



Design and Statistical Considerations for Cluster Randomized Trials in Community Settings: Lessons From Across the NCI Community Oncology Research Program (NCORP)

Anna C. Snavely, PhD and Kathryn E. Weaver, PhD, MPH



A program of the National Cancer Institute
of the National Institutes of Health



Wake Forest University
School of Medicine



Atrium Health
Wake Forest Baptist

Disclosures

No relevant disclosures

Agenda

- Introduction
 - NCORP and Cancer Care Delivery Research (CCDR) Overview
 - Cluster Randomized Trials in Community Settings
- Case Studies
 - Case Study 1: Stephanie Pugh, PhD (NRG NCORP)
 - Case Study 2: Heather Gunn, PhD (Alliance NCORP)
 - Case Study 3: Emily Dressler, PhD (Wake Forest NCORP)
- Panel Discussion
 - Led by Kathryn Weaver, PhD, MPH (mPI Wake Forest NCORP)

NCORP Overall Goal

To bring cancer clinical trials and cancer care delivery research (CCDR) to people in their own communities, thereby generating a broadly applicable evidence base that contributes to improved patient outcomes and a reduction in cancer disparities.

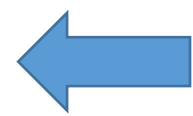
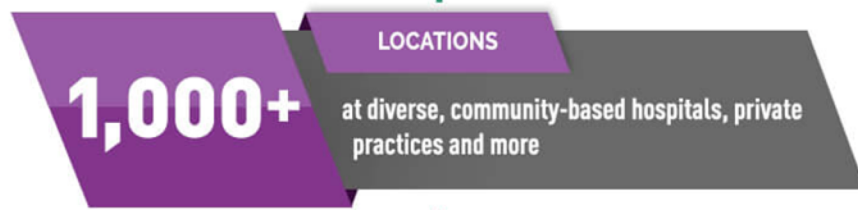
The **NCI Community Oncology Research Program (NCORP)** brings cancer research studies and results to patients in a variety of community settings across the United States.



for



including



Advancing Research. Improving Lives.™



NCORP Research Agenda

Type	Focus
Cancer Prevention	Identify/evaluate interventions to reduce cancer risk and incidence
Cancer Control	Reduce incidence/co-morbidity of cancer and its treatment, enhance quality of life
Cancer Care Delivery Research	Intervene on patient, clinician, and organizational factors that influence care delivery

Cancer Care Delivery Research (CCDR)

- Multidisciplinary science that seeks to improve clinical outcomes and patient well-being by intervening on patient, clinician, and organizational factors that influence care delivery
- Involves multiple stakeholders/perspectives
 - Practices/providers
 - Patients
 - Caregivers
- Examines how care is currently and/or should be delivered
- Interventions are multi-level (not typically exclusively directed at patients)
- Evolution of Cancer Care Delivery Research in the NCI Community Oncology Research Program (Geiger AM et al. J Natl Cancer Inst, June 2020)

CCDR Connections with Implementation Science

- Integration of evidence-based practices lags behind
 - Screening tests not consistently ordered for the right people
 - Over, under, and non-recommended treatment for cancer
 - Poor symptom/late effects management (under and over treatment)
 - Underuse of evidence-based supportive care
 - People at high risk (familial genetic syndromes) not identified and screened
- Interest in implementation and de-implementation
- Pragmatic and comparative effectiveness trials
- Expand testing of interventions from academic to community sites



Quality
Chasm

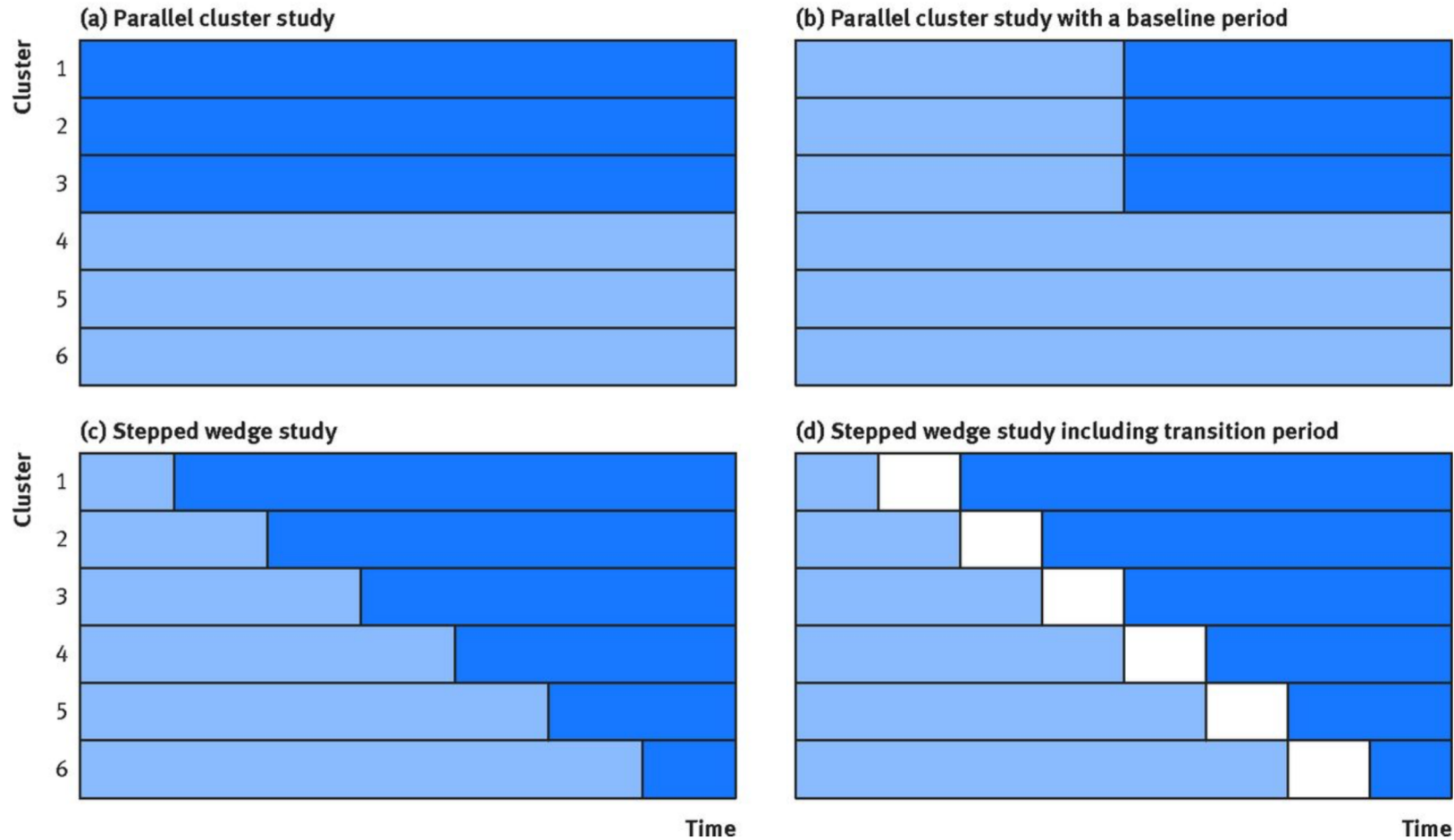
How should
cancer care
be delivered?

