Principles and Practice of Clinical Research
A Global Journal in Clinical Research

PPCR
ISSN: 2378-1890

Cost-Effectiveness Analysis to Inform Randomized Controlled Trial Design in Chronic Pain Research: Methods for Guiding Decisions on the Addition of a Run-In Period

H. Rafferty[1,A,#], E. Rocha[1,B,#], P. Gonzalez-Mego[1,2], CL. Ramos[1], MM. El-Hagrassy[1,2], ME. Gunduz[1,2], E. Uygur-Kucukseymen[1,2], H. Zehry[1], SS. Chaudhari[1], PEP. Teixeira[1,5,6], GR. Rosa[1], AL. Zaninotto[1], C. Connor[2,8,10], U. Eden[8], C. Ramos-Estebanez[7], F. Fregni[1,2,4,B], L. Dipietro[3,B], T. Wagner*[3,9,B]

#These authors have contributed equally to this work.
**Corresponding author: Timothy Wagner, Highland Instruments, Inc Cambridge, MA, USA. Email: twagner@mit.edu.*

Rest of author's affiliation at the end of the manuscript.

Received April 7, 2022; accepted May 14, 2022; published August 10, 2022.

Abstract:

Introduction: *Run-In (RI) periods can be used to improve the validity of randomized controlled trials (RCTs), but their utility in Chronic Pain (CP) RCTs is debated. Cost-effectiveness analysis (CEA) methods are commonly used in evaluating the results of RCTs, but they are seldom used for designing RCTs. We present a step-by-step overview to objectively design RCTs via CEA methods and specifically determine the cost effectiveness of a RI period in a CP RCT.*

Methods: *We applied the CEA methodology to data obtained from several noninvasive brain stimulation CP RCTs, specifically focusing on (1) defining the CEA research question, (2) identifying RCT phases and cost ingredients, (3) discounting, (4) modeling the stochastic nature of the RCT, and (5) performing sensitivity analyses. We assessed the average cost-effectiveness ratios and incremental cost effectiveness ratios of varied RCT designs and the impact on cost-effectiveness by the inclusion of a RI period vs. No-Run-In (NRI) period.*

Results: *We demonstrated the potential impact of varying the number of institutions, number of patients that could be accommodated per institution, cost and effectiveness discounts, RCT component costs, and patient adherence characteristics on varied RI and NRI RCT designs. In the specific CP RCT designs that we analyzed, we demonstrated that lower patient adherence, lower baseline assessment costs, and higher treatment costs all necessitated the inclusion of an RI period to be cost-effective compared to NRI RCT designs.*

Conclusions: *Clinical trialists can optimize CP RCT study designs and make informed decisions regarding RI period inclusion/exclusion via CEA methods.*

Keywords: cost-effectiveness analysis (CEA), randomized controlled trial design (RCT design), Chronic Pain Trials, Run-In Period, Noninvasive Brain Stimulation, Neuromodulation

DOI: http://dx.doi.org/10.21801/ppcrj.2022.82.5

INTRODUCTION

Run-In (RI) periods in Randomized Controlled Trials (RCTs) represent a defined time between enrollment and randomization during which all individuals receive the same intervention (i.e., placebo, active arm, or no intervention) (Brittain & Wittes, 1990; Cipriani & Geddes, 2010). This methodology aids to identify subjects that should be excluded from the trial (e.g., placebo responders, non-compliant, intervention tolerance, etc.) and/or to enrich treatment response prior to randomization (Dworkin et al., 2010; Dworkin et al., 2012; Gewandter et al., 2014; Pablos-Méndez, Barr, & Shea, 1998). Thereby, RI periods may improve the power and validity of RCTs (Hattori et al., 2019; Hewitt et al.,

2011; Kim et al., 2018; Pablos-Méndez et al., 1998; Steiner et al., 2011; Yarlas et al., 2015).

The qualitative advantages and disadvantages of including an RI in Pain RCTs, from initial proof of concept studies to confirmatory studies, has been highlighted by numerous groups (e.g., (Dworkin et al., 2010; Dworkin et al., 2012; Gewandter et al., 2014)). However, designing an efficient and effective RI period for a Pain RCT generally requires addressing numerous questions balancing "validity, generalizability, and efficiency" (Pablos-Méndez et al., 1998). Further limitations involve the general lack of quantitative standardized methods to even determine justification of an RI in a given Pain RCT.

In health-economics, Decision-Support Tools (DSTs) are frequently employed to guide decisions that maximize resources. Cost Effectiveness Analysis (CEA) is one of the most frequently used DSTs in healthcare (Organization, 2003). CEAs allow comparison of different treatment scenarios, whereby health benefits are defined in natural units and costs in monetary units, and commonly used for assessing the cost of health benefits of new vs. established treatments (Sox & Higgins, 2013). Herein, we demonstrate how a CEA approach can be used to optimize the use of resources in RCT design, with a focus on the need for an RI period in Pain RCTs. We depict a stepwise CEA approach through real world examples on

