Inspired by research. Driven by compassion. Inspiré par la recherche. Guidé par la compassion.

THE STEPPED WEDGE CLUSTER RANDOMIZED TRIAL: FRIEND OR FOE?

NIH PRAGMATIC TRIALS COLLABORATORY SPECIAL GRAND ROUNDS SERIES, 9 DEC 2022

MONICA TALJAARD AND DAVID MAGNUS

The Ottawa Hospital | L'Hôpital d'Ottawa

www.ottawahospital.on.ca | Affiliated with • Affilié à uOttawa

ACKNOWLEDGE-MENTS

Dr. Karla Hemming

University of Birmingham

OUTLINE

1. Introduction to Cluster Randomized Trials (CRTs)
2. Four key methodological complications of Stepped Wedge CRTs (SW-CRTs)
3. Five common methodological arguments for SW-CRTs (and why they may not work)

Dr. Magnus:

4. Common ethical arguments for SW-CRTs (and why they may not work)

INTRODUCTION

- Two main types of clinical trials:
  - Patient randomized trial
  - Cluster randomized trial (CRT)

- A patient randomized trial is always preferable
  - Use cluster randomization only when no other choice

KEY CHARACTERISTICS OF CLUSTER RANDOMIZATION

- Cluster randomized vs. patient randomized trials:
  - Always require larger sample sizes
  - Have higher risks of bias
  - More vulnerable to chance imbalances between arms
  - More complicated to design and analyze

OPEN ACCESS Freely available online

PLOS | MEDICINE

Guidelines and Guidance

The Ottawa Statement on the Ethical Design and Conduct of Cluster Randomized Trials

Charles Weijer[1,2,3*], Jeremy M. Grimshaw[1,4,5], Martin P. Eccles[6], Andrew D. McRae[1,3,7], Angela White[1], Jamie C. Brehaut[4,8], Monica Taljaard[1,4,8], the Ottawa Ethics of Cluster Randomized Trials Consensus Group[¶]

Justifying the Cluster Randomized Design

Recommendation 1: Researchers should provide a clear rationale for the use of the cluster randomized design and adopt statistical methods appropriate for this design.

RATIONALE FOR CLUSTER RANDOMIZATION

- Intervention is a cluster-level intervention 
- Research question of interest pertains to cluster-level effects 
- To avoid contamination

Hemming K, Taljaard M, Moerbeek M, Forbes A. Contamination: How much can an individually randomized trial tolerate? Stat Med. 2021 Jun 30;40(14):3329-3351.

RATIONALE FOR CLUSTER RANDOMIZATION

- Intervention is a cluster-level intervention 
- Research question of interest pertains to cluster-level effects 
- To avoid contamination
- “To be more pragmatic” 

London AJ, Taljaard M, Weijer C. Loopholes in the Research Ethics System? Informed Consent Waivers in Cluster Randomized Trials with Individual-Level Intervention. Ethics Hum Res. 2020 Nov;42(6):21-28.

OPEN ACCESS Freely available online

PLOS | MEDICINE

Guidelines and Guidance

The
Con

Charles
Jamie
Group

White[1],
ensus

a
d
r

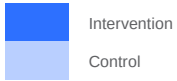
Justifying the cluster randomised design

A cluster randomised trial is more complex to design and conduct and statistically inefficient than an individually randomised trial, and is vulnerable to multiple sources of biases. For these reasons, researchers should clearly justify their choice of cluster rather than individual randomisation (recommendation 1). Acceptable reasons include the evaluation of a cluster level intervention or group effects of an intervention; the need to avoid experimental contamination, reduce costs, enhance compliance, or secure cooperation of investigators; and administrative convenience. Researchers should not adopt this design in a veiled attempt to sidestep the requirements for informed consent.