CCDR and Cluster Randomization

- Common for intervention to happen at the site or provider level in CCDR studies
- Or intervention may happen at the patient level, but contamination would be a problem if multiple strategies were in place at the same site or for same provider
- Cluster randomized trials are common in CCDR!

Cluster Randomized Designs

■ Cluster exposed to intervention ■ Cluster unexposed to intervention (control) □ Cluster in transition period



Intra-Cluster Correlation (ICC)

- The responses from patients within a cluster are likely to be more similar than those patients from different clusters, so we can't assume independence
- The degree of similarity, or clustering, is commonly quantified by the intra-cluster correlation coefficient (ICC) and is denoted by ρ
- Power/sample size calculations must take the ICC into account, so standard formulas for individually randomized trials don't apply
- Larger ICC - Larger Sample Size
 - Often power for ICC 0.03-0.05

Power and Sample Size

- A consequence of clustering is that the information gained is less than that in an individually randomized trial of the same size
- A sample size calculated assuming individual randomization can be inflated by a Design Effect (DE) to reach the required level of statistical power under cluster randomization

Fixed Cluster Size

$$DE = 1 + (n - 1)\rho$$

Variable Cluster Size

$$DE = 1 + \{(CV^2 + 1)\bar{n} - 1\}\rho$$

n = number per cluster, \bar{n} = average number per cluster, CV = coefficient of variation in cluster size, ρ = ICC

Feasibility

There are a large number of sites in the NCORP network (~1000), but they have less research infrastructure than academic centers.

- Can the intervention be done at community sites?
- What constitutes a cluster?
- Will all clusters be available for randomization at the same time?
- Does each cluster see the same patient volume?
- What happens if a cluster drops out?
- How do you deal with cluster heterogeneity?



Advancing Research. Improving Lives.™

Case Study #1: Projecting number of practices and accrual timelines in NRG-CC007CD

Stephanie Pugh, PhD
Deputy Director, NRG NCORP SDMC
Director of Statistics, ACR

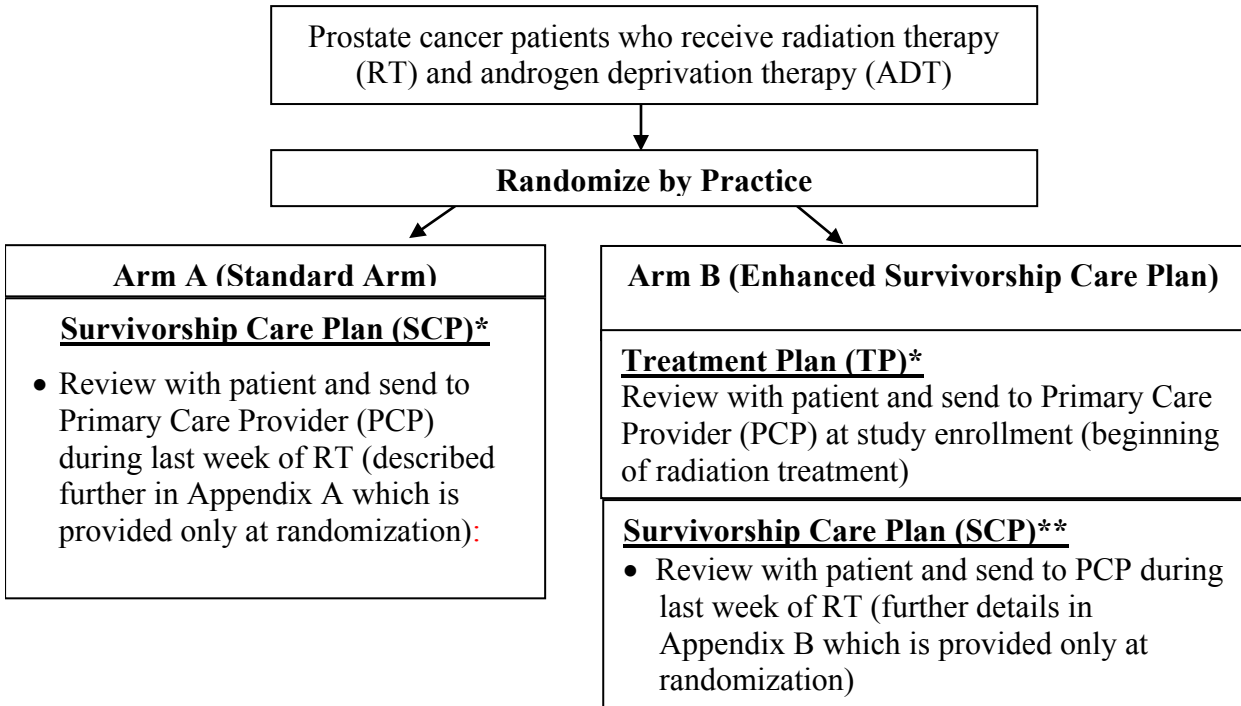
Society for Clinical Trials
May 23, 2022