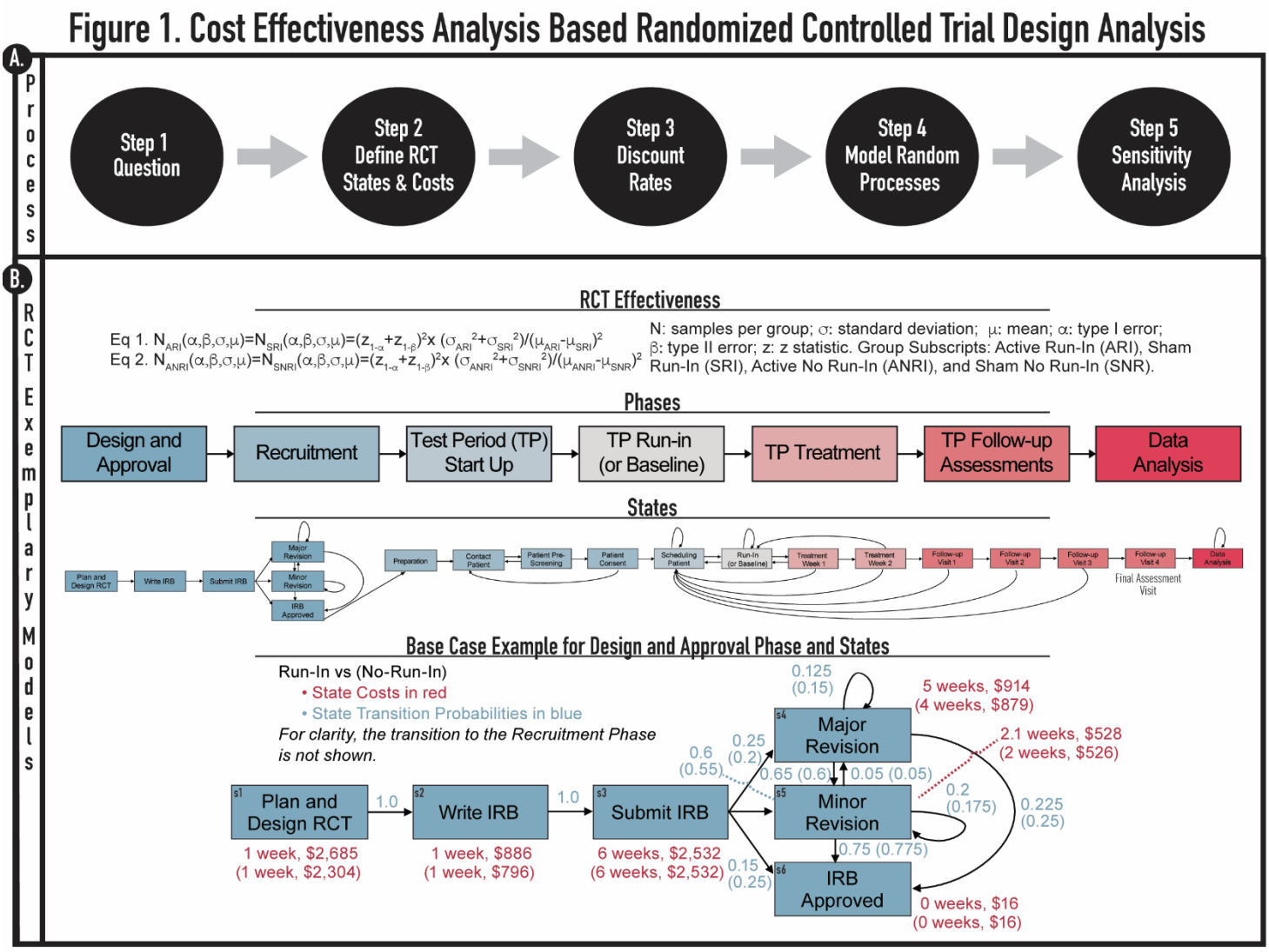


Figure 1. CEA based RCT design analysis example. A. General Process Steps. B. The bottom panel provides examples of the Effectiveness Definition and RCT Designs examined in our CEA analysis. Note, Effectiveness Definitions can be based on any metric(s) that characterizes the trial design. For our specific examples we use the total patient sample size Nt, with the simplifying assumption that patient number needed is the same between the RI and NRI cohorts (i.e., patient therapeutic response characteristics are the same regardless of RCT design). We assess a generalized RCT Design with phases that are composed of a number of states, which are characterized by costs and state transition probabilities that differ between RCTs with/without an RI (NRI case in parentheses). For this example, we enlarge the Design and Approval Phase and States in the bottom row for easier viewing. See Table 1 for complete state list and characteristics.

data obtained from a number of noninvasive brain stimulation (NIBS) Chronic Pain (CP) RCTs and an Opioid Use Disorder RCT which included CP patients (NCT02954432, NCT02330315, NCT02723929, NCT01404052, NCT03625752, NCT04206215, NCT04379115).

METHODS

Our approach informs decision-making on RI suitability in the design of a CP RCT (see **Figure 1.A**). We will detail the methodology for a CEA Based RCT Design, focused on: 1. Defining the research question; 2. Identifying the RCT States and costs; 3. Discounting (cost and effectiveness); 4. Modeling the random nature of the RCT; and 5. Performing a sensitivity analysis. We follow with the analysis of exemplary results based on data from previous NIBS studies. We use this data to exemplify the approach and highlight how RCT design criteria can affect cost effectiveness, but not to make specific indication or therapy recommendations.

CEA Step-1: The first step of our CEA Based RCT Design Analysis is defining the research question. Generating a well-defined research question entail: 1. Considering source material; 2. Identifying the CEA audience; 3. Defining effectiveness; 4. Highlighting potential comparators; 5. Setting the structure/format of results; and 6. Identifying the methods/tools necessary for answering the research question. For our CEA Based RCT Design Analysis examples, we explore the following general research question:

'Will an RCT with an RI period be more cost effective than an RCT with No-Run-In (NRI) period for demonstrating a *statistically significant* and *clinically important difference* in the *Mean Change of a Primary Outcome measured* at a Baseline and Final Assessment visit, between Intervention A and Placebo A on Indication X?' (See **Appendix-Table 1** for further details).

Inherent to the research question is the definition of 'Effectiveness', which is dependent on the intervention being studied, preliminary data sets available, and the RCT goals (i.e., the definition of 'Effectiveness' is unique to each specific RCT design). For our specific CEA examples, we define our metric of 'Effectiveness' as the total patient sample size, N_t (Active + Sham groups), (see **Figure 1B**), necessary to demonstrate a statistically significant and clinically important difference in the primary outcome. Herein, the primary outcome is the mean change in Visual Analogue Scale (VAS) scores of Pain between Active and Sham NIBS treatments in CP patients. As a real-world example, using data from past NIBS studies to solve Eqs.1 and 2 in **Figure 1B**, a sample size of 26 patients (13 Active, 13 Sham) is necessary to demonstrate a statistically significant, clinically important differences between Active and Sham treatments on the VAS endpoint (Salaffi, Stancati, Silvestri, Ciapetti, & Grassi, 2004). For this example, we based the standard deviation of the therapeutic effects on NIBS CP studies (of Osteoarthritis (OA) of the knee (Moreno-Duarte et al., 2013; Wagner & Dipietro, 2018)), a clinically important difference of the VAS OA pain metric of 33% (as defined by (Salaffi et al., 2004) as a "much better" clinically important difference), an 0.80 power, and an alpha of 0.05; we also made the simplifying assumption that patient characteristics in the two cohorts are the same (see **Appendix-Table 2** for additional details).

CEA Step-2: The next step is identifying the fundamental Phases and States of the different RCT designs that will be compared, including the expected time and cost elements for each. First, one should identify the expected Phases of the RCT with and without an RI. For our examples, we break the RCT into the following Phases: 1. Design and Approval; 2. Recruitment; 3. Test Period (including sub-Phases a. Startup/Equipment Purchase, b. NRI Baseline period or RI Assessment period (for these models, we model the RI as a 1 week period designed to exclude patients who don't meet RCT requirements (e.g., compliance with office visit attendance during RI period), see below), c. Treatment, d. Follow-up; and 4. Data Analysis (**Figure 1B**). We then break each Phase down further into individual States to more easily account for granular differences in the different RCT designs (**Figure 1B**). For determining State costs, we employ Levin's ingredients method (Levin & McEwan, 2001). Every RCT element is assumed to have a cost, and all direct and indirect RCT costs should be accounted for. For determining State time durations, one can use data from past-studies or expert opinion (herein, we use the former (see trial list above), which are used to determine time and cost ranges for the CEA assessments (note, all studies and procedures were approved by the Institutional Review Boards at the study sites and written informed consent was obtained from all participants)). Thus, for each RCT design, one should define the individual Phases and States, the costs of each State, and the time necessary to complete each State. See **Figure 1B** and **Table 1** for 'Base-Case' values, developed from past NIBS trials, which serve as the basis of exemplary CEAs explored herein (see **Appendix-Table 3** for a cost worksheet).

