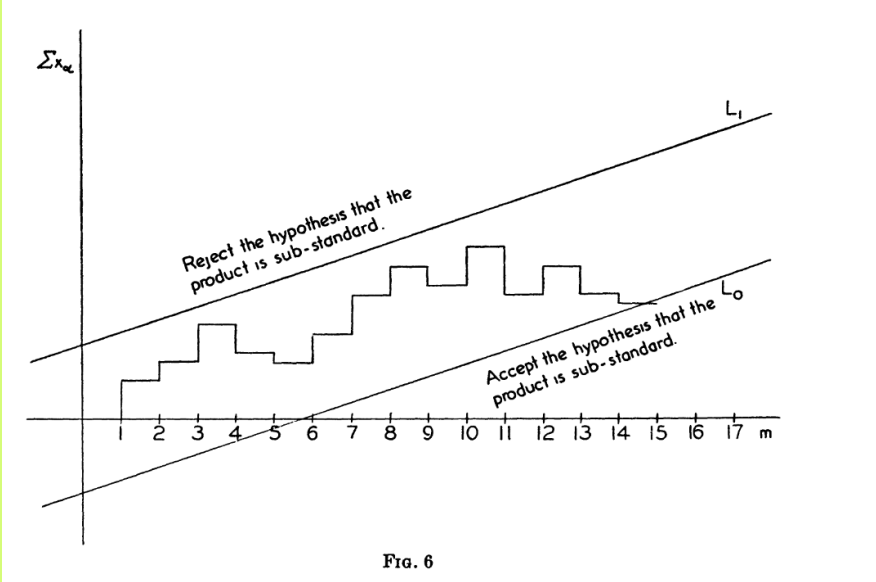
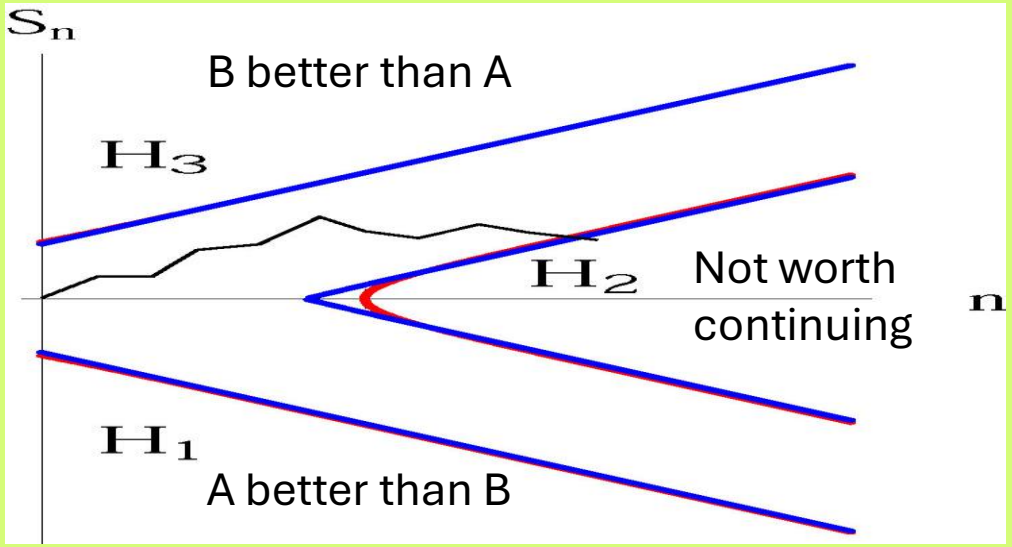


Fig. 6



Wald A. Sequential Tests of Statistical Hypotheses, The Annals of Mathematical Statistics , Jun., 1945, Vol. 16, No. **2 (Jun., 1945),** pp. 117-186

SPRT Optimality Theorem: (Wald) Among all tests with a given bound on the error rate, the SPRT minimizes the expected number of trials

Limitations of Sequential Trials

- Requires rapid outcome
  - e.g. immediate cessation of a symptom, bone marrow engraftment within 30 days etc.
  - Oxygenation response to intervention
  - Biomarker changes (may be useful for rare chronic diseases)
- May not enter boundary and can lead to larger sample sizes than fixed trials, but the average sample size is smaller!
- Assumes accrual is limitless and sequential with homogeneous population

Designs In Rare Diseases

- Rare serious diseases present two problems
  - Small available sample
  - Reluctance to use a placebo
- A series of underpowered studies **are not** the answer
- With almost all rare diseases - precision of the estimates available for planning may be less than adequate
  - Sample size adjustments may be worthwhile in these instances – hence adaptive designs

Sample Size Reassessment Decided Upon Initial Assumptions (blinded SSR) or Observed Results (Unblinded SSR)

- Based on initial assumptions

  So called nuisance parameters only

  – Placebo rate of events or variance of estimates, σ^2

  - **No Type I error penalty**

- Based on observed results
  - Estimated mean differences or effect size at interim
  - **Important Type I error penalties must be considered**
  - Caution in observing part of the whole - assumes patient population the same over time

Why Adaptive Designs For Dose Finding?

- At final analysis we find out that:
  - no doses are effective OR
  - we missed obtaining a significant result because our original assumptions were incorrect (usually optimistic)
- Standard Dose Ranging Design
  - known entity, but lacks flexibility
- Adaptive Design
  - May save both resources and time if there are clear signs that Rx does not work!
  - Allows for addition of more patients to a promising dose
    - Protects against underestimate of the variance
  - Potential to get to decision quicker,
  - May provide more information on doses of interest
  - Statistical validity maintained despite changing plans

Three Basic Concepts

- Classical issues: Interim monitoring to stop as soon as possible for Safety, Efficacy or Futility
- Futility Studies – small studies designed to assess if going forward makes any sense
- Short term small trials to arrive at what treatments to pursue or to alter design going forward
  - Sequential Trials (Old stuff)
  - Adaptive Designs (Some Old, some New)
    - Adaptive Designs -current BIG BUZZ Word – spoken as novel, innovative, NEW but….
  - Small trials (Not New when poorly planned – New with good planning)

Possible issues due the early stopping of a trial

Problem	Reason
Lack of credibility	Small trials are not convincing
Lack of realism	Dramatic treatment difference is implausible
Imprecision	Wide confidence interval for treatment effect
Bias	Trial is liable to stop on a random high or low
Speed	Time spent and information obtained may be insufficient to allow consideration of overall balance of costs and benefits. Stopping early can seem more important than completing the trial
Pressure	Unduly enthusiastic recommendation for practice may follow
Mistakes	Risk exists for false-positive or false negative results

Pocock SJ. When to stop a clinical trial. BMJ 1992;305:235-240

Interim analyses and adaptations are performed for many reasons

- to stop enrollment in the control arm so that all future enrollment is in the test regimen.
- to stop all enrollment because of disappointing results.
- to increase enrollment to reach a larger sample size.
- All such decision points **must be planned and pre-specified**.
- extra burden on the monitoring and data management groups.