DIFFERENT TYPES OF CRT DESIGNS

Parallel arm CRT

	Time
Cluster	1
1	
…	
…	
…	
…	
K	



DIFFERENT TYPES OF CRT DESIGNS

Parallel arm CRT

	Time
Cluster	1
1	
…	
…	
…	
…	
K	

Intervention

Control

Parallel arm before and after CRT

	Time	
Cluster	1	2
1		
…		
…		
…		
…		
K		

DIFFERENT TYPES OF CRT DESIGNS

Parallel arm CRT

	Time
Cluster	1
1	
…	
…	
…	
…	
K	

Intervention

Control

Parallel arm before and after CRT

	Time	
Cluster	1	2
1		
…		
…		
…		
…		
K		

Cluster cross-over

	Time	
Cluster	1	2
1		
…		
…		
…		
…		
K		

DIFFERENT TYPES OF CRT DESIGNS

Parallel arm CRT

	Time
Cluster	1
1	
…	
…	
…	
…	
K	

Intervention

Control

Parallel arm before and after CRT

	Time	
Cluster	1	2
1		
…		
…		
…		
…		
K		

Cluster cross-over

	Time	
Cluster	1	2
1		
…		
…		
…		
…		
K		

Stepped wedge

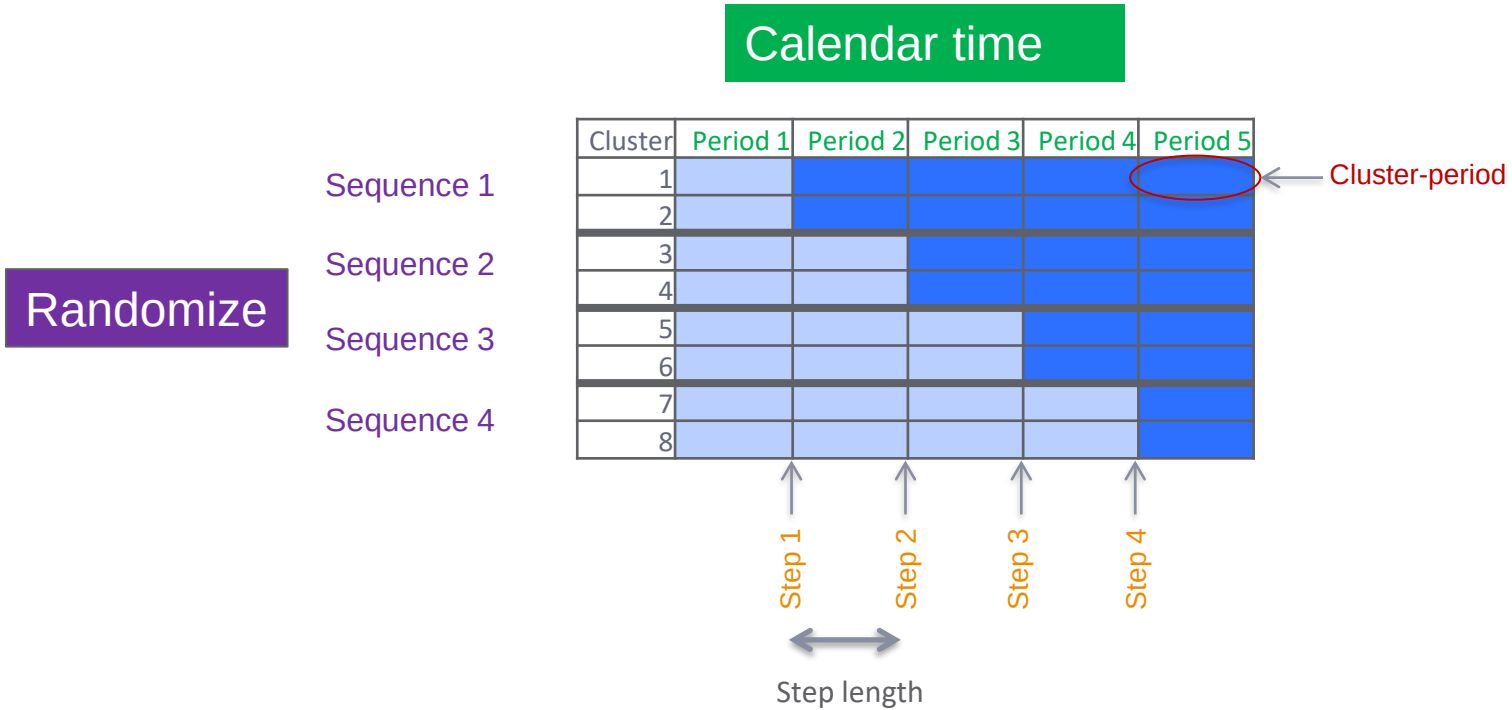
	Time			
Cluster	1	2	3	4
1				
…				
…				
…				
…				
K				