Phase		State	State #	Dollar Cost (USD) RI	Dollar Cost (USD) NRI	Time Cost (weeks) RI	Time Cost (weeks) NRI
Design/Approval		Plan and Design RCT	1	$ 3,398.26	$ 3,007.59	1	1
		Write IRB	2	$ 1,241.86	$ 1,147.99	1	1
		Submit IRB	3	$ 2,849.99	$ 2,849.99	6	6
		Minor Revision	4	$ 539.05	$ 535.96	2.1	2
		Major Revision	5	$ 938.80	$ 903.59	5	4
		IRB Approval (Process Approval)	6	$ 336.62	$ 336.62	1/7	1/7
Recruitment		Training/Preparation	7	$ 918.82	$ 918.82	1	1
		Contact Patient	8	$ 2.08	$ 2.08	0.000595	0.000595
		Phone Pre-Screening	9	$ 6.25	$ 6.25	0.00179	0.00179
		In Person Consenting	10	$ 494.00	$ 486.00	0.5	0.5
Test Period	TP: Start Up	Training/Prep Equip	11	$ 28,377.10	$ 28,175.71	1	1
	TP: Baseline	Scheduling	12	$ 112.18	$ 112.18	1	1
		NRI Baseline	13NRI		$ 424.97		1
		Run-In	13RI	$ 868.73		1	
	TP: Treat-ment	Treatment Week 1	14	$ 1,562.68	$ 1,562.68	1	1
		Treatment Week 2	15	$ 1,498.54	$ 1,498.54	1	1
	TP: Follow-up	Follow-up 1	16	$ 375.62	$ 375.62	1	1
		Follow-up 2	17	$ 375.62	$ 375.62	1	1
		Follow-up 3	18	$ 375.62	$ 375.62	2	2
		Follow-up 4	19	$ 468.15	$ 468.15	2	2
Data analysis		Data-Analysis	20	$ 14,468.53	$ 13,990.79	1	1

Table 1. 'Base-Case' Model States and Costs: Herein, we list each of the States for the No-Run-In RCT design (NRI) and the Run-In RCT design (RI), and the total cost per State in monetary and time units (see Appendix-Table 3 for further details)("Note, due to the methodological nature of this paper, we do not provide institutional specific salaries, overhead, or cost values, but instead we use averages across multiple institutions and/or published rates from the NIH where available, including: https://www.opm.gov/policy-data-oversight/pay-leave/pay-administration/fact-sheets/computing-hourly-rates-of-pay-using-the-2087-hour-divisor;https://grants.nih.gov/grants/guide/notice-files/NOT-OD-19-036.html; https://grants.nih.gov/grants/policy/salcap_summary.htm,")

CEA Step-3: The next step is determining the Discount rates for Cost and Effectiveness. Discounted Cost (DC) and Discounted Effectiveness (DE) are based on the concept that a dollar and effectiveness as utility, respectively, are worth more today than they are in the future, and are given by the following equations:

$$DE=\frac{E}{(1+r_d)^T} \quad \text{and} \quad DC=\sum_{t=1}^{t=T}\frac{C_t}{(1+i_d)^t}$$

where i_d is the cost discount rate, r_d is the effectiveness discount rate, t is the time, C is cost, and E is the measure of effectiveness (herein defined as N_t).

CEA studies focused on RCT trial health effect outcomes are typically discounted between ~1.5% to 5% for effectiveness and ~3 to 6% for costs (Gravelle & Smith, 2001; Organization, 2003; Treasury). However, there are no standard and accepted values for discounts for CEA based RCT Design Analyses. Thus, for our specific real-world examples, we evaluate r_d of 5% and i_d of 6% for the 'Base-Case' CEA assessments (but vary them during sensitivity analysis- see below).

CEA Step-4: Next, in order to conduct an effective analysis of the total RCT costs the random nature of a clinical trial needs to be accounted for. The cost ingredient method described above, with the predicted flow from individual States, provides the foundation for using Markov Models to account for the State flow of the RCT (Norris, 1997). Markov models are stochastic processes that undergo transitions from one State (X_n=i) to another State (X_{n+1}=j) and are characterized by the property that probability P_{ij} to transition from State i to State j is equal to $P(X_{n+1}=j|X_n=i,X_{n-1}=,\ldots,X_o=i_o)= P(X_{n+1}=j|X_n=i)$, or simply stated that the past and future are conditionally

independent given the present. A State transition probability matrix,

$$P=\begin{bmatrix} P00 & \cdots & P0n \\ \vdots & \ddots & \vdots \\ Pn0 & \cdots & Pnn \end{bmatrix}$$

serves as the basis for modeling the State flow of our RCT models (Norris, 1997). The values of the transition probabilities can be directly gathered from study data or modeled based on past studies and/or expert opinion (for our models, we use data from past RCTs with and without RIs (see **Figure 1** and **Appendix-Table 4**).

Next, in order to simulate the State flow, we implement a Monte Carlo Simulation (MCS) method. MCSs are a class of computational algorithms that can be used to analyze stochastic systems to establish the odds for a variety of outcomes. MCS typically involves 3 steps: 1) Randomly generate M inputs (or scenarios); 2) Run a simulation N times for each scenario on an RCT model being analyzed (herein, we implemented N=1000); and 3) Collect and analyze the simulation outputs (For further details the reader is referred to Raffa (2016)).

Specifically, for each of the States in the example models, we have defined a cost and time (**Table 1**) and transition probabilities to move from the State to another State (**Figure 1** and **Appendix-Table 4**). This allows one to conduct an MCS and determine the distributions of the State terms (e.g., cost, days, number of steps through the model, etc.), from which descriptive statistics can be developed (e.g., Mean Discounted Cost (MDC) and Mean Discounted Effectiveness (MDE)). While we focused on statistics related to the mean, other measures of central tendency (e.g., median) or variability (e.g., variance, entropy) could also be used to determine limits on cost-effectiveness (see below).

Finally, MCS allows one to explore RCT design criteria by building upon the fundamental model scenarios (**Figure 2**). In our CEA Based RCT Design Analysis, we assess varied state costs, state transition probabilities, number of institutions, patient evaluation capacity, and patient consent capacity. For our 'Base-Case' we model costs (**Table 1**), P_{ij}'s (**Appendix-Table 4**), discounts (i_d=6%, r_d=5%), and 3 patients simultaneously evaluated at one institution with up to 7 consented per week (modeled directly from past NIBS studies of comparable size which were used to define our CEA examples). We vary the range of these design parameters as part of the sensitivity analysis (see below).

CEA Step-5: The final step is conducting a sensitivity analysis. Sensitivity analyses evaluate how changes in model inputs affect model outputs. To exemplify our CEA methodology, we implement a deterministic sensitivity analysis focused on criteria often considered during the RCT design process. Specifically, we investigated the RCT design impact on RI and NRI cost-effectiveness by varying the:

- number of institutions conducting the trial (varied from 1 to 4),
- number of patients that can be evaluated simultaneously at an institution (varied from 1 to 5),
- number of patients that can be consented weekly at an institution (varied from 1 to 7),
- discount rates (r_d varied from 0 to 50% in 5% steps and id from 0 to 10% in 1% steps),
- costs (0.5-3x in 0.5 steps for different states), and
- State transition probabilities (varying $P(X_{n+1}=j|X_n=i)$ by 25% relative step changes- for example, if there was a 'Base-Case' 75% transition likelihood for Treatment-Week State ➡ Follow-up State, we would investigate 75%, 56.25%, 37.5% and 18.75% transition probabilities with complementary changes to the other possible States (e.g., Scheduling or Run-In States)).

In addition to assessing a sample size N=26 which defines the 'Base-Case' Model (see Step 1) we also assess N=1 and N=104 (which represents a sample size to necessary to demonstrate a minimal clinically important difference in VAS Pain scores, of 15%, between Active and SHAM conditions (Salaffi et al., 2004) in the primary endpoint of the research question, solved per Eqs. 1-2, as detailed in Step-1) (See **Appendix-Table 2** for further details).