Cautions

- Logistics issues critical to adaptive designs
- Must establish a DSMB with a specific charter and rules for actions more complete than for pivotal trials
- Should have adaptation performed by an independent third party with no conflicts of interest issues
  - (intellectual almost more than financial)
- During adaptation considerations, unblind only people that are necessary to make or implement a decision
- Patient recruitment is not interrupted during adaptation consideration and must be factored into ultimate benefit

Interim Analyses or Adaptation entail careful planning of the protocol

- Exacting detail of the statistical design and analysis that can be fixed in advance is provided in the protocol:
    - number of interim analyses or adaptations (to control Type I errors)
    - information rates (how much of the data are available)
        - Too soon → too little information generally not warranted for power assessment until at least 50% of the study is over.
    - stopping guidelines → doesn't necessitate stopping necessitates considering stopping!
    - Tests or methods utilized in these assessments should be specified
- The time of the Interim Analysis is unknown to the investigators, if possible. Can impact recruitment, etc.

Adaptive Designs Often Ignore the Consequences of Entrance Criteria

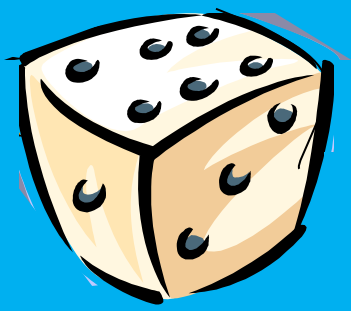
- Can create false positive treatments
  - Placebo effects
  - Regression toward the mean effects
- Mask treatment effects due to additional therapies
- Assume recruitment is homogeneous over time
  - Often a problem of the prevalent pool versus incident pool

Regression Toward the Mean
Let the die face be a clinical Scale AND inclusion criteria as die roll of 1 to 3 and assume higher scores are better

Roll a Die 60 times

- 1 – 10 times
- 2 – 10 times
- 3 – 10 times
- 4 – 10 times
- 5 – 10 times
- 6 – 10 times

Mean = 2.0 among 30 eligible



What Now Happens to the 30 Enrolled Patients at the Next Visit?

Roll a Die 30 times again for their post enrollment visit

- 1 – 5 times =105/30
- 2 – 5 times
- 3 – 5 times
- 4 – 5 times
- 5 – 5 times
- 6 – 5 times

mean=(5*1+5*2+…+5*6)/30

Mean = 3.5

Average improvement:

= new minus old mean

= 3.5 - 2.0

= **1.5 units**

50% of these “patients” improved and were “responders”!

Does Simvastatin Reduce Gd Lesions?

- **An Open-Label, Single Arm Study of Simvastatin as a Therapy for Multiple Sclerosis (MS).**
- Vollmer T, Key L, Durkalski V, Tyor W, Corboy J, Markovic-Plese S, Preiningerova J, Rizzo M, Singh I. Oral simvastatin treatment in relapsing-remitting multiple sclerosis. Lancet. 2004 May 15;363(9421):1607-8. PubMed PMID: 15145635.

...clinically diagnosed with RRMS were invited to participate. Subjects underwent a pre-treatment period of three months. Subjects that had at least one Gd-enhanced lesion were eligible for six months of treatment with simvastatin....

Results: mean number of gadolinium lesions

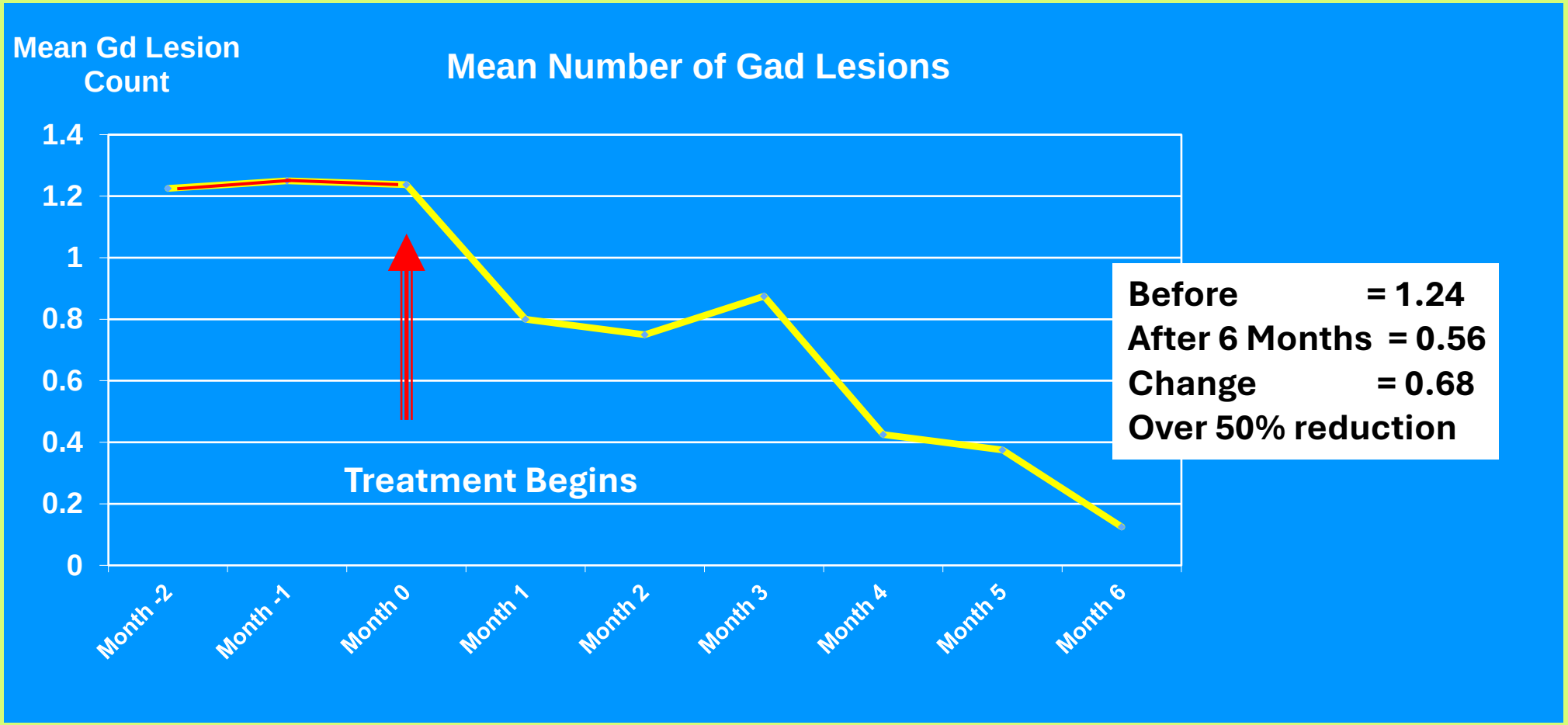
pre-treatment = 2.35 lesions

post-treatment = 1.31 lesions

reduction of 44%!

- ...CONCLUSIONS: Preliminary data suggest that daily treatment with 80mg of simvastatin may be safe and effective for the treatment of RR-MS. Randomized-controlled studies will need to be conducted to definitively ascertain the effectiveness of this treatment.