Disclosures

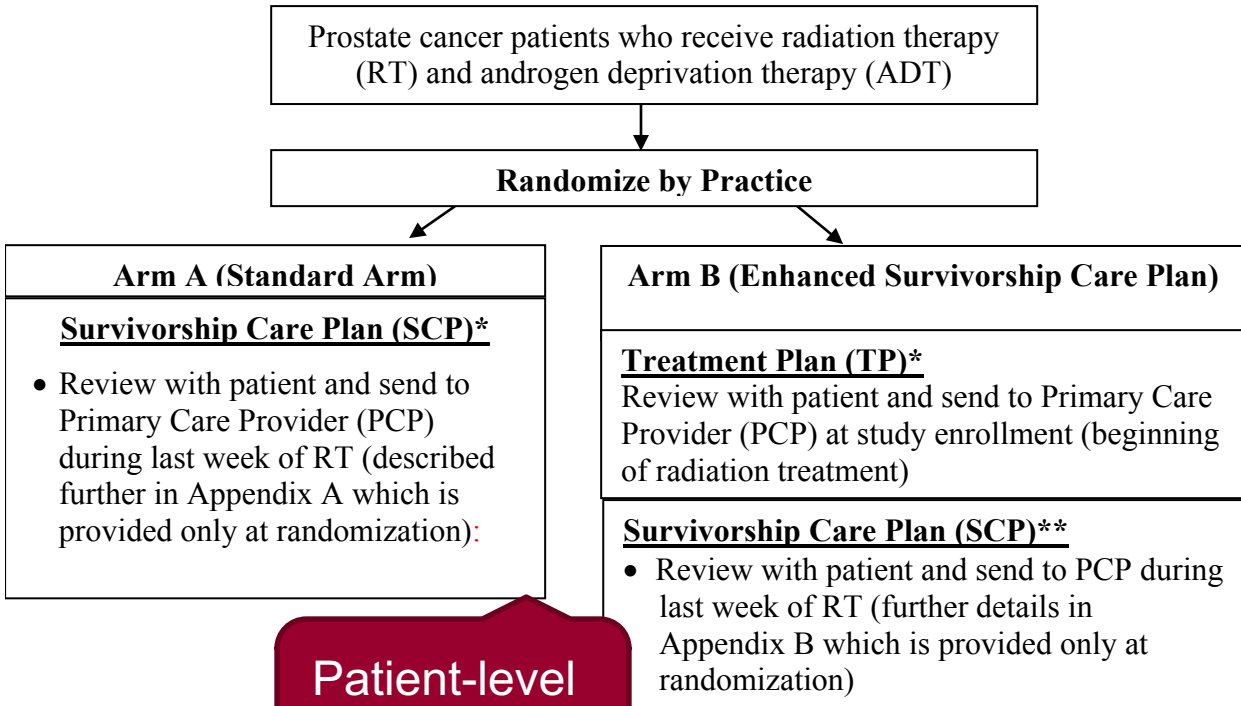
No relevant disclosures

NRG-CC007CD



- **Primary endpoint:** proportion of patients who saw PCP and had fasting blood glucose & fasting cholesterol checked in Year 2 (13-24 months) after completing RT
- **Sample size:** 544 patients from 35 practices
- **Actual enrollment:** 547 patients from 34 practices

NRG-CC007CD



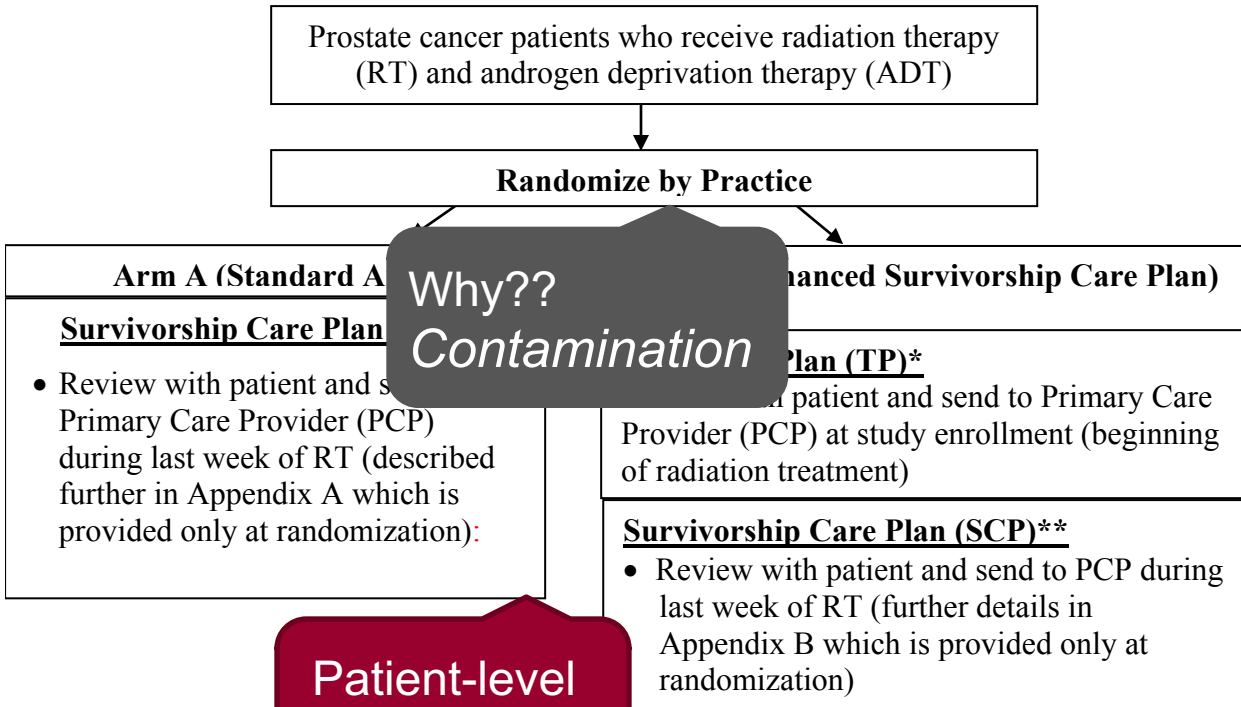
5
Patient-level treatment

- **Primary endpoint:** proportion of patients who saw PCP and had fasting blood glucose & fasting cholesterol checked in Year 2 (13-24 months) after completing RT

5
Patient-level analysis

- **Actual enrollment:** 547 patients from 34 practices

NRG-CC007CD



Why??
Contamination

Patient-level treatment

- **Primary endpoint:** proportion of patients who saw PCP and had fasting blood glucose & fasting cholesterol checked in Year 2 (13-24 months) after completing RT