Finally, we calculate the:

- Incremental Cost Effectiveness Ratio (ICER), equal to $(MDC_{RI}-MDC_{NRI})/(MDE_{RI}-MDE_{NRI})$, and
- Average Cost Effectiveness Ratio (ACER), equal to MDC/MDE,
- for the different RCT designs we are assessing.

Although ICERs are conventionally recommended in CEA studies analyzing health outcomes (Briggs & Fenn, 1997; Hoch & Dewa, 2008), ICER instability can result when compared Effectiveness metrics are small (as would be expected for ours based on equivalent Nt's) (Bang & Zhao, 2012, 2014; Organization, 2003). Thus, we focus on the ACERs, and per our research question based on an all-or-none question of cost effectiveness, we also assess the difference in average cost effectiveness ratios for the RI and NRI RCT Designs (Laska, Meisner, & Siegel, 1997). We compared $ACER_{RI}$ to $ACER_{NRI}$ to determine cost-effectiveness per the ACERs (where an

Figure 2. Markov/MCS Methods Analysis Design and Approval Phase For An RCT With No-Run-In Example

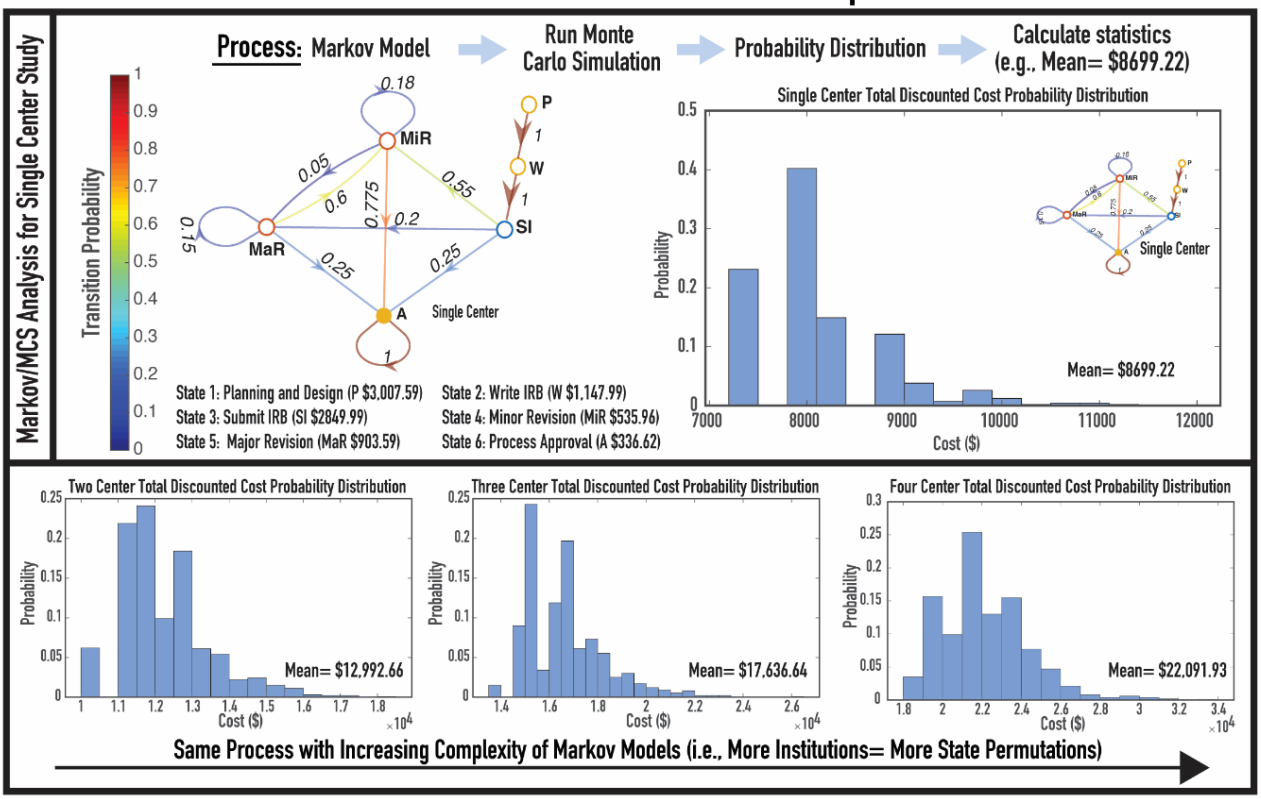


Figure 2. Markov Models, coupled with MCS, allow one to easily develop and assess complicated RCT designs based on the initial State building blocks. Focusing on just the Planning and Approval Phase of the RCT design, one can see how quickly the complexity increases from a 'Single-Center Study' to a 'Four-Center-Study' design. The methodology we present allows one to build up and assess increasingly complicated permutations of RCT designs, and systematically evaluate the cost effectiveness criteria, such as for example the total mean discounted cost for each scenario.

$ACER_{RI}<ACER_{NRI}$ would indicate RI cost-effectiveness per the ACERs). We also assess ICER cost-effectiveness decision criteria per (Briggs & Fenn, 1997), similar to what would be determined by CEA plane analysis (Bang & Zhao, 2014; Briggs & Fenn, 1997). Comparisons are made between the resulting ACER and ICER decision criteria.

RESULTS

Below we compare the CEA results from the 'Base-Case' model and those characterized by the range of variables defined above.

For the 'Base-Case' exemplary model, we found that the NRI Design was more cost effective than the RI Design per the ACER analysis (see **Figure 3A**). The sensitivity analysis demonstrated that the ACERs for the RI designs remained consistently greater than that of the NRI designs when varying the number of institutions, number of patients that could be consented per week/institution, number of patients that could be assessed simultaneously/institution, and/or cost discounts (with the other variables fixed at the 'Base-Case' values) (for example see **Figure 3B**, **Appendix-Table 5**, Appendix-Figure 1A). Similarly, ICER analysis comparing the NRI and RI cost effectiveness, as dictated by (Briggs & Fenn, 1997), would outright reject an NRI RCT Design or represent a trade-off condition in cost-effectiveness, with minimal increases of effectiveness at high costs. Although the above variables demonstrate a limited impact on the RI vs NRI design decisions for these exemplary models, they still have an impact on the overall RCT design cost effectiveness (See **Figure 3.C** and **Appendix-Figure 1B** where we compare the cost effectiveness of each NRI and RI RCT design to the NRI 'Base Case' (versus directly comparing each RI to NRI RCT designs for each specific design criteria as in the other results)).