“Treatment” Applied in a Similar Manner to a Placebo Group from an MS Trial



Cognitive rehabilitation and aerobic exercise for cognitive impairment in people with progressive multiple sclerosis (CogEx): a randomised, blinded, sham-controlled trial

*Anthony Feinstein, Maria Pia Amato, Giampaolo Brichetto, Jeremy Chataway, Nancy D Chiaravalloti, Gary Cutter, Ulrik Dalgas, John DeLuca, Rachel Farrell, Peter Feys, Massimo Filippi, Jennifer Freeman, Matilde Inglese, Cecilia Meza, Robert W Motl, Maria A Rocca, Brian M Sandroff, Amber Salter, on behalf of the CogEx Research Team**

Lancet Neurol 2023; 22: 912–24

See Comment page 875

*Members listed in the appendix

Department of Psychiatry, University of Toronto, Toronto, ON, Canada (Prof A Feinstein MD, C Meza MA); **Sunnybrook Health Sciences Centre, Toronto, ON, Canada** (Prof A Feinstein, C Meza); **Department NEUROFARBA, Section Neurosciences, University of Florence, Florence, Italy** (Prof M P Amato MD); **IRCCS Fondazione Don Carlo Gnocchi, Florence, Italy** (Prof M P Amato); **Scientific Research Area, Italian Multiple Sclerosis Foundation, Genoa, Italy** (G Brichetto MD); **AISM Rehabilitation Service, Italian Multiple Sclerosis Society, Genoa, Italy** (G Brichetto); **Queen Square Multiple Sclerosis Centre, Department of Neuroinflammation, UCL Queen Square Institute of Neurology, Faculty of Brain Sciences, University College London, London, UK** (Prof J Chataway MD, Prof R Farrell MD); **National Institute for Health Research, University College London Hospitals, Biomedical Research Centre, London, UK** (Prof J Chataway, Prof R Farrell); **Kessler Foundation, East Hanover, NJ, USA** (Prof N D Chiaravalloti PhD, Brian M Sandroff PhD, Prof J DeLuca PhD); **Department of Physical Medicine & Rehabilitation, Rutgers New Jersey Medical School, Newark, NJ, USA** (Prof N D Chiaravalloti, Brian M Sandroff, Prof J DeLuca); **Department of Biostatistics,**

Summary

Background Cognitive dysfunction in people with relapsing-remitting multiple sclerosis can improve with cognitive rehabilitation or exercise. Similar effects have not been clearly shown in people with progressive multiple sclerosis. We aimed to investigate the individual and synergistic effects of cognitive rehabilitation and exercise in patients with progressive multiple sclerosis.

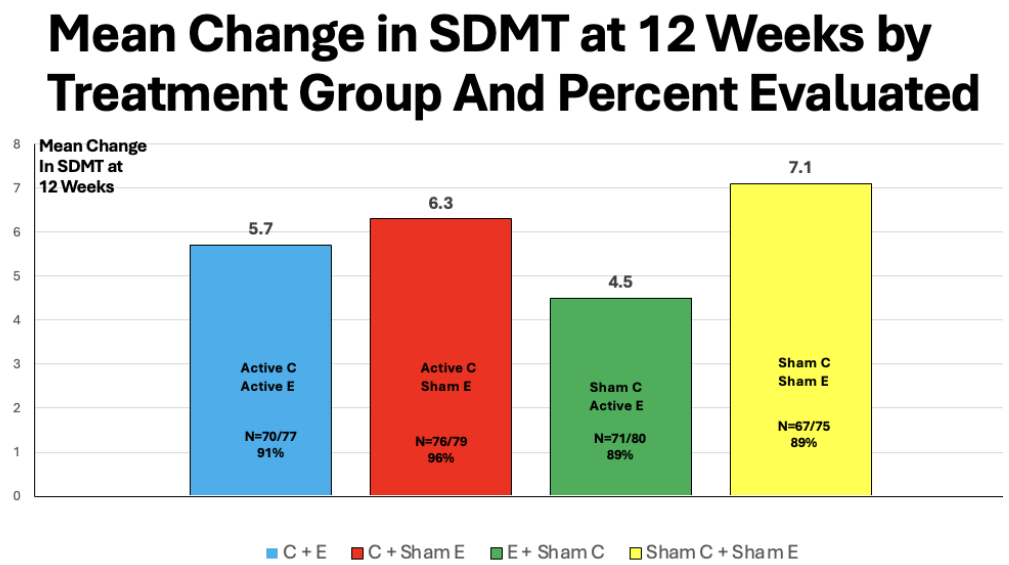
Methods CogEx was a randomised, sham-controlled trial completed in 11 hospital clinics, universities, and rehabilitation centres in Belgium, Canada, Denmark, Italy, UK, and USA. Patients with progressive multiple sclerosis were eligible for inclusion if they were aged 25–65 years and had an Expanded Disability Status Scale (EDSS) score of less than 7. All had impaired processing speed defined as a performance of 1·282 SD or greater below normative data on the Symbol Digit modalities Tests (SDMT). Participants were randomly assigned (1:1:1:1), using an interactive web-response system accessed online from each centre, to cognitive rehabilitation plus exercise, cognitive rehabilitation plus sham exercise, exercise plus sham cognitive rehabilitation, or sham exercise plus sham cognitive rehabilitation. The study statistician created the randomisation sequence that was stratified by centre. Participants, outcome assessors, and investigators were blinded to group allocation. The study statistician was masked to treatment during analysis only. Interventions were conducted two times per week for 12 weeks: cognitive rehabilitation used an individualised, computer-based, incremental approach to improve processing speed; sham cognitive rehabilitation consisted of internet training provided individually; the exercise intervention involved individualised aerobic training using a recumbent arm–leg stepper; and the sham exercise involved stretching and balance tasks without inducing cardiovascular strain. The primary outcome measure was processing speed measured by SDMT at 12 weeks; least squares mean differences were compared between groups using linear mixed model in all participants who had a 12-week assessment. The trial is registered with ClinicalTrials.gov, NCT03679468, and is completed.

Findings Between Dec 14, 2018, and April 2, 2022, 311 people with progressive multiple sclerosis were enrolled and 284 (91%) completed the 12-week assessment (117/311 [38%] male and 194/311 [62%] female). The least squares mean group differences in SDMT at 12 weeks did not differ between groups (p=0·85). Compared with the sham cognitive rehabilitation and sham exercise group (n=67), differences were −1·30 (95% CI −3·75 to 1·16) for the cognitive rehabilitation plus exercise group (n=70); −2·78 (−5·23 to −0·33) for the sham cognitive rehabilitation plus exercise group (n=71); and −0·71 (−3·11 to 1·70) for the cognitive rehabilitation plus sham exercise group (n=76). 11 adverse events possibly related to the interventions occurred, six in the exercise plus sham cognitive rehabilitation group (pain, dizziness, and falls), two in the cognitive rehabilitation plus sham exercise group (headache and pain), two in the cognitive rehabilitation and exercise group (increased fatigue and pain), and one in the dual sham group (fall).

Interpretation Combined cognitive rehabilitation plus exercise does not seem to improve processing speed in people with progressive multiple sclerosis. However, our sham interventions were not inactive. Studies comparing interventions with a non-intervention group are needed to investigate whether clinically meaningful improvements in processing speed might be attainable in people with progressive multiple sclerosis.

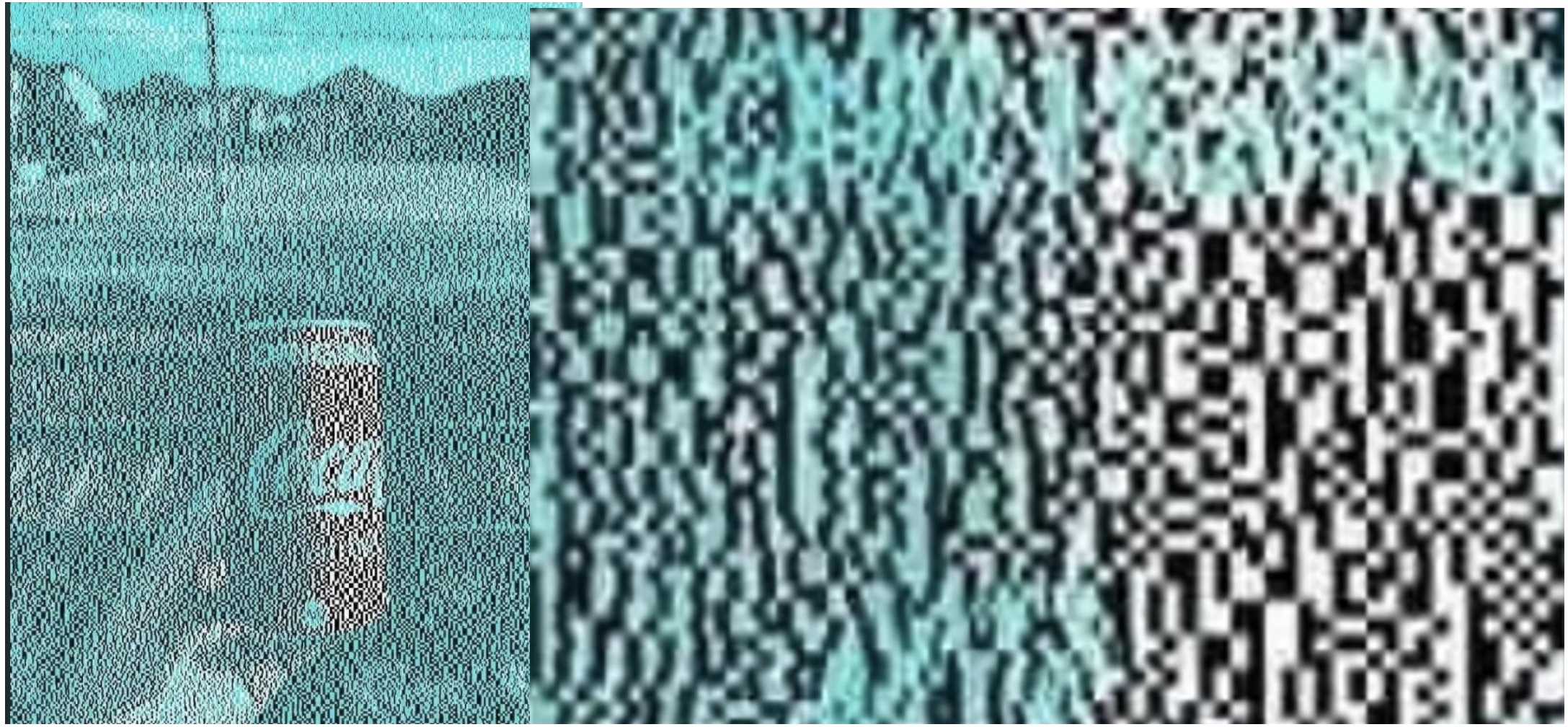
Funding MS Canada.

	Total	Cognitive rehabilitation plus exercise	Cognitive rehabilitation plus sham exercise	Exercise plus sham cognitive rehabilitation	Sham cognitive rehabilitation plus sham exercise	p value*
Number of participants						
Baseline	311	77	79	80	75	..
12 weeks	284	70	76	71	67	..
6 months	258	65	68	65	60	..
Cognitive outcomes						
SDMT						
Baseline	33·4 (8·2)	32·2 (8·6)	33·0 (7·4)	35·1 (8·1)	33·3 (8·4)	..
12 weeks†	39·3 (11·5)	38·0 (11·9)	39·1 (10·3)	39·9 (11·1)	40·2 (12·8)	0·85
6 months‡	36·8 (11·6)	35·8 (11·1)	35·9 (12·5)	37·9 (10·3)	37·8 (12·4)	..
Difference in SDMT§						
Baseline to 12 weeks¶	5·9 (7·5)	5·7 (7·2)	6·3 (6·6)	4·5 (7·5)	7·1 (8·6)	0·23



Mean Change in SDMT at 12 Weeks by Treatment Group And Percent Evaluated

Processing Speed impacts cognition, but may not be measuring the cognitive problems of interest



Ethics

- Withdrawal of data by an individual – what does it mean to have the right to withdraw your data?
  - In screening?
  - At end of trial?
  - At time of making data public?
- During a trial what rules to protect a participant and their impact?
  - Withdraw at first event
  - Eliminates multiple occurrences
  - Shortens exposure time

Designs In Rare Diseases

- Limiting exposure time can be an acceptable approach when the events occur reasonably rapidly to minimize exposure to placebos

- While not optimal from a statistical perspective using k to 1 randomization or matching may be preferred to 1 to 1 randomization from the clinical perspective.

Is a Placebo Tolerable or for How Long?

- The traditional designs follow patients on their treatment assignment until a specified time or number of medically relevant events have accrued.
- This strategy is not considered tenable for many studies because indefinite placebo exposure is considered an excessive burden or risk for the study subjects.
- These are not statistical considerations rather decisions of the heart!

RESEARCH ARTICLE

Early Treatment with Vigabatrin Does Not Decrease Focal Seizures or Improve Cognition in Tuberous Sclerosis Complex: The PREVeNT Trial

Elizabeth Martina Bebin, MD, MPA [1], Jurriaan M. Peters, MD, PhD [2], Brenda E. Porter, MD, PhD,[3] Tarrant O. McPherson, PhD,[4] Sarah O'Kelley, PhD,[5] Mustafa Sahin, MD, PhD [2,6], Katherine S. Taub, MD,[7] Rajsekar Rajaraman, MD,[8] Stephanie C. Randle, MD,[9] William M. McClintock, MD,[10] Mary Kay Koenig, MD,[11] Mike D. Frost, MD,[12] Hope A. Northrup, MD,[11] Klaus Werner, MD, PhD,[13] Danielle A. Nolan, MD,[14] Michael Wong, MD, PhD,[15] Jessica L. Krefting, RN, BSN,[1] Fred Biasini, PhD,[5] Kalyani Peri, MS,[16] Gary Cutter, PhD,[16] and Darcy A. Krueger, MD, PhD, [17] on behalf of the PREVeNT Study Group

Objective: This study was undertaken to test the hypothesis that early vigabatrin treatment in tuberous sclerosis complex (TSC) infants improves neurocognitive outcome at 24 months of age.
Methods: A phase IIb multicenter randomized double-blind placebo-controlled trial was conducted of vigabatrin at first epileptiform electroencephalogram (EEG) versus vigabatrin at seizure onset in infants with TSC. Primary outcome was Bayley Scales of Infant and Toddler Development, Third Edition (Bayley-III) cognitive assessment score at 24 months. Secondary outcomes were prevalence of drug-resistant epilepsy, additional developmental outcomes, and safety of vigabatrin.
Results: Of 84 infants enrolled, 12 were screen failures, 4 went straight to open label vigabatrin, and 12 were not randomized (normal EEG throughout). Fifty-six were randomized to early vigabatrin (n = 29) or placebo (n = 27). Nineteen of 27 in the placebo arm transitioned to open label vigabatrin, with a median delay of 44 days after randomization.

View this article online at wileyonlinelibrary.com. DOI: 10.1002/ana.26778

Received May 24, 2023, and in revised form Aug 7, 2023. Accepted for publication Aug 20, 2023.

Address correspondence to Dr Bebin, Department of Neurology, University of Alabama, 1719 6th Avenue South CIRC 312 Birmingham, AL 35294.
E-mail: ebebin@uabmc.edu

Members of the PREVeNT Study Group who contributed to the completion of this trial are available as an online Supplementary file.

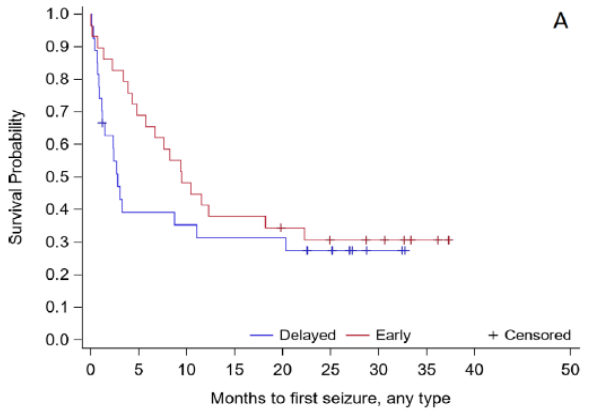
From the [1]Department of Neurology, University of Alabama at Birmingham, Birmingham, AL, USA; [2]Department of Neurology, Boston Children's Hospital, Harvard Medical School, Boston, MA, USA; [3]Department of Neurology, Stanford University, Stanford, CA, USA; [4]Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA, USA; [5]Department of Psychology, University of Alabama at Birmingham, Birmingham, AL, USA; [6]Rosamund Stone Zander Translational Neuroscience Center, Boston Children's Hospital, Harvard Medical School, Harvard University, Boston, MA, USA; [7]Department of Pediatrics, Children's Hospital of Philadelphia, Philadelphia, PA, USA; [8]Department of Pediatrics and Psychiatry and Biobehavioral Sciences, University of California, Los Angeles, Los Angeles, CA, USA; [9]Department of Pediatrics, Division Pediatric Neurology and Epilepsy, Seattle Children's Hospital, Seattle, WA, USA; [10]Department of Pediatrics, Division of Neurology, Children's National Medical Center, Washington, DC, USA; [11]Department of Pediatrics, McGovern Medical School at University of Texas Health Science Center at Houston and Children's Memorial Hermann Hospital, Houston, TX, USA; [12]Minnesota Epilepsy Group, P.A., Minnesota Epilepsy Group, Roseville, MN, USA; [13]Department of Pediatrics, Duke University, Durham, NC, USA; [14]Beaumont Florence and Richard McBrien Pediatric Neuroscience Center, Beaumont Hospital, Royal Oak, MI, USA; [15]Department of Neuroscience, Washington University in Saint Louis, Saint Louis, MO, USA; [16]Department of Biostatistics, University of Alabama at Birmingham, Birmingham, AL, USA; and [17]Department of Pediatrics, University of Cincinnati, Cincinnati, OH, USA

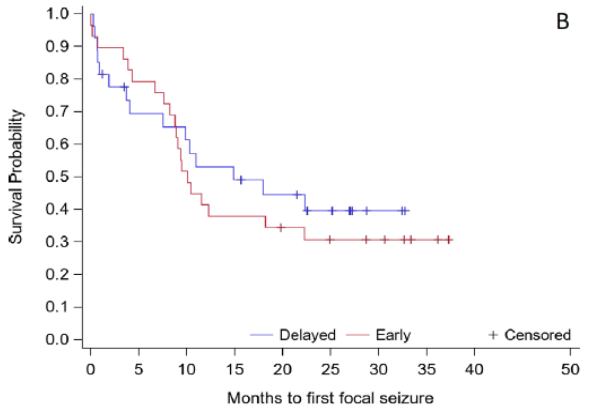
Additional supporting information can be found in the online version of this article.

This trial used the strategy to protect a child as soon as possible!

If a child developed a seizure, they were offered the Vigabatrin and Followed in the study as per protocol

Seizure Results





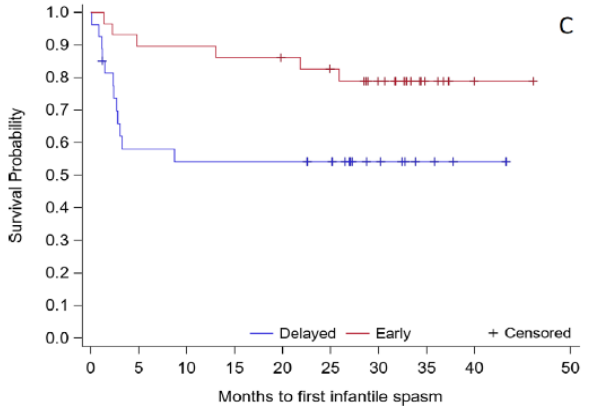
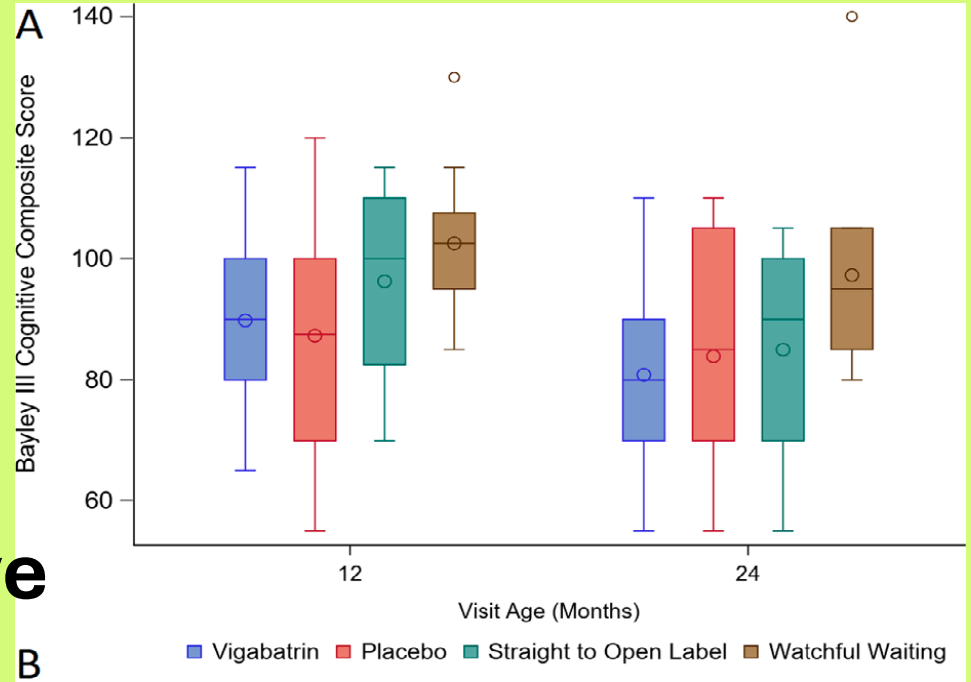
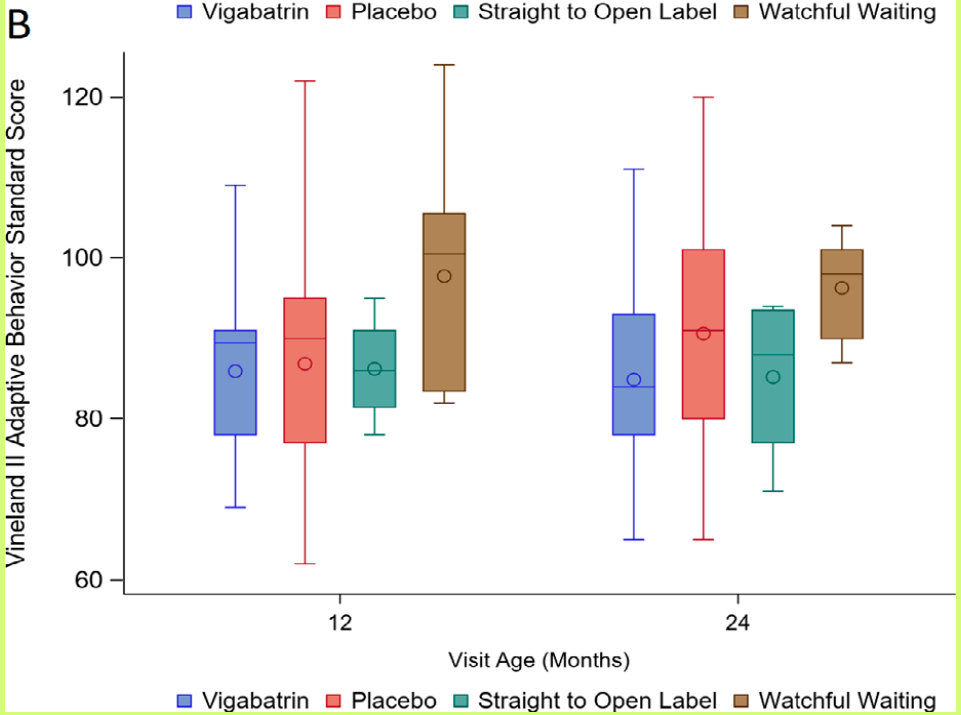


FIGURE 3: Time from randomization to first (A) any type of seizure, (B) focal seizure, (C) infantile spasms.

Cognitive Results





- Participants in the placebo group were transitioned to open label vigabatrin in blinded fashion following the occurrence of a first clinical or electrographic seizure for ethical reasons.

- **In this latter group, seizure onset and vigabatrin treatment initiation was between 3 to 610 days (median = 44 days, interquartile range = 21–90) after randomization.**

- As such, **participants in both groups** were receiving the vigabatrin treatment for a large portion of the randomized study period by the time they had reached the age of 24 months.

- On average (639-44)/639 =93% of the time – Cognitive benefit needed to be conveyed by preventing 1 episode.

N of 1 Trials
A Feasible Design for Rare Diseases

- Basically, clinical medicine as in practice with a twist – designed and statistical evaluation

- Can be done over a group of participants (how likely x successes out of N trials)

- Within participant as well as among participant evaluations – patient serves as their own control

- Great reference: Diamond Project https://orda.shef.ac.uk/articles/report/Guidance_for_designing_n-of-1_trials_the_DIAMOND_project_/22004438/1?file=39055724

Essential Characteristics of N-of-1 Studies

The specific design of the trial depends on the question of interest. The N-of-1 trial design is suitable only in situations where:

- There is substantial uncertainty about the optimal treatment path for a patient:
  - Lack of evidence to support clinical decision (often in a rare condition)
  - Evidence of heterogeneity in the effectiveness of a treatment (e.g. contradictory or mixed effects reported)
  - Patient characteristics are not represented in existing clinical trials or guidelines for their condition (e.g. comorbidities, age, concurrent medication)

- The clinical condition is chronic or frequently recurs and thus is essentially similar over time (allows for crossover trials within each participant)

- The treatment being considered demonstrates measurable outcomes within a short period

- Both patient and clinician are committed to the effort required to undertake a trial.

In contrast, the **N-of-1 trial design is NOT suitable** for acute conditions or ones that progress rapidly; and are more challenging for treatments that have a slow onset and long carryover effects once Treatment begins or ceases.

Design and Implementation of N-of-1 Trials:

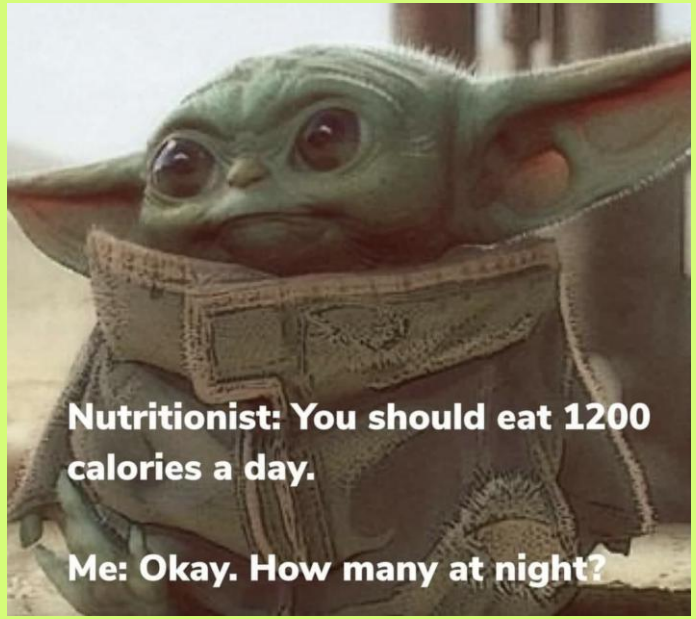
- Treatment assignment needs to be balanced across treatment conditions, using either randomization or counterbalancing, along with blocking
- Blind treatment assignment when feasible
- Invoke appropriate measures to deal with potential bias due to carryover and slow onset effects
- Perform multiple assessments within treatment periods
- Consider adaptive trial designs and sequential stopping rules to maximize allocation to promising treatments
- Use appropriate statistical method to analyze outcome data, taking into consideration important features of time-series data, including autocorrelation, time trend, and repeated measures within treatment periods
- Borrow from strength

Why A Futility Trial

- How can we rule out treatments that are not going to lead anywhere?
- Use existing data to set a hurdle that we need to exceed
- Compare a single group study to the historical results and make a decision:
  - Historical proportion worsening in 6 or 12 months – say p_0
  - Use the binomial Distribution to reject that the test treatment is better
    - i.e. what is the probability of observing x or fewer worsenings out of “n” patients when the expected number is p_0
- Futility trials are not the same as a futility analyses! Stopping a study for futility is simply saying there is no reason to continue as we are highly unlikely to reject the null hypothesis and we should stop.

Futility Trials – Doing More with Fewer

- **In futility trials:**
  - **The null hypothesis is NEGATIVE -**
    - **the treatment is worse not better or it will not increase treatment successes.**
  - **You plan to reject that it is not the same or worse THUS NOT Futile.**
  - Benefits are reduced sample size; speed of rejection of poorly performing therapies AND can be compared to historical controls to further reduce sample size.
  - If the treatment does achieve the objective number of successes, it is declared **non-futile and considered to be worthy of further investigation.**
  - Futility trials cannot really prove efficacy - it is necessary to establish efficacy/effectiveness in a follow-up RCT.



Futility Trials Can Be Adaptive

- Futility trials are small enough to be done at 1 site by a single investigator

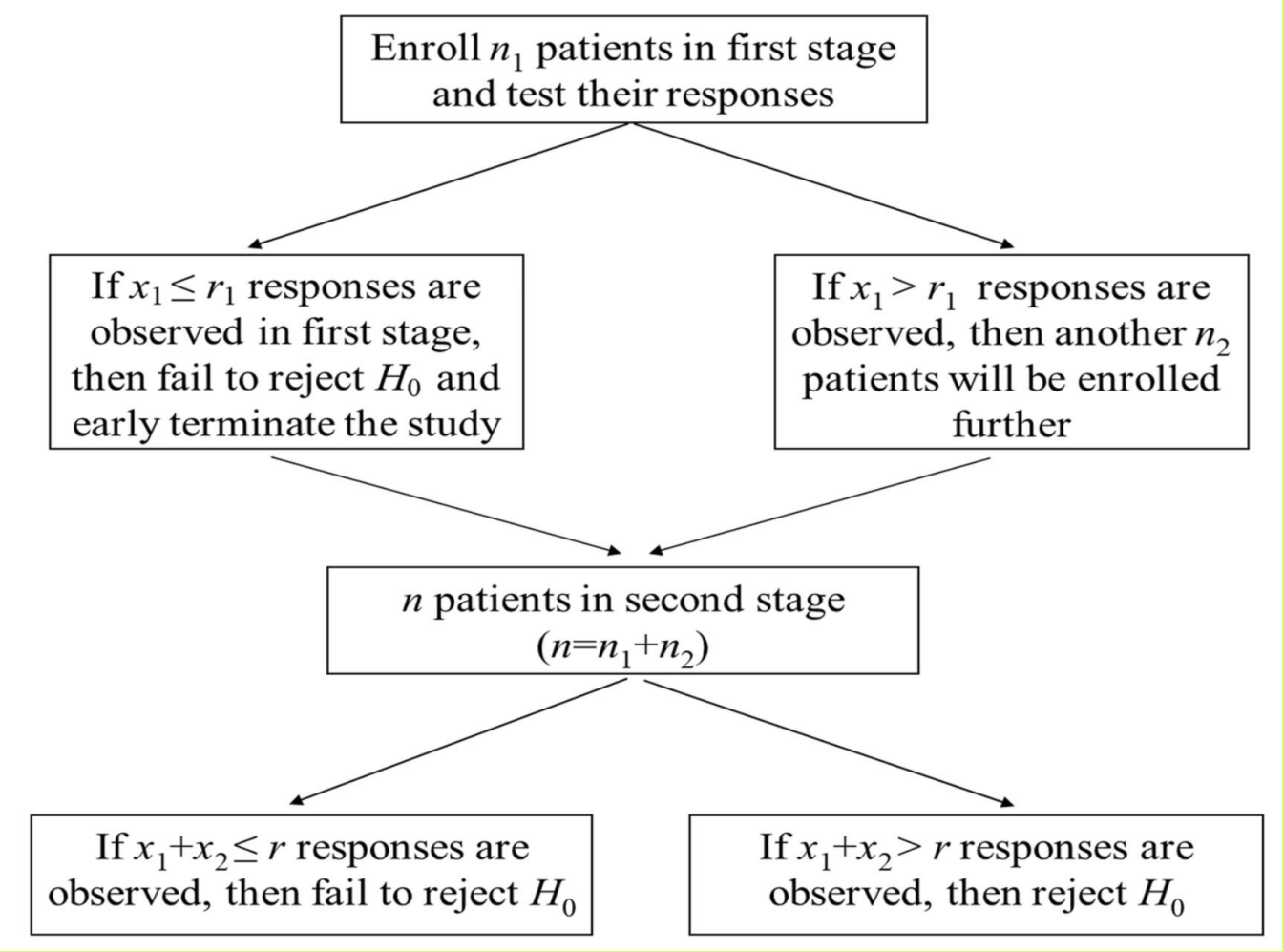
- Very useful for testing repurposed drugs that pharma and reviewers are often not very interested in or positive in and thus rarely recommend funding.

- They offer independent research with appropriate rigor to provide feasibility data for definitive trials

So what is a Simon Two Stage Design

- We want to rule out a futile treatment as soon as possible!
- So the two stage approach takes an interim look and if the results don't justify going on at this first stage – we stop and bag the test treatment
- The mathematics adjust for the fact that we are looking at part of the data and have to take that into account when we complete the trial.
- Another related type of design is Fleming's Two Stage Design which is essentially the same BUT allows stopping early for overwhelming efficacy (we'll ignore this as it is generally unlikely to be needed)

Flowchart for Simon 2-Stage Design



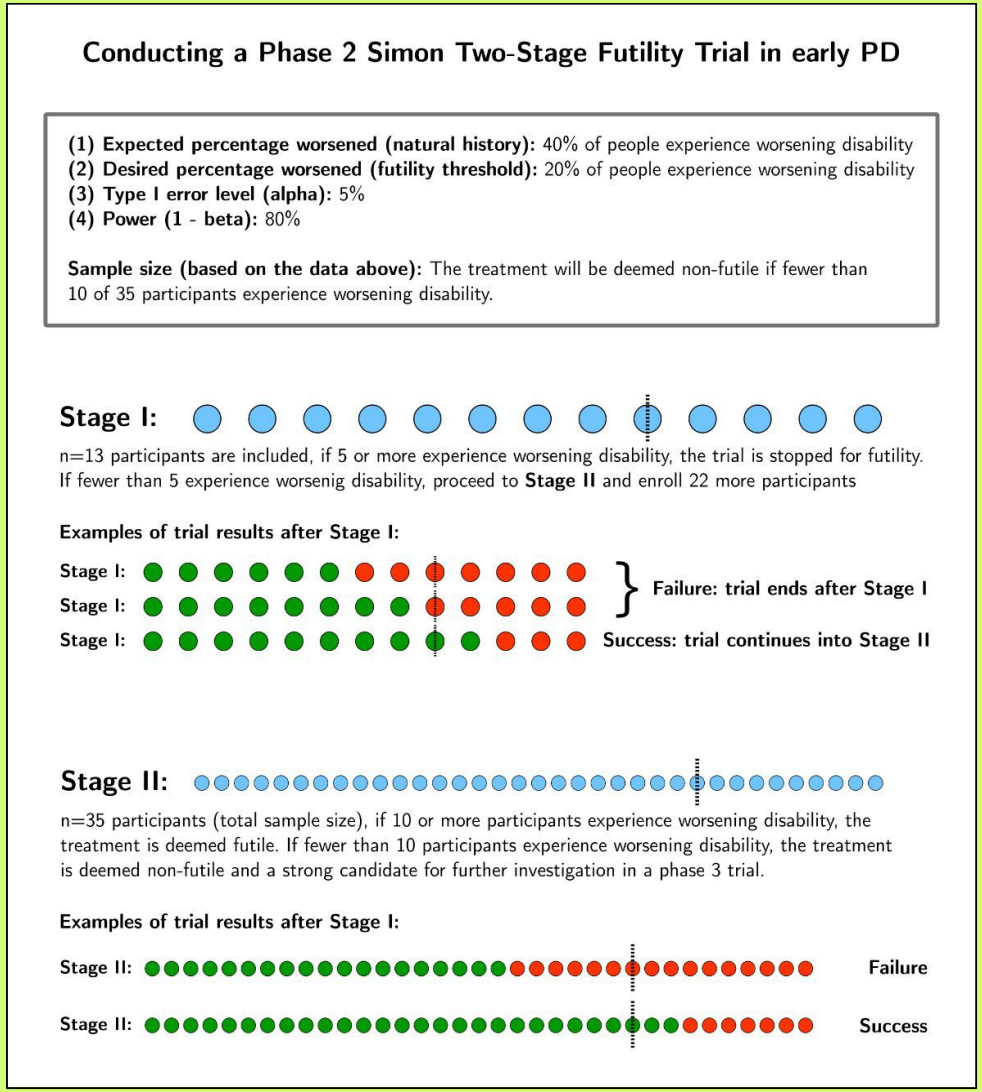
Gehan was the first to suggest a two-stage model to judge on futility in 1961 for cancer trials, but his approach had the drawback at even with only a 5% chance of being successful, it would judge non-futility 51% of the time.

Fleming proposed a new two- or three stage futility design in 1982 for cancer trials. His design allows for early termination of the study after a first stage, if the treatment is either extremely effective or extremely ineffective. It was this latter problem to stop for failure was only possible is the most extreme cases,

Simon published his paper on two related futility designs in 1989.

The **Optimal** Simon-2-stage model is designed to produce *the smallest possible sample size* for the first study phase, and the **Minimax** model is designed to reduce to overall number of patients in the study.

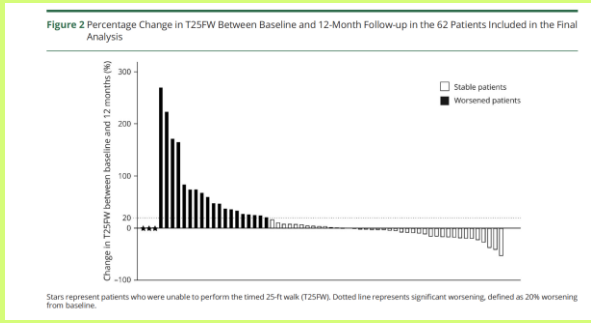
For example, a Minimax Simon-2-stage model (type I error rate of 5% and 90% power) would require 45 patients, and the first stage would include 24 patients when testing a treatment against an historical control assuming 40% versus the control of 20%. The treatment would be declared non-futile if 6 or more patients had PFS in the first phase, and if 14 or more patients had PFS after the second stage.



Two Multiple Sclerosis Futility Trials

Domperidone in Progressive MS

- In the first stage, 8 of 30 (27%) patients had experienced significant worsening in T25FW performance. Because the number of worsened patients was lower than the prespecified futility threshold of ≥12 of 30 patients, the trial continued into its second phase.
- After a year of domperidone treatment**, 22 of 62 patients (35%)** had significant worsening of their T25FW**, which is close to the 40% historical rate and futility could not be rejected.**



Hydroxychloroquine for Progressive MS

At Stage 1 -only 2 of 13 (15%) participants experienced worsening of the T25FW. this was lower than the prespecified futility threshold of 5 or more of 13, the trial continued into its second phase.

In conclusion, in people with PPMS without overt focal inflammatory disease activity, HCQ treatment was associated with reduced disability worsening. **HCQ is a promising treatment candidate for PPMS and should be investigated further in randomized controlled clinical trials.**

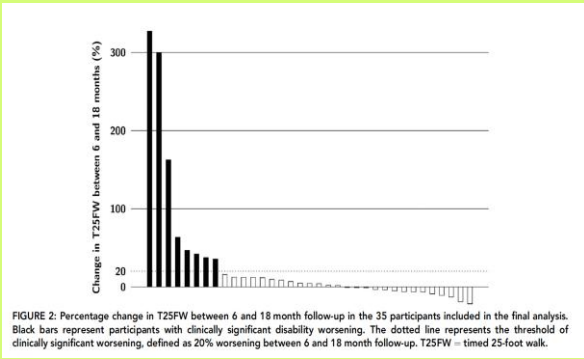
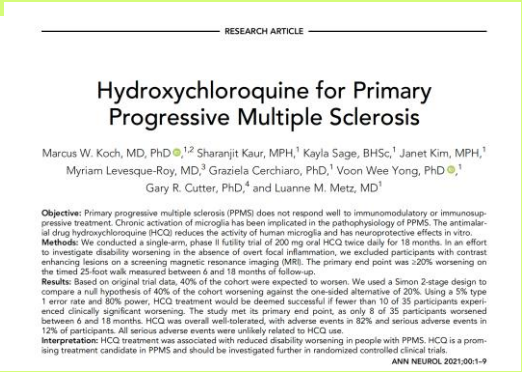


FIGURE 2: Percentage change in T25FW between 6 and 18 month follow-up in the 35 participants included in the final analysis. Black bars represent participants with clinically significant disability worsening. The dotted line represents the threshold of clinically significant worsening, defined as 20% worsening between 6 and 18 month follow-up. T25FW = timed 25-foot walk.



BARRIERS TO USE OF ADAPTIVE DESIGNS

However, before any sample size re-estimation technique can be practically implemented, there are also logistical barriers that need to be overcome.

These include:

- Budget Administration
- Information Technology
- Protocol Issues

BARRIERS TO USE OF ADAPTIVE DESIGNS

Problem: Current funding mechanisms make it difficult to include an adaptive design since final sample size may not be known at outset.

This causes logistical problems in

- Setting up a budget
- Determining the number of sites

Solution? Discussions will need to take place among sponsors to determine how to gain the advantages of adaptive designs within current funding framework.

BARRIERS TO USE OF ADAPTIVE DESIGNS

Problem: Adaptive designs require a high degree of transparency with respect to decision procedures.

Solution? The extent to which sample size re-estimation is planned should be described a priori in detail, if possible.

Hung et al. (2006):
"At the very least, the regulatory agencies need to know every detail of how the trial proceeded during its conduct and adaptations."

BARRIERS TO USE OF ADAPTIVE DESIGNS

Problem: Adaptations have potential to convey knowledge to observers based on actions taken as a result of interim results

Solution? Document details of re-estimation procedure somewhere other than the protocol?

Carefully define who should have access to the proposed procedures for re-estimating sample size.

Requires greatly expanding the responsibilities of DSMBs or independent ad hoc group.

May need sponsor representation in the process due to the nature of the adaptation decisions being made (ultimately $$$)

BARRIERS TO USE OF ADAPTIVE DESIGNS

Problem: Who should perform the calculations?

Solution? Fisher et al. (2001) suggest that an independent statistician should produce DSMB reports for clinical trials.

Previous discussions regarding the need for an independent statistician have revolved around issues of bias introduced with the traditional interim monitoring approach.

As adaptive designs become more popular, the debate regarding the independent statistician may become more important in such settings.

BARRIERS TO USE OF ADAPTIVE DESIGNS

Problem: Methods for design and analysis of adaptive designs are often more computationally complex.

Solution? Customized software programs may be required.

Availability of additional commercial software solutions for the design and analysis of adaptive designs will increase the feasibility of implementing these methods.

IN CONCLUSION: Be thoughtful, be careful and think!


Thanks For
Listening