Patient-level analysis

- **Actual enrollment:** 547 patients from 34 practices

Issues Related to Cluster Randomization

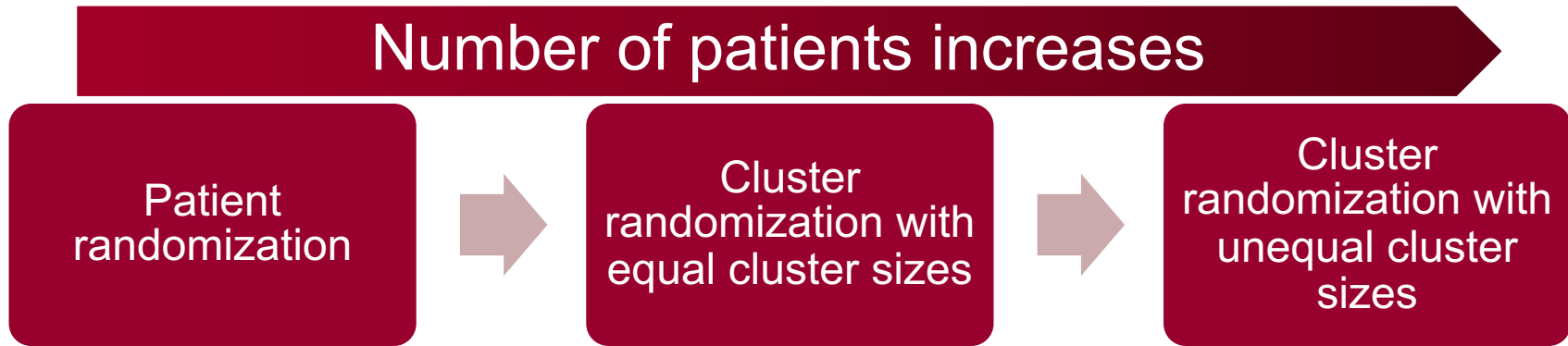
- How to define a “practice”
- Number of practices to randomize
- Unequal enrollment by practice
- How to handle randomized practices without enrollment

Defining Practices

- Single NCORP affiliate or sub-affiliate; or NCORP affiliates and/or sub-affiliates that share the same physicians and/or staff, but are in different locations
- Once a site is randomized as part of a practice, it cannot join a different practice, even in the case of no accrual
 - Arm-specific appendices are provided since knowledge of the enhanced SCP arm could contaminate the control arm
- Some sites randomized to a single practice didn't share staff/physicians

Unequal Enrollment by Practice

- Ideally, **each practice** would enroll the **same number of patients**; however, that is **unlikely** in the NCORP/NCTN
 - Sites are of varying sizes and locations
 - Practice eligibility required seeing 10 eligible patients/year



Less clusters = more patients

Sample Size Calculation

Calculate sample size assuming patient-level randomization then increase by a design effect (DE) that considers clustering and unequal cluster sizes.

$$DE = 1 + ((CV^2 + 1)\bar{n} - 1)\rho$$

ρ = *intra-cluster correlation (ICC)*

CV = *coefficient of variation* = $\frac{\text{std dev of practice enrollment}}{\text{mean of practice enrollment}}$

\bar{n} = *patients per practice*

Sample Size Calculation

Calculate sample size assuming patient-level randomization then increase by a design effect (DE) that considers clustering and unequal cluster sizes.

$$DE = 1 + ((CV^2 + 1)\bar{n} - 1)\rho$$

$\rho =$ *intra-cluster correlation (ICC)*

Correlation within clusters

$CV =$ *coefficient of variation* = $\frac{\text{std dev of practice enrollment}}{\text{mean of practice enrollment}}$

$\bar{n} =$ *patients per practice*

Sample Size Calculation

Calculate sample size assuming patient-level randomization then increase by a design effect (DE) that considers clustering and unequal cluster sizes.

$$DE = 1 + ((CV^2 + 1)\bar{n} - 1)\rho$$

Variation in cluster size

ρ = *intra-cluster correlation (ICC)*

CV = *coefficient of variation* = $\frac{\text{std dev of practice enrollment}}{\text{mean of practice enrollment}}$

\bar{n} = *average number of patients per practice*

Number of Practices

- Sample size:
 - Initial : **504 patients** from **50 practices** with max of 15 patients/practice
 - Revised : **544 patients** from **35 practices** with max of 25 patients/practice
- **2018: 32 NCORPs & 10 NCORP MUs** as NRG Members
 - Over 900 affiliates & sub-affiliates
- Participation from 20 NCORPs

Practices per NCORP



■ One/NCORP ■ Two/NCORP
■ Three/NCORP ■ Four/NCORP

Less clusters = more patients

Sample Size Re-Estimation

- After simulating likely accrual scenarios based on current accrual by each practice, **slightly lower CV is reasonable.**
- ICC was reduced since the primary endpoint is dependent on the patient's **PCP/cardiologist** who is not participating in the trial and is likely to be **different from patient to patient within the same practice**, thus **decreasing** the correlation within clusters

Initial Sample Size

- CV = 0.50
- ICC=0.05
- Avg # patients = 9



Revised Sample Size

- CV = 0.45
- ICC=0.04
- Avg # patients = 14

Unequal Enrollment by Practice

547 patients enrolled

Distribution of patients/practice:

Mean = 15.6

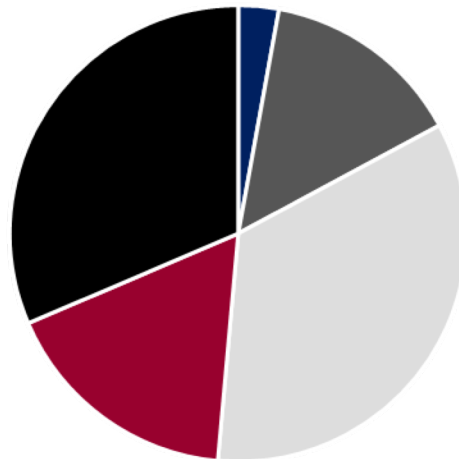
Median = 15

Q1-Q3 = 9-25

Range = 0 – 25

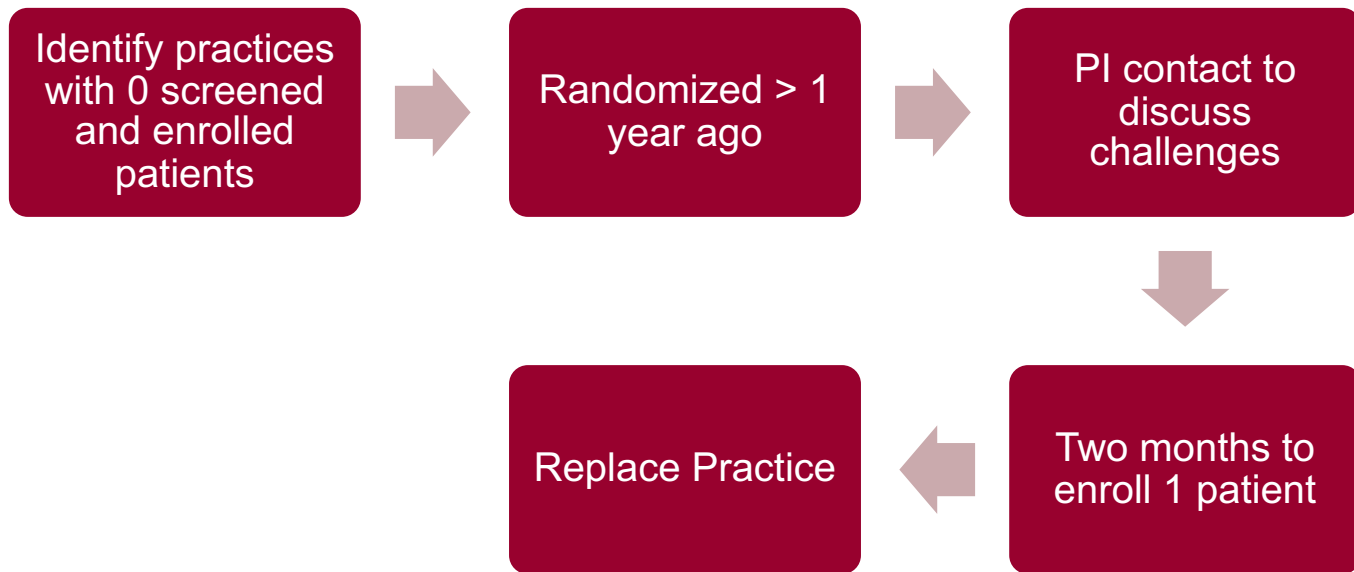
CV = 0.54

Accrual by Practice



■ 0 patients ■ 1-6 patients ■ 7-15 patients ■ 16-24 patients ■ 25 patients

Replacing Practices



- 8 practices meet criteria
- Replaced 3 practices
 - 2 closed out the study at their IRB
- Last practice to be randomized that replaced a non-enrolling practice did not enroll any patients

Lessons Learned

- **Cluster replacement should be defined in the protocol from activation**
 - Replacement should occur early in the trial
 - Consider keeping the trial open to allow all practices to enroll; especially those replacement practices
 - Specify criteria for study closure
- **Although more clusters lead to less patients, number of clusters must be reasonable in the network**



PughS@NRGOncology.org



Case Study 2: Logistics of stepped wedge design and practice stratification in A231701CD (Breast Surgical Decision Aid)

Heather J. Gunn, PhD
Mayo Clinic, Alliance NCORP

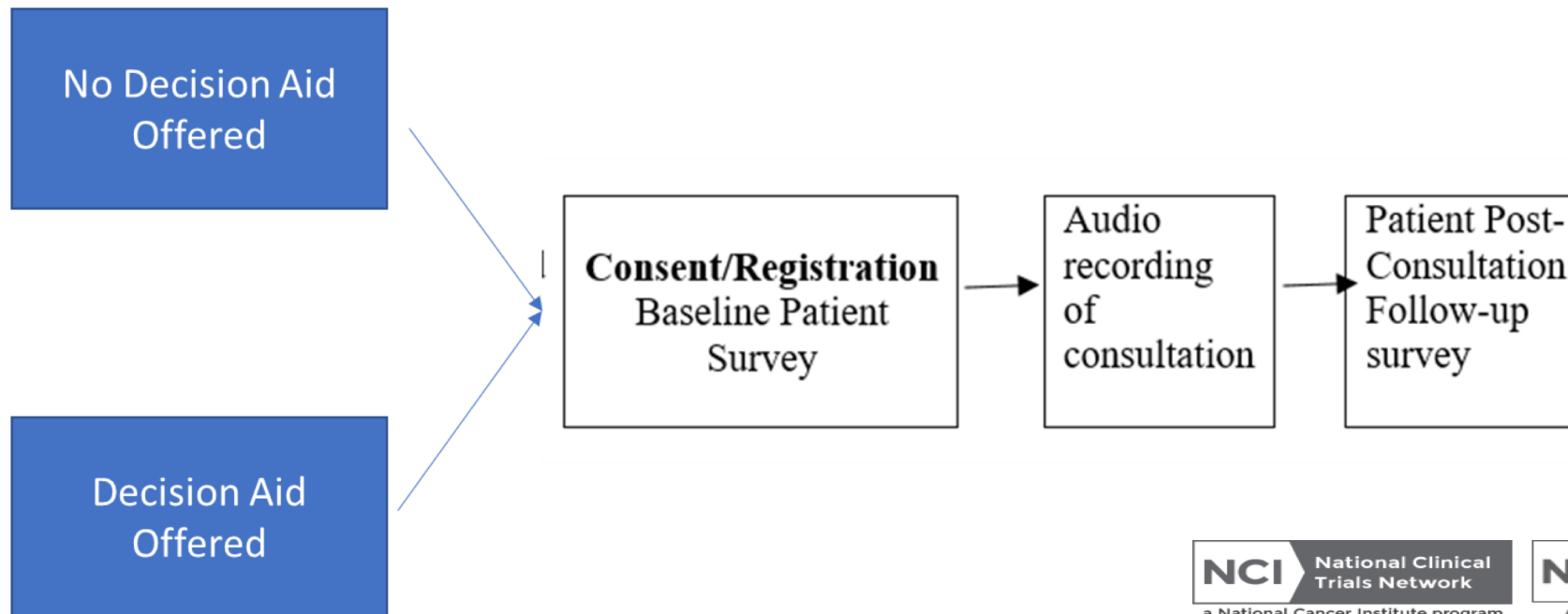
Society for Clinical Trials, May 23, 2023