The Average and Incremental Cost-Effectiveness Ratios for the different RI vs NRI designs showed their greatest variation as functions of the Phase Costs and

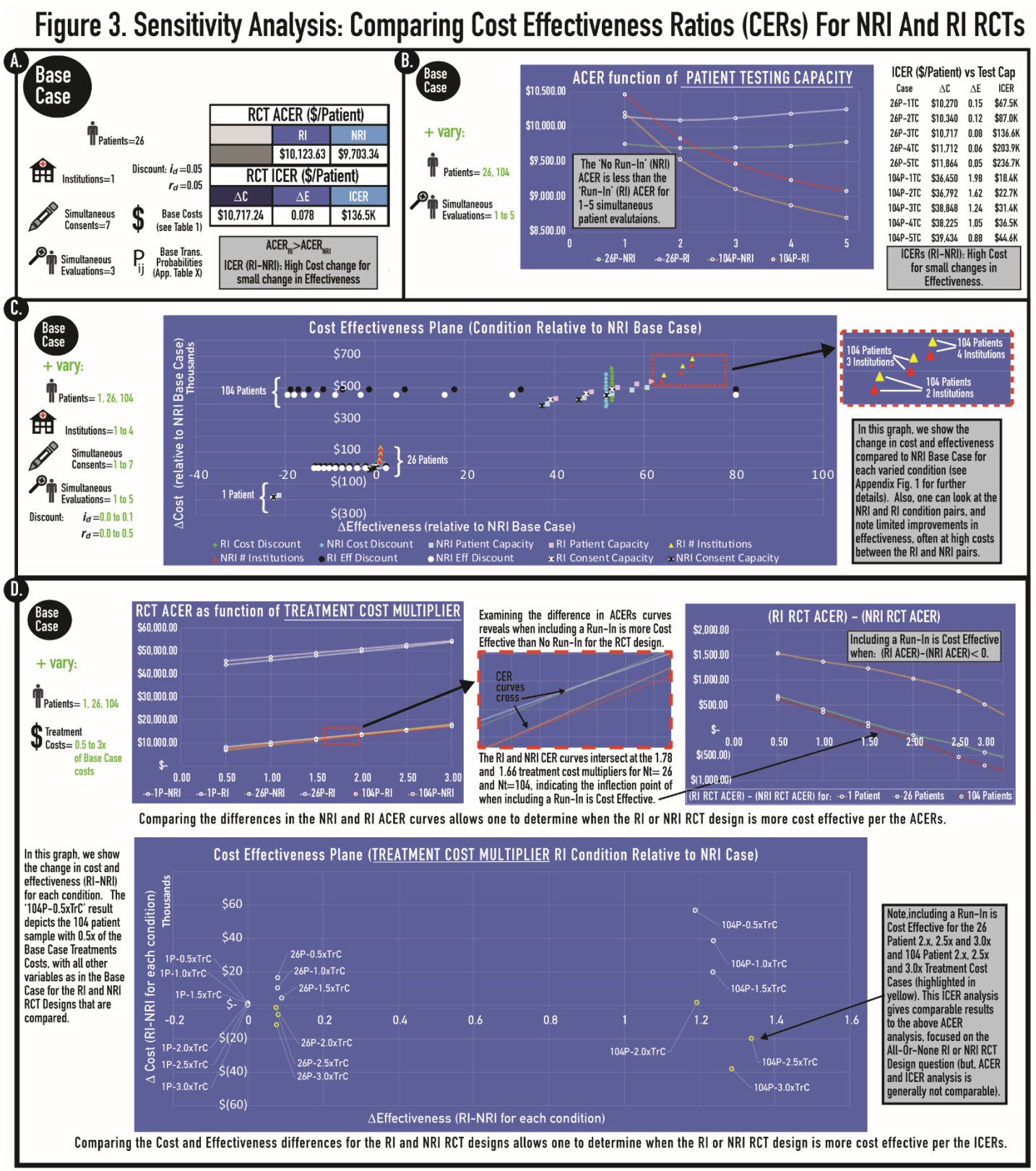


Figure 3. A. 'Base-Case' NRI and RI CERs **B.** CERs as a function of simultaneous patient testing capacity for the 'Base-Case' and N=104 models (with all other variables set to those levels of 'Base-Case', note 26P-2TC indicates 26 patients, 2 patient testing capacity). **C.** CE Plane comparing varied RCT Design Criteria relative to the NRI 'Base-Case'. **D.** CERs as a function of Treatment Phase Costs for N=1, 26, and 104 models.

State Transition probabilities, particularly when analyzing the Test Period (TP) states (while keeping other variables fixed as defined by the 'Base-Case' unless otherwise noted).

When varying the NRI Baseline and RI costs (states 13NRI and 13RI, see **Table 4**), but keeping their relative relationship equal, we found that the RI ACER was less than the NRI design for N=26 and N=104 when the Baseline and RI costs were approximately 0.6 and 0.63 times those of the 'Base-Case' costs, respectively. Furthermore, the ICERs for both indicated that the RI design was Cost-Effective when the Baseline/Run-In costs were approximately 0.5x the 'Base-Case' costs. When varying the RI and NRI Treatment costs (states 14 and 15, see **Table 4**), but keeping their relative relationship equal, we found that the RI RCT design was more cost effective than the NRI design for N=26, and N=104 when the Treatment costs were approximately 1.78 and 1.66 times greater respectively than those of the 'Base-Case' per the ACER analysis, which was comparable to the ICER analysis (see **Figure 3D**). When varying the RI and NRI Follow-Up costs (states 16-19) we found that including an RI was not more cost effective as the follow-up costs varied from 0.5-3.0x 'Base-Case' costs; although adding an RI lead towards cost-effectiveness as a function of increased follow-up costs (see **Appendix Figure 2**). When varying the RI and NRI costs for the other states (while keeping the other values fixed), the RI designs only became cost effective compared to the NRI designs when the ratio of NRI/RI Phase costs was artificially skewed so the burden of including an RI was ignored (e.g., assuming Recruiting RI patients cost 0.5 that of NRI patients). Generally, when examining multiple cost Phase changes simultaneously (and/or with changes in discount rates, transition rates, etc.), the above effects compound themselves (e.g., decreasing the Baseline costs while increasing Treatment costs increases the overall cost effectiveness of the RI vs NRI design).

While varying the NRI and RI State transition probabilities, we demonstrated that the RI design is more cost effective than the NRI design as function of decreased likelihood of transitioning to later RCT States before reaching the final follow-up (while keeping the RI transition probabilities constant), and the complementary effect while varying the NRI transition probabilities (**Figure 4**) (Note: figure focuses on just the TP for clarity). For these exemplary results, the RCT design was more cost-effective than the NRI design when the likelihood of transitioning to the next State decreases by just 25% of the 'Base-Case' NRI design (and becomes more cost-effective with later State changes (e.g., State 16 vs 15) and increased probability changes). The situation becomes increasingly complicated as one begins varying multiple State transition probabilities and other sensitivity criteria, such that the above effects generally compound themselves (see **Appendix-Figure 3**). The transition probability data demonstrates that an increased likelihood of patient dropout necessitates RI inclusion to be cost-effective.

Finally, when varying the effectiveness discount, at N_t=104 the RCT RI design became cost-effective per the ACER analysis when r_d reached ~16% (with lowest ICER at r_d=15%) (see below and **Appendix-Table 5F**).