JUSTIFICATION FOR CHOICE OF SW-CRT

- CONSORT extension for SW-CRTs:
  - Provide a clear rationale for cluster randomization
  - Provide a clear rationale for the stepped wedge roll-out

Hemming K, Taljaard M, McKenzie J E, Hooper R, Copas A, Thompson J A et al. Reporting of stepped wedge cluster randomised trials: extension of the CONSORT 2010 statement with explanation and elaboration *BMJ* 2018; 363:k1614

TERMINOLOGY



EXAMPLE

ORIGINAL RESEARCH ARTICLE

Point-of-Care Hemostatic Testing in Cardiac Surgery

A Stepped-Wedge Clustered Randomized Controlled Trial

BACKGROUND: Cardiac surgery is frequently complicated by coagulopathic bleeding that is difficult to optimally manage using standard hemostatic testing. We hypothesized that point-of-care hemostatic testing within the context of an integrated transfusion algorithm would improve the management of coagulopathy in cardiac surgery and thereby reduce blood transfusions.

METHODS: We conducted a pragmatic multicenter stepped-wedge cluster randomized controlled trial of a point-of-care–based transfusion algorithm in consecutive patients undergoing cardiac surgery with cardiopulmonary bypass at 12 hospitals from October 6, 2014, to May 1, 2015. Following a 1-month data collection at all participating hospitals, a transfusion algorithm incorporating point-of-care hemostatic testing was sequentially implemented at 2 hospitals at a time in 1-month intervals, with the implementation order randomly assigned. No other aspects of care were modified. The primary outcome was red blood cell transfusion from surgery to postoperative day 7. Other outcomes included transfusion of other blood products, major bleeding, and major complications. The analysis adjusted for secular time trends, within-hospital clustering, and patient-level risk factors. All outcomes and analyses were prespecified before study initiation.

Keyvan Karkouti, MD
Jeannie Callum, MD
Duminda N. Wijeysundera, MD, PhD
Vivek Rao, MD, PhD
Mark Crowther, MD
Hilary P. Grocott, MD
Ruxandra Pinto, PhD
Damon C. Scales, MD, PhD
TACS Investigators

EXAMPLE: SUMMARY

- **Background:** Major bleeding is a frequent complication of cardiac surgery. Point-of-care tests have faster turnaround times allowing for more rapid management of bleeding.
- **Objective**: Does a point-of-care blood test within the context of an integrated transfusion algorithm reduce red blood cell transfusions?
- **Design:** SW-CRT at 12 hospitals (7,402 surgeries) over 7 months
- **Intervention:** Transfusion algorithm incorporating point-of-care blood test
- **Primary outcome:** Red blood cell transfusion within 7 days
- **Results:** The intervention reduced rates of red blood cell transfusion (adjusted relative risk 0.91, 95% CI 0.85 to 0.98; p=0.02)

EXAMPLE: DESIGN

Group 6 N=2 Hospitals	n=144	n=140	n=150	n=132	n=130	n=168	n=114
Group 5 N=2 Hospitals	n=192	n=197	n=227	n=214	n=211	n=258	n=203
Group 4 N=2 Hospitals	n=189	n=175	n=183	n=178	n=171	n=209	n=135
Group 3 N=2 Hospitals	n=136	n=121	n=122	n=136	n=115	n=146	n=135
Group 2 N=2 Hospitals	n=172	n=170	n=171	n=174	n=164	n=214	n=170
Group 1 N=2 Hospitals	n=204	n=220	n=216	n=220	n=214	n=250	n=212
Total (n=7402)	**n=1037**	**n=1023**	**n=1069**	**n=1054**	**n=1005**	**n=1245**	**n=969**
Period	**Baseline** Oct 1 2014– Nov 2 2014	**Step 1** Nov 3 2014 – Nov 30 2014	**Step 2** Dec 1 2014 – Jan 4 2015	**Step 3** Jan 5 2015 – Feb 1 2015	**Step 4** Feb 2 2015 – Mar 1 2015	**Step 5** Mar 2 2015 – Apr 5 2015	**Follow-up** Apr 6 2015 – May 1, 2015