Disclosures

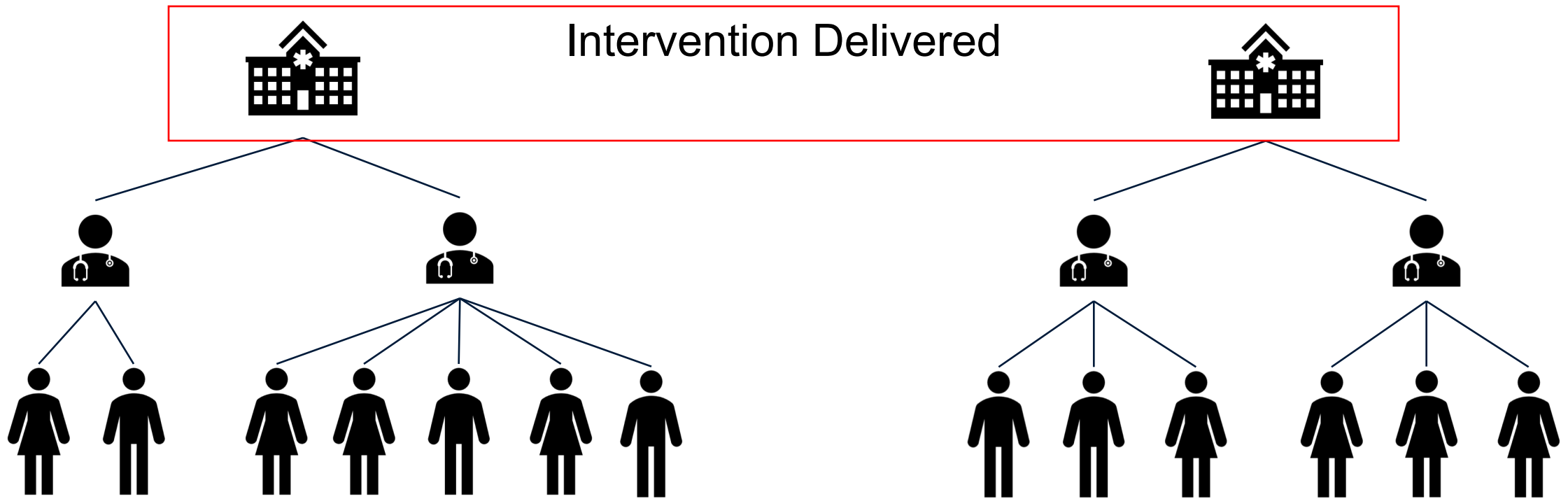
- No relevant disclosures

Study Design

- Intervention: breast cancer surgery decision aid
- Primary Outcomes: patient engagement in decision making as measured by the PEPPI-5 self-efficacy scale and Active Patient Participation behaviors



Cluster Structure



Continuous Recruitment Stepped Wedge Design

Wave	1	2	3	4	5	6	7
Clinics 9-10							
Clinics 7-8							
Clinics 5-6							
Clinics 3-4							
Clinics 1-2							

Usual care group	Implementation period (no data collection)	Intervention group
------------------	---	--------------------

Practice Stratification

- Need to identify sites and have them commit prior to study activation
- For this study, patient volume of biggest concern
- 10 clinics stratified by practice type
 - Group 1: Wisconsin NCORP, Montana Cancer Consortium NCORP, Bay Area Tumor Institute NCORP, Carle Cancer Center NCORP, Hawaii M/U NCORP
 - Group 2: University of New Mexico M/U NCORP, Medical University of South Carolina M/U NCORP, Virginia Commonwealth University Massey Cancer Center M/U NCORP, Stroger Hospital of Cook County M/U NCORP, Georgia Cares M/U NCORP

Challenges of this trial

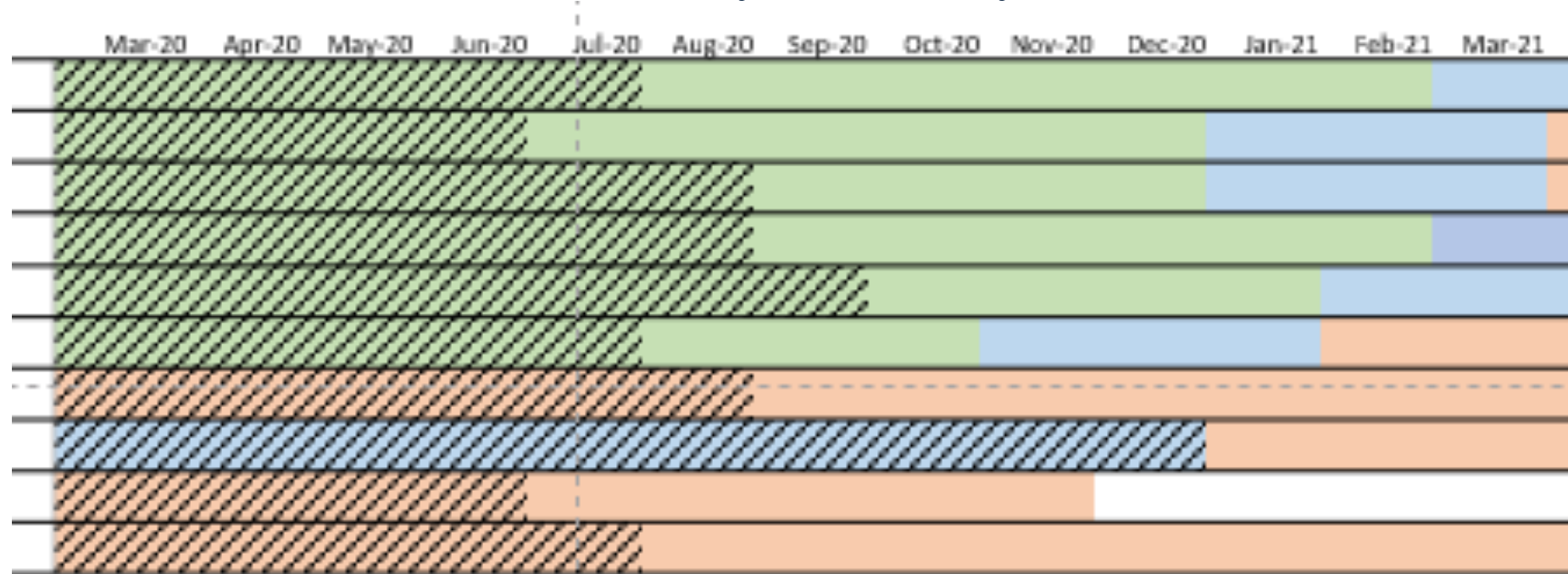
- Timing of the transition
- Meeting accrual targets
- COVID (staff turnover, analytic changes)
 - First patient accrued 2019, last patient accrued 2021

Timing of transition

- Asynchrony at start
 - Local IRB review
 - Local site priorities for activating a study
 - Staffing challenges
- What is the best solution?
 - Extend wave 1 until all sites ready?
 - Make sites wait (and risk losing them for the study)?
 - Just get started?