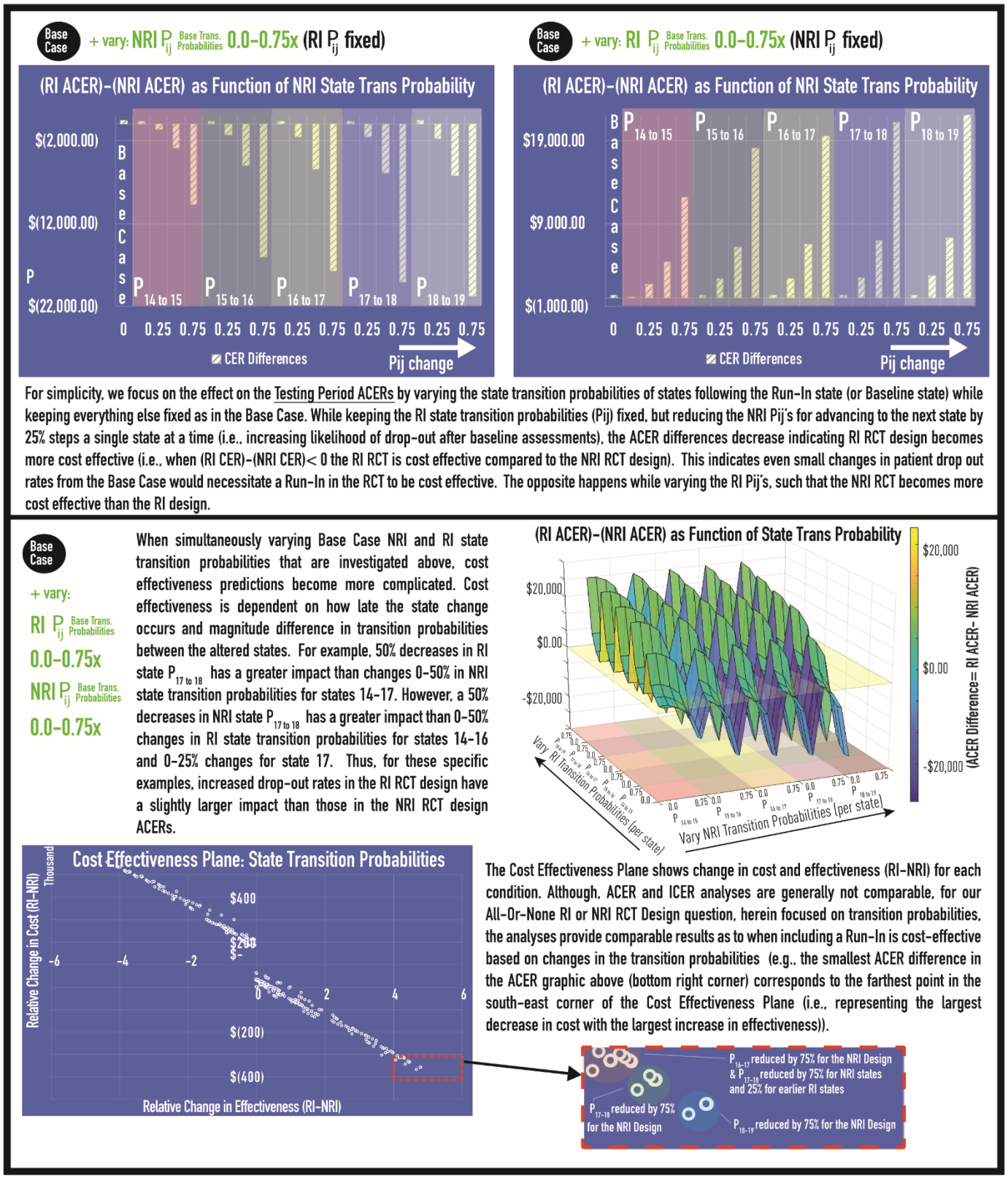
DISCUSSION

Generally, healthcare CEAs are used to compare the health effects of different treatments determined through RCTs. Herein, we demonstrated a method to use CEAs to assess the Design of such RCTs. We presented a step-by-step guide for conducting a CEA for including/excluding an RI in a CP RCT with the goal of assisting researchers with their own RCT designs. The framework allows one to objectively quantify the impact of RCT design variables on the different phases of an RCT trial and on the overall cost-effectiveness of the trial.

Past studies have explored aspects of the CEA based RCT design methodologies we presented herein, but have not provided a method to fully optimize RCT design and/or determine RI need based on RCT cost effectiveness (Brittain & Wittes, 1990; Eisenstein et al., 2008; Huynh et al., 2014; Schechtman & Gordon, 1993; Schroy et al., 2009). For example, Schroy (Schroy et al., 2009) and Huynh (Huynh et al., 2014) utilized CEA design methodologies to analyze RCT recruitment strategies, and thus did not consider RIs or full trial optimization. While Eisenstein (Eisenstein et al., 2008) assessed the impact of cost reductions across multiple design criteria for theoretical RCTs, Eisenstein did not explore RIs or RCT cost-effectiveness. The works of Brittain (Brittain & Wittes, 1990) and Schechtman (Schechtman & Gordon, 1993) most closely resemble this study, particularly in terms of study goals. They assessed the Cost Effectiveness of RI periods, but based on computational "analogues" (e.g., "the randomized and eligible subjects ratio" (Schechtman & Gordon, 1993)).

While our findings confirmed a number of their general results (e.g., RIs are likely to be cost-effective when: "per patient costs during the post-randomization as compared to the screening period are high" (Schechtman & Gordon, 1993)), their computational methods did not allow for trial design optimization due to the limited number of variables assessed via their equation based 'analogue' methodology.

Figure 4. CERs As Function Of State Transition Probability



For simplicity, we focus on the effect on the Testing Period ACERs by varying the state transition probabilities of states following the Run-In state (or Baseline state) while keeping everything else fixed as in the Base Case. While keeping the RI state transition probabilities (Pij) fixed, but reducing the NRI Pij's for advancing to the next state by 25% steps a single state at a time (i.e., increasing likelihood of drop-out after baseline assessments), the ACER differences decrease indicating RI RCT design becomes more cost effective (i.e., when (RI CER)-(NRI CER)< 0 the RI RCT is cost effective compared to the NRI RCT design). This indicates even small changes in patient drop out rates from the Base Case would necessitate a Run-In in the RCT to be cost effective. The opposite happens while varying the RI Pij's, such that the NRI RCT becomes more cost effective than the RI design.

Base Case

+ vary:

RI P_{ij} Base Trans. Probabilities

0.0–0.75x

NRI P_{ij} Base Trans. Probabilities

0.0–0.75x

When simultaneously varying Base Case NRI and RI state transition probabilities that are investigated above, cost effectiveness predictions become more complicated. Cost effectiveness is dependent on how late the state change occurs and magnitude difference in transition probabilities between the altered states. For example, 50% decreases in RI state $P_{17\ to\ 18}$ has a greater impact than changes 0-50% in NRI state transition probabilities for states 14-17. However, a 50% decreases in NRI state $P_{17\ to\ 18}$ has a greater impact than 0-50% changes in RI state transition probabilities for states 14-16 and 0-25% changes for state 17. Thus, for these specific examples, increased drop-out rates in the RI RCT design have a slightly larger impact than those in the NRI RCT design ACERs.

The Cost Effectiveness Plane shows change in cost and effectiveness (RI-NRI) for each condition. Although, ACER and ICER analyses are generally not comparable, for our All-Or-None RI or NRI RCT Design question, herein focused on transition probabilities, the analyses provide comparable results as to when including a Run-In is cost-effective based on changes in the transition probabilities (e.g., the smallest ACER difference in the ACER graphic above (bottom right corner) corresponds to the farthest point in the south-east corner of the Cost Effectiveness Plane (i.e., representing the largest decrease in cost with the largest increase in effectiveness)).

Figure 4. CERs as function of state transition probability

Thus, expanding upon these past works, we presented an expanded CEA based RCT design method which can determine RI cost effectiveness based on any quantifiable trial element.