Control Phase:

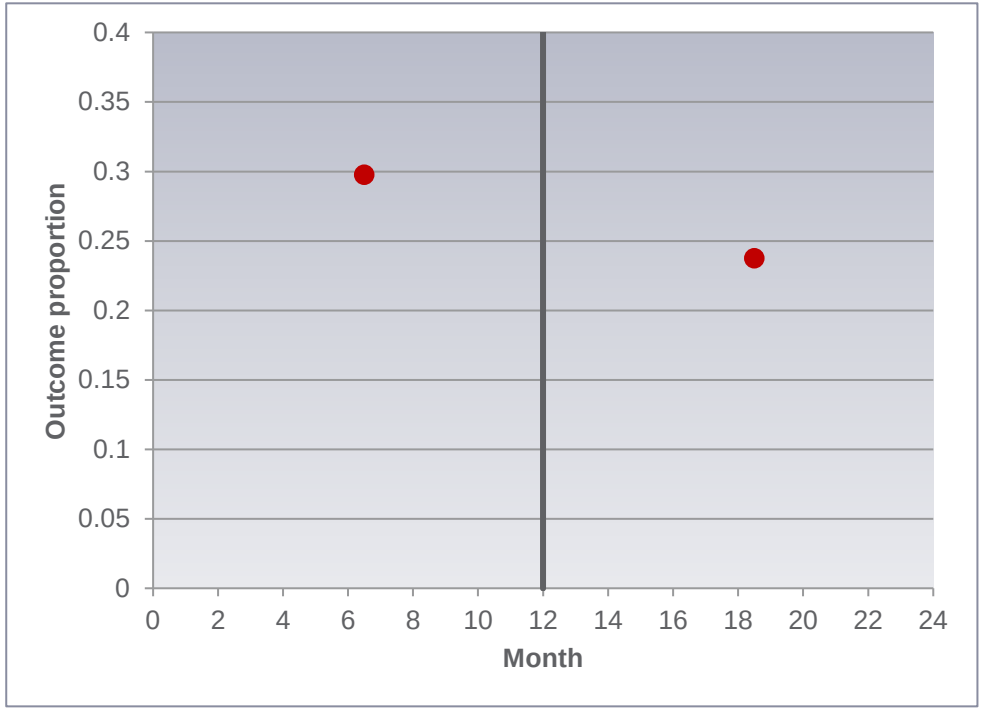
Intervention Phase:

FOUR KEY METHODOLOGICAL COMPLICATIONS OF SW-CRTs

- SW-CRTs have several characteristics that complicate their design and analysis
  1. Confounding by time
  2. Within-cluster contamination
  3. Time-varying intervention effect
  4. Complex intracluster correlations