Timing of transition

- Asynchrony “after” COVID-19 pandemic
 - Every site paused the study at the same time, but sites reopened accrual at different times based on when they were ready and able



Accrual Targets

- Regular accrual across waves to get balance of usual care and intervention
- In reality, accrual varied month to month (staff turnover due to COVID played a huge role)
- Wave Extension Rule: In a wave, if the current accrual for a given site is $<2/3$ of the stated accrual goal for that site at the end of the wave, then the wave will be extended by 3 weeks for that site
 - Increased risk of secular trends
 - Extended time period of study

Clinic specific accrual targets per wave

Stratum 1	Estimated Annual Volume	Target Accrual Goals		
		No. of Patients per Wave	No. of Patients per Month	No. of Patients per Week
Clinic 9	120	15	6	1.5
Clinic 7	120	15	6	1.5
Clinic 5	200	17	7	1.75
Clinic 3	250	17	7	1.75
Clinic 1	240	17	7	1.75
Stratum 2				
Clinic 10	175	15	6	1.5
Clinic 8	300	20	8	2.0
Clinic 6	300	20	8	2.0
Clinic 4	250	20	8	2.0
Clinic 2	300	20	8	2.0
Average No. of Patients		17.5	7.1	1.78

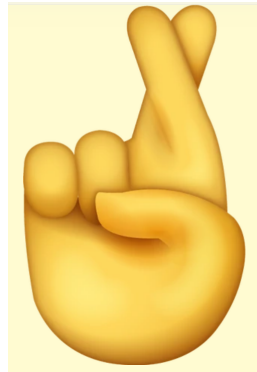
Analytic Change due to COVID

- Sample size reduced from 1,050 to 563 due to low accrual
 - Luckily, observed ICC was smaller than expected
- Timing is important when conducting a stepped wedge
 - Secular trends need to be adjusted for (time confounded with intervention)
 - COVID a huge secular trend

$$y_{ijk} = \beta_0 + \theta x_{ijk} + \beta_1 t_{1ijk} + \beta_2 t_{1ijk}^2 + \beta_3 z_{ijk} + \beta_4 t_{2ijk} + u_k + u_{j(k)} + e_{i(j(k))}$$

Site dropout

- This can be a huge issue for a stepped wedge design
 - Low accrual better than no accrual
- Luckily, this didn't happen with this study
- How did we plan for this?



Key Takeaways

- Need a site champion who will be committed to making it work
- A design that allows for more flexibility, plan with conservative expectations

Questions?



Case Study on Unintended Crossover

Emily V. Dressler, PhD

Associate Professor, Department of Biostatistics and Data Science

Wake Forest School of Medicine



Disclosures

No relevant disclosures

Introduction

- Practice randomized trials collect implementation adoption and fidelity outcomes using mixed methods from sites, providers, and/or participants.
- OaSiS is a hybrid effectiveness-implementation study looking at integrating tobacco use treatment strategies for current smokers undergoing lung cancer screening at time of the visit
 - 14 different options for practices to integrate into their workflow



OaSiS (WF-20817CD) Study Aims

Goal: To identify the best ways to implement evidence-based smoking cessation support at the point-of-care for lung screening patients.

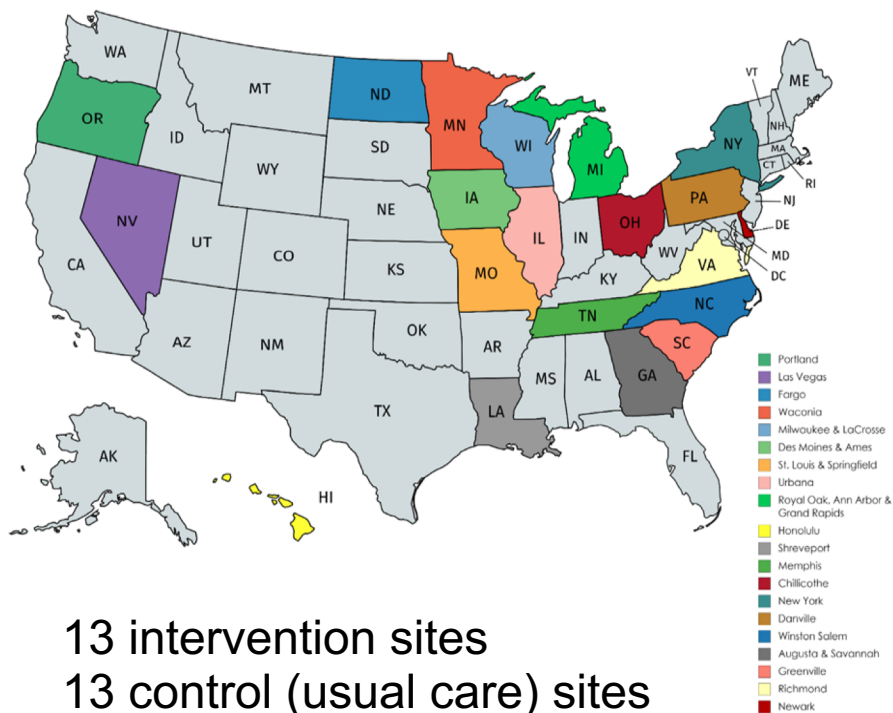
Aim 1: Effectiveness Aim... Evaluate a multi-faceted implementation strategy to improve smoking cessation rates among patients who present for lung cancer screening in community-based NCORP practices.

Aim 2: Implementation Aim... Characterize the adoption and adaptation of evidence-based tobacco cessation strategies



Study Design & Implementation Strategies

National Practice RCT



Implementation Strategies



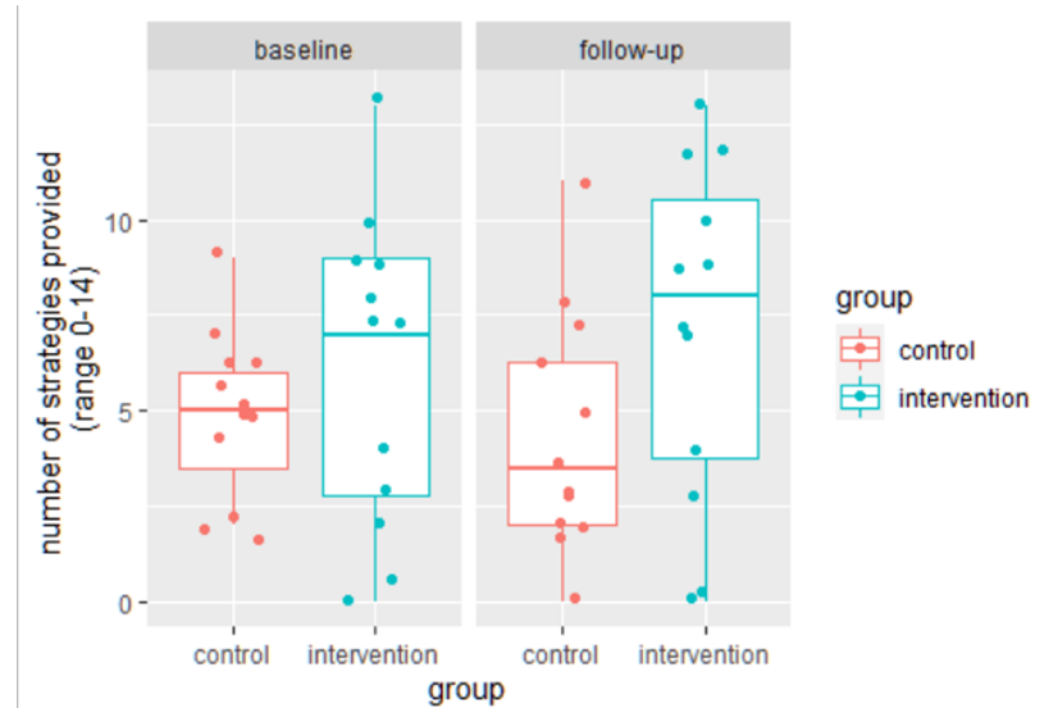
A program of the National Cancer Institute of the National Institutes of Health



Site Reported Cessation Support

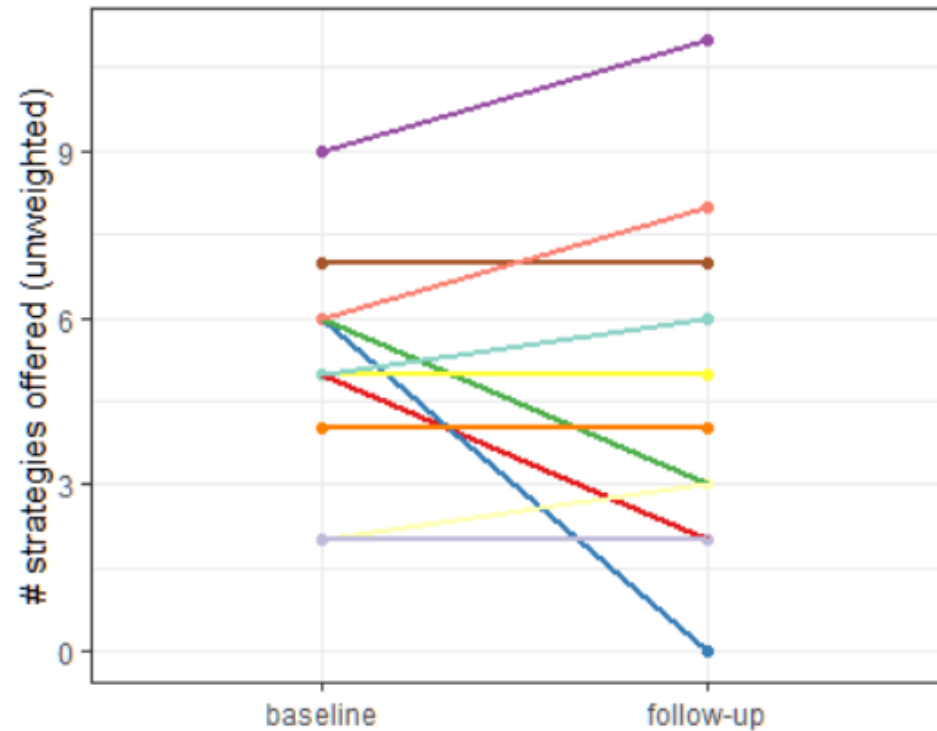
Takeaways...

- Considerable site level heterogeneity in # of strategies offered.
- No difference at baseline in # strategies provided.
- At follow-up...
 - # of strategies in control group decreased
 - # of strategies in intervention group increased
 - Not statistically significant.

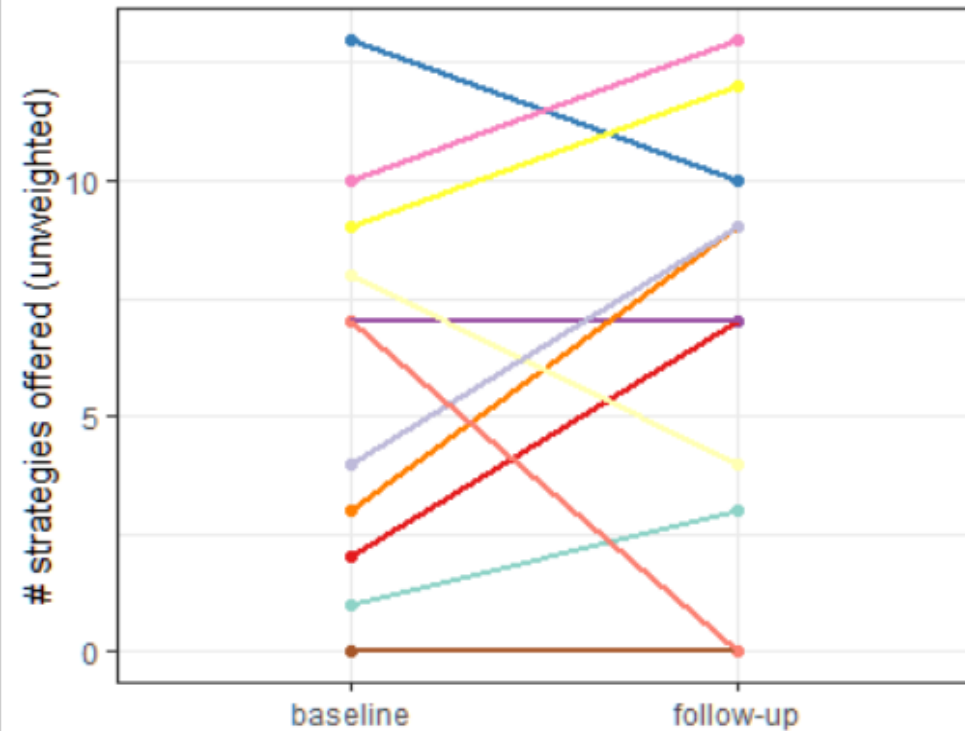


Changes in # services over time

Control



Intervention



Unintended Crossover

- Things happened during the study that caused practices to change approaches
 - Staff Turnover
 - New Trial Enrollment Opportunities
 - Institutional Policy Changes
- How best should we gather this data?
 - How often and from which stakeholders?