We demonstrated the real-world impact of specific RCT design variables (e.g., institutional capacity, treatment costs, sample sizes, and trial length) on the inclusion/exclusion of an RI in RCT Designs developed from past NIBS trials. For example, when the 'Base-Case' models' treatment period costs were increased by 78-100%, an RI became necessary for the RCT to be cost-effective (per the ACER and ICER analysis); thus, more expensive and/or prolonged treatments would necessitate an RI in the RCT design. As another example, when the 'Base-Case' model patient drop-out rate increased by 25%, the need for an RI became necessary for the RCT to be cost effective. More challenging patient populations, common to Pain trials, would thus necessitate an RI. However, if the RCT included an enhanced adherence plan (such as offer commuting costs/parking, flexibility with scheduling, reduced time commitment) resulting in a large adherence to the study, the RI may not be cost effective (note, our data was developed from studies that maintained such measures; but as demonstrated through the modeling, benefits can quickly wane dependent on patient characteristics and trial size). Even in cases where including an RI was not cost effective, the CEA Based RCT Design methods demonstrated how trial variables can be optimized. For example, one can see that increasing the number of institutions improves effectiveness in the N_t=104 vs N_t= 26 cases, with slight increases costs for both sample sizes (in both the NRI and RI conditions). This represents a design trade-off decision between increased personnel/facility costs, decreased trial duration, and increased effectiveness.

Although each RCT should be tailored and optimized for patient type and treatment modality, we anticipate that a number of results will generalize across most Pain RCTs. For example, as above, RIs can improve cost effectiveness in RCTs with patient populations with lower adherence and in larger/longer trials where greater resources would be expended in patients that are likely to not complete the trial. There are also a number of quantitative trends that should be noted. For our examples, as C_{RI}>C_{NRI}, when the RI RCT duration is shorter than the NRI RCT duration (seen with the increased sample sizes) and the effectiveness discount is raised then $ACER_{NRI}$ > $ACER_{RI}$, but as incremental differences in effectiveness become lower the ICERs can become quite high (**Appendix-Table 5F**). Ultimately though, what a researcher models for the value of r_d inherently depends on the researcher's RCT goal (if a researcher was evaluating an RCT of a life-saving treatment for immediate use, r_d could be exceedingly high), but r_d will not be an RCT design criteria that can be altered by the researcher for cost effectiveness optimization.

Our analysis has some limitations. First, we assumed that the patient characteristics in the two cohorts were the same, and that including an RI had no impact on patient response to treatment (i.e., we did not consider the fact that patients who were randomized after the Run-In period may have exhibited different clinical characteristics from those who initially enrolled (Pablos-Méndez et al., 1998)). Second, for the models we analyzed a generic 1-week RI design, which would exclude patients with poor compliance, although different RI designs and inclusion/exclusion criteria should be differentially accounted for in future studies. Additionally, as the CEA analysis was based on an RCT design which 'demonstrated statistically significant and clinically important difference in the Mean Change of a Primary Outcome measured at a Baseline and Final Assessment visit', the effectiveness definition does not fully account for partly- or non-compliant patients (which could also require intention-to-treat analysis in the RCT design). Furthermore, sensitivity analysis was confined to a deterministic parameter analysis, neither probabilistic parameter assessments or model uncertainty were assessed herein (nor did we include the potential for intention-to-treat analysis in the RCT design) (Cost-Effectiveness in Health and Medicine, 2016). Therefore, our analysis entails simplifying assumptions around the patient characteristics, effectiveness considerations, and sensitivity analysis that should be addressed in future work. Finally, future studies should explore hybrid CEA models that assesses the cost effectiveness of the RCT trial results and RCT trial design simultaneously, and further incorporate methods of statistical inference into the methodology. Analyses focused on assessing potentially skewed data sets, such as comparing incremental cost effectiveness ratios based both on the means and medians of effectiveness and cost distributions, should also be explored in future studies (Bang & Zhao, 2014).

CONCLUSION

To the best of our knowledge, this is one of the first in-depth analyses of the cost-effectiveness of the inclusion of an RI period for CP RCTs. Using the methods outlined herein, researchers can design cost-effective RCTs with/without an RI based on an objective assessment of RCT design variables. Furthermore, while we focused on questions related to RCT RI periods, our CEA based RCT

design methods can be applied for the optimization of an RCT design for any indication and/or intervention. Ultimately the methods outlined herein will allow researchers to maximize resource use for specific trials, reduce waste, and/or free up resources for future studies.

Grant support

Research reported in this publication, funding the clinical trials that informed the datasets of the CEA models and/or modeling of patient behaviors, were supported in part by the National Institute of Health NIA (Award Number R44AG055360), NIDDK (Award Number DK117710), NIDA (Award Number R44DA049685), NIAMS (Award Number R44AR076885), and NCCIH (Award Number AT008637A). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Trial registration number & trial register

NCT02954432, NCT02330315, NCT02723929, NCT01404052, NCT03625752, DA049685 and AR076885.

Author Affiliations

1 Neuromodulation Center and Center for Clinical Research Learning, Spaulding Rehabilitation Hospital and Massachusetts General Hospital, Charlestown, MA, USA.
2 Harvard Medical School, Boston, MA, USA
3 Highland Instruments, Inc. Cambridge, MA, USA
4 Harvard T.H. Chan School of Public Health, Boston, MA, USA
5 MGH Institute of Health Professions, Boston, MA, USA
6 Instituto Wilson Mello, Campinas, Sao Paulo, Brazil
7 Case Western Reserve University, Cleveland, Ohio
8 Boston University, Boston, MA, USA
9 Harvard-MIT Division of Health Sciences and Technology, Cambridge, MA, USA
10 Department of Anesthesiology, Perioperative and Pain Medicine, Brigham & Women's Hospital, Boston, MA.
a HR and ER served as co-first authors.
b FF, LD, and TW served as co-last authors sharing senior author responsibilities.

Conflict of Interest Statement

Dr. Timothy Wagner and Dr. Laura Dipietro are officers at Highland Instruments, a medical device company. They have patents pending or issued, personally or as officers in the company, related to imaging, brain stimulation, diagnostics, modeling, and simulation. FF is the editor-in-chief of the Principles and Practice of Clinical Research journal, and PG is part of the editorial team. Therefore, they excused themselves from the peer-review process and followed the journal guidelines for peer-reviewing when an editor co-authors a manuscript. They did not influence the editorial process and final publication decision.

REFERENCES

Bang, H., & Zhao, H. (2012). Average cost-effectiveness ratio with censored data. *J Biopharm Stat, 22*(2), 401-415. doi:10.1080/10543406.2010.544437

Bang, H., & Zhao, H. (2014). Cost-effectiveness analysis: a proposal of new reporting standards in statistical analysis. *J Biopharm Stat, 24*(2), 443-460. doi:10.1080/10543406.2013.860157

Briggs, A., & Fenn, P. (1997). Trying to do better than average: a commentary on 'statistical inference for cost-effectiveness ratios'. *Health Econ, 6*(5), 491-495. doi:10.1002/(sici)1099-1050(199709)6:5<491::aid-hec293>3.0.co;2-r

Brittain, E., & Wittes, J. (1990). The run-in period in clinical trials. The effect of misclassification on efficiency. *Control Clin Trials, 11*(5), 327-338. doi:10.1016/0197-2456(90)90174-z

Cipriani, A., & Geddes, J. R. (2010). What is a run-in phase? *Epidemiol Psichiatr Soc, 19*(1), 21-22. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/20486419

Cost-Effectiveness in Health and Medicine. (2016). (P. J. Neumann, G. D. Sanders, L. B. Russell, J. E. Siegel, & T. G. Ganiats Eds. 2nd ed.): Oxford University Press.