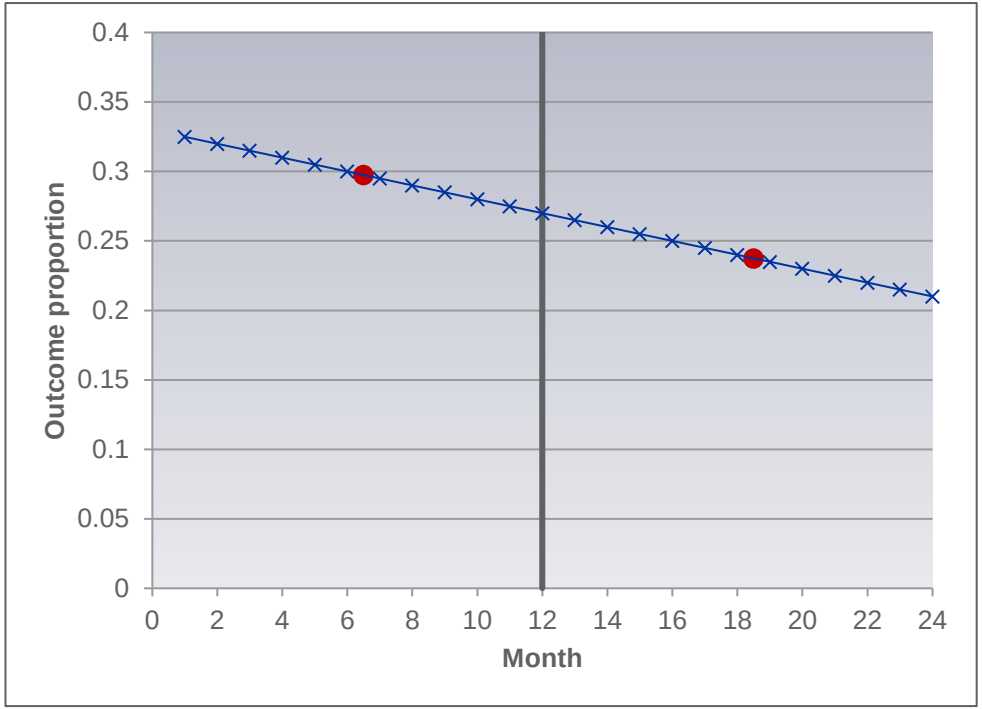
(1) CONFOUNDING BY TIME

- Before-and-after study: intervention is *completely* confounded with time



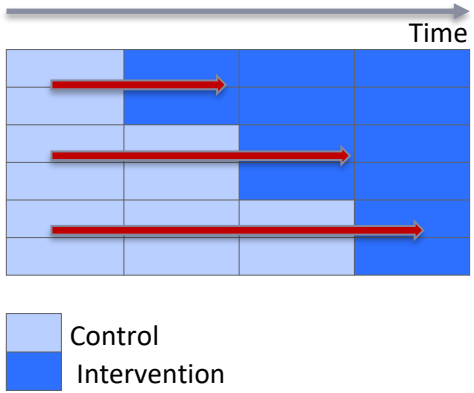
(1) CONFOUNDING BY TIME

- Before-and-after study: intervention is *completely* confounded with time



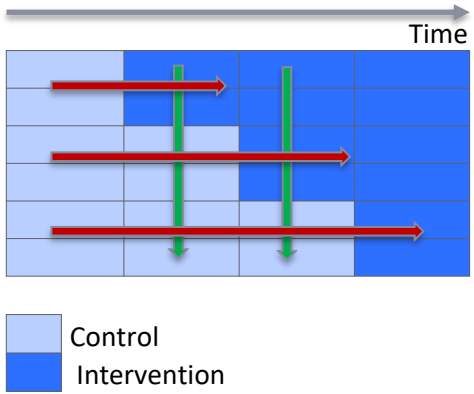
(1) CONFOUNDING BY TIME

- SW-CRT includes a mixture of within and between-cluster comparisons



(1) CONFOUNDING BY TIME

- SW-CRT includes a mixture of within and between-cluster comparisons
- Intervention is *partially* confounded with time
- Requires a model-based analysis adjusting for time



EXAMPLE

Effect of a novel vital sign device on maternal mortality and morbidity in low-resource settings: a pragmatic, stepped-wedge, cluster-randomised controlled trial

Nicola Vousden, Elodie Lawley, Hannah L Nathan, Paul T Seed, Muchabayiwa Francis Gidiri, Shivaprasad Goudar, Jane Sandall, Lucy C Chappell, Andrew H Shennan*, on behalf of the CRADLE Trial Collaborative Group†*

Summary

Background In 2015, an estimated 303 000 women died in pregnancy and childbirth. Obstetric haemorrhage, sepsis, and hypertensive disorders of pregnancy account for more than 50% of maternal deaths worldwide. There are effective treatments for these pregnancy complications, but they require early detection by measurement of vital signs and timely administration to save lives. The primary aim of this trial was to determine whether implementation of the CRADLE Vital Sign Alert and an education package into community and facility maternity care in low-resource settings could reduce a composite of all-cause maternal mortality or major morbidity (eclampsia and hysterectomy).

Methods We did a pragmatic, stepped-wedge, cluster-randomised controlled trial in ten clusters across Africa, India, and Haiti, introducing the device into routine maternity care. Each cluster contained at least one secondary or eferral facilities. Clusters crossed over from existing routine care to the CRADLE at 2-monthly intervals, with CRADLE devices replacing existing equipment at the computer-generated randomly allocated sequence determined the order in which tion. Because of the nature of the intervention, this trial was not masked. Data time periods of 1 month. The primary composite outcome was at least one of This study is registered with the ISRCTN registry,

Lancet Glob Health 2019; 7: e347–56

See Comment page e290

*Contributed equally

†Members listed at the end of the Article and in appendix

Department of Women and Children's Health, School of Life Course Sciences, Faculty of Life Sciences and Medicine, King's College London, London, UK (N Vousden MBBS, E Lawley BSc, H L Nathan PhD, P T Seed CSTAT, J Sandall PhD, Prof L C Chappell FRCOG, Prof A H Shennan FRCOG); **Department of Obstetrics and Gynaecology, College of Health Sciences, University of Zimbabwe, Harare, Zimbabwe**

Before adjusting for time:
OR=0.92, 95% CI 0.86 to 0.97; p=0.0056

After adjusting for time:
OR=1.22, 95% CI 0.73 to 2.06; p=0.45

(2) WITHIN-CLUSTER CONTAMINATION

- Increased risk of within-cluster contamination
- At patient-level
  - Patients recruited in the control period but remain in cluster after cross-over
- At cluster-level
  - Intervention implemented earlier than scheduled (they can't wait)
  - Intervention implemented later than scheduled (logistical challenges)

Copas AJ e.a. (2015) Designing a stepped wedge trial: three main designs, carry-over effects and randomisation approaches. *Trials*; 16:352

ELSEVIER

Journal of Clinical Epidemiology 148 (2022) 93–103

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

Recruitment and implementation challenges were common in stepped-wedge cluster randomized trials: Results from a methodological review

Agnes Caille[a,b,c,*], Monica Taljaard[d,e], Floriane Le Vilain—Abraham[b,c], Alexis Le Moigne[a], Andrew J. Copas[f], Florence Tubach[a,1], Agnes Dechartres[a,1]

[a]*Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique, AP-HP, Hôpital Pitié-Salpêtrière, Département de Santé Publique, Paris, France*

[b]*Université de Tours, Université de Nantes, INSERM, SPHERE U1246, Tours, France*

[c]*INSERM CIC 1415, CHRU de Tours, Tours, France*

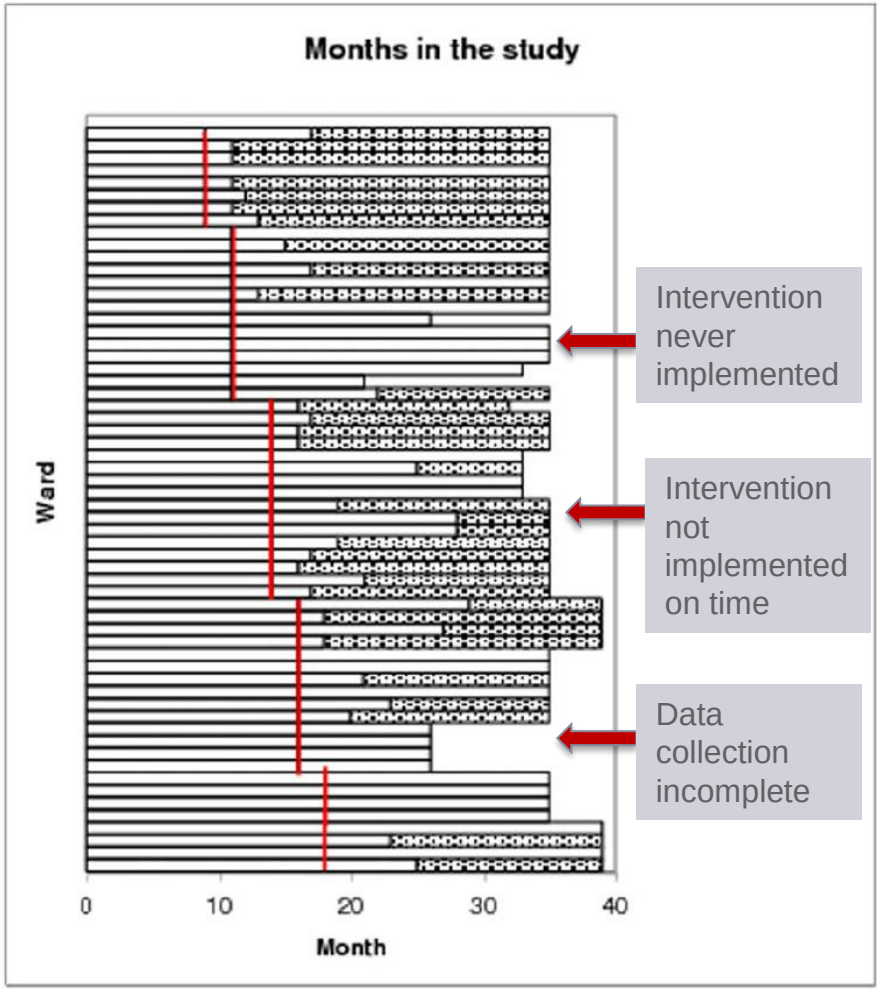
[d]*Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Canada*

[e]*School of Epidemiology and Public Health, University of Ottawa, Ottawa, Canada*

[f]*MRC Clinical Trials Unit at University College London, London, United Kingdom*

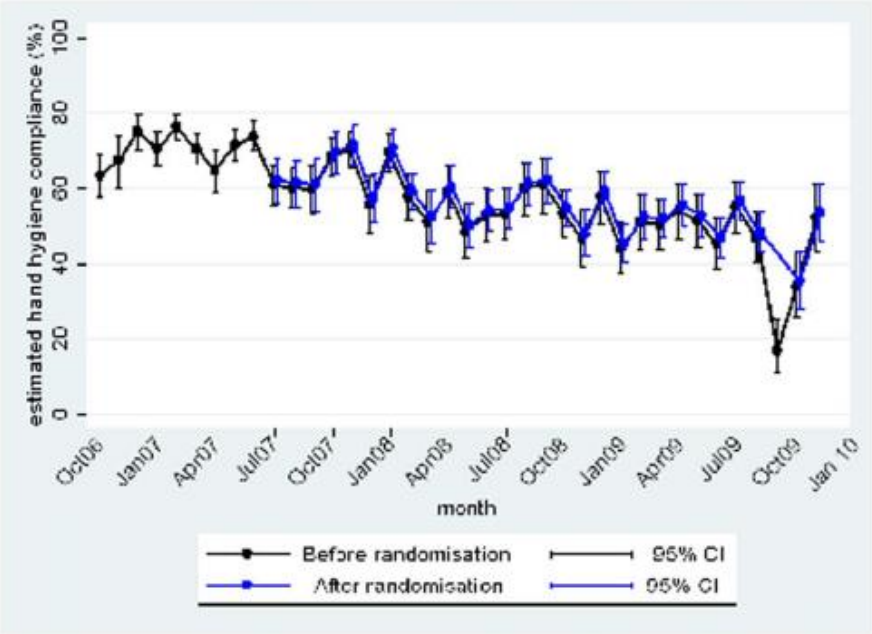
Accepted 20 April 2022; Published online 26 April 2022

At least one implementation challenge reported in 44% of 55 SW-CRTs published 2019-2020

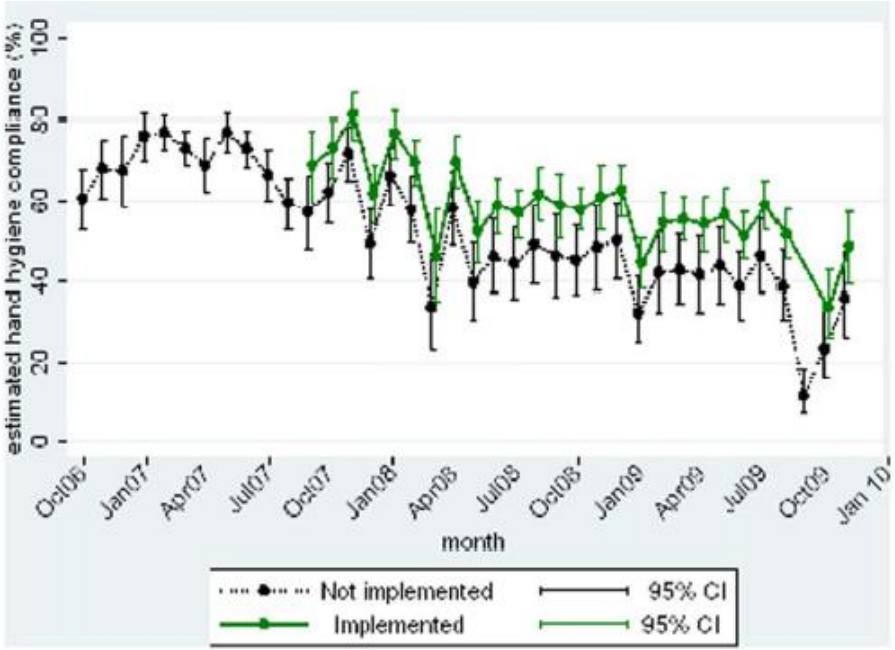


Fuller et al. The feedback Intervention Trial – Improving Hand Hygiene Compliance in UK Healthcare Workers. *Plos One* 2012;7(10):e41617

Intention-to-treat analysis



Per-protocol analysis.

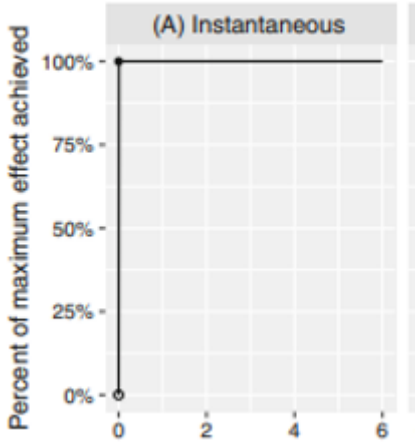


Hand hygiene compliance intervention
OR=1.06 (0.87 to 1.27), p=0.5

Hand hygiene compliance intervention
OR=1.67 (1.26 to 2.22), p<0.001

(3) TIME-VARYING INTERVENTION EFFECT

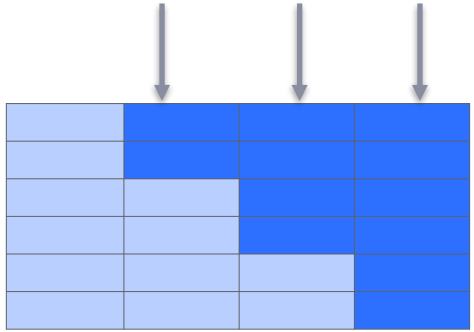
- Standard analytical approach assumes intervention works immediately and keeps working



Kenny, A, Voldal, E, Xia, F, Heagerty, PJ, Hughes, JP. Analysis of stepped wedge cluster randomized trials in the presence of a time-varying treatment effect. *Statistics in Medicine*. 2022; 41(22): 4311– 4339.

(3) TIME-VARYING INTERVENTION EFFECT

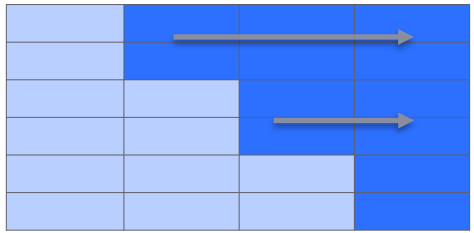
- True effect may vary with calendar time
  - Seasonal variation, external events



Copas AJ e.a. (2015) Designing a stepped wedge trial: three main designs, carry-over effects and randomisation approaches. *Trials*; 16:352

(3) TIME-VARYING INTERVENTION EFFECT

- True effect may vary with calendar time
  - Seasonal variation, external events
- True effect may vary with duration of exposure
  - Increase with more experience
  - Weaken over time
- An analysis which assumes a time-fixed intervention effect may be biased



Copas AJ e.a. (2015) Designing a stepped wedge trial: three main designs, carry-over effects and randomisation approaches. *Trials*; 16:352

BIAS IN ESTIMATED INTERVENTION EFFECT