Dworkin, R. H., Turk, D. C., Peirce-Sandner, S., Baron, R., Bellamy, N., Burke, L. B., . . . Witter, J. (2010). Research design considerations for confirmatory chronic pain clinical trials: IMMPACT recommendations. *Pain, 149*(2), 177-193. doi:10.1016/j.pain.2010.02.018

Dworkin, R. H., Turk, D. C., Peirce-Sandner, S., Burke, L. B., Farrar, J. T., Gilron, I., . . . Ziegler, D. (2012). Considerations for improving assay sensitivity in chronic pain clinical trials: IMMPACT recommendations. *Pain, 153*(6), 1148-1158. doi:10.1016/j.pain.2012.03.003

Eisenstein, E. L., Collins, R., Cracknell, B. S., Podesta, O., Reid, E. D., Sandercock, P., . . . Diaz, R. (2008). Sensible approaches for reducing clinical trial costs. *Clin Trials, 5*(1), 75-84. doi:10.1177/1740774507087551

Gewandter, J. S., Dworkin, R. H., Turk, D. C., McDermott, M. P., Baron, R., Gastonguay, M. R., . . . White, R. E. (2014). Research designs for proof-of-concept chronic pain clinical trials: IMMPACT recommendations. *Pain, 155*(9), 1683-1695. doi:10.1016/j.pain.2014.05.025

Gravelle, H., & Smith, D. (2001). Discounting for health effects in cost–benefit and cost-effectiveness analysis. In *Health Economics* (Vol. 10, pp. 587-599).

Hattori, N., Takeda, A., Takeda, S., Nishimura, A., Kitagawa, T., Mochizuki, H., . . . Takahashi, R. (2019). Rasagiline monotherapy in early Parkinson's disease: A phase 3, randomized study in Japan. *Parkinsonism Relat Disord, 60*, 146-152. doi:10.1016/j.parkreldis.2018.08.024

Hewitt, D. J., Ho, T. W., Galer, B., Backonja, M., Markovitz, P., Gammaitoni, A., . . . Wang, H. (2011). Impact of responder definition on the enriched enrollment randomized withdrawal trial design for establishing proof of concept in neuropathic pain. *Pain, 152*(3), 514-521. doi:10.1016/j.pain.2010.10.050

Hoch, J. S., & Dewa, C. S. (2008). A clinician's guide to correct cost-effectiveness analysis: think incremental not average. *Can J Psychiatry, 53*(4), 267-274. doi:10.1177/070674370805300408

Huynh, L., Johns, B., Liu, S. H., Vedula, S. S., Li, T., & Puhan, M. A. (2014). Cost-effectiveness of health research study participant recruitment strategies: a systematic review. *Clin Trials, 11*(5), 576-583. doi:10.1177/1740774514540371

Kim, M. K., Lee, S. Y., Park, H. S., Yoon, H. J., Kim, S. H., Cho, Y. J., . . . Park, C. S. (2018). A Randomized, Multicenter, Double-blind, Phase III Study to Evaluate the Efficacy on Allergic Rhinitis and Safety of a Combination Therapy of Montelukast and Levocetirizine in Patients With Asthma and Allergic Rhinitis. *Clin Ther, 40*(7), 1096-1107.e1091. doi:10.1016/j.clinthera.2018.04.021

Laska, E. M., Meisner, M., & Siegel, C. (1997). The usefulness of average cost-effective ratios. *Health Econ, 6*(5), 497-504. doi:10.1002/(sici)1099-1050(199709)6:5<497::aid-hec298>3.0.co;2-v

Levin, H. M., & McEwan, P. J. (2001). Cost-Effectiveness Analysis: Methods and Applications. In (2nd Edition ed.): Sage Publications Inc.

Moreno-Duarte, I., Diaz-Cruz, C., Doruk, D., Coutinho, L., Luque, L., Dipietro, L., . . . Fregni, F. (2013, September 16-18, 2013). *Effects of Electrosonic Stimulation on the Perception of Chronic Pain Due to Osteoarthritis of the Knee.* Paper presented at the 5th International Symposium on Neuromodulation, Sao Paulo, Brazil.

Norris, J. R. (1997). *Markov Chains (Cambridge Series in Statistical and Probabilistic Mathematics)*: Cambridge University Press.

Organization, W. H. (2003). *WHO GUIDE TO COST-EFFECTIVENESS ANALYSIS*. Retrieved from Geneva, Switzerland:

Pablos-Méndez, A., Barr, R. G., & Shea, S. (1998). Run-in periods in randomized trials: implications for the application of results in clinical practice. *JAMA, 279*(3), 222-225. doi:10.1001/jama.279.3.222

Raffa, M. K. a. J. (2016). Markov Models and Cost Effectiveness Analysis: Applications in Medical Research. In *MIT Critical Data, Secondary Analysis of Electronic Health Records,* (pp. 428).

Salaffi, F., Stancati, A., Silvestri, C. A., Ciapetti, A., & Grassi, W. (2004). Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. *Eur J Pain, 8*(4), 283-291. doi:10.1016/j.ejpain.2003.09.004

Schechtman, K. B., & Gordon, M. E. (1993). A comprehensive algorithm for determining whether a run-in strategy will be a cost-effective design modification in a randomized clinical trial. *Stat Med, 12*(2), 111-128. doi:10.1002/sim.4780120204

Schroy, P. C., Glick, J. T., Robinson, P., Lydotes, M. A., Heeren, T. C., Prout, M., . . . Wong, J. B. (2009). A cost-effectiveness analysis of subject recruitment strategies in the HIPAA era: results from a colorectal cancer screening adherence trial. *Clin Trials, 6*(6), 597-609. doi:10.1177/1740774509346703

Sox , H., & Higgins, M., Douglas. (2013). Cost-effectiveness analysis and cost–benefit analysis. In *Medical Decision Making* (Second Edition ed.): John Wiley & Sons, Ltd.

Steiner, D. J., Sitar, S., Wen, W., Sawyerr, G., Munera, C., Ripa, S. R., & Landau, C. (2011). Efficacy and safety of the seven-day buprenorphine transdermal system in opioid-naïve patients with moderate to severe chronic low back pain: an enriched, randomized, double-blind, placebo-controlled study. *J Pain Symptom Manage, 42*(6), 903-917. doi:10.1016/j.jpainsymman.2011.04.006

Treasury, H. *The Green Book: central government guidance on appraisal and evaluation* Retrieved from

Wagner, T., & Dipietro, L. (2018). Novel Methods of Transcranial Stimulation: Electrosonic Stimulation. In *Neuromodulation: Comprehensive Textbook of Principles, Technologies, and Therapies* (pp. 1619-1626): Elsevier.

Yarlas, A., Miller, K., Wen, W., Lynch, S. Y., Munera, C., Pergolizzi, J. V., . . . Ripa, S. R. (2015). Buprenorphine transdermal system compared with placebo reduces interference in functioning for chronic low back pain. *Postgrad Med, 127*(1), 38-45. doi:10.1080/00325481.2014.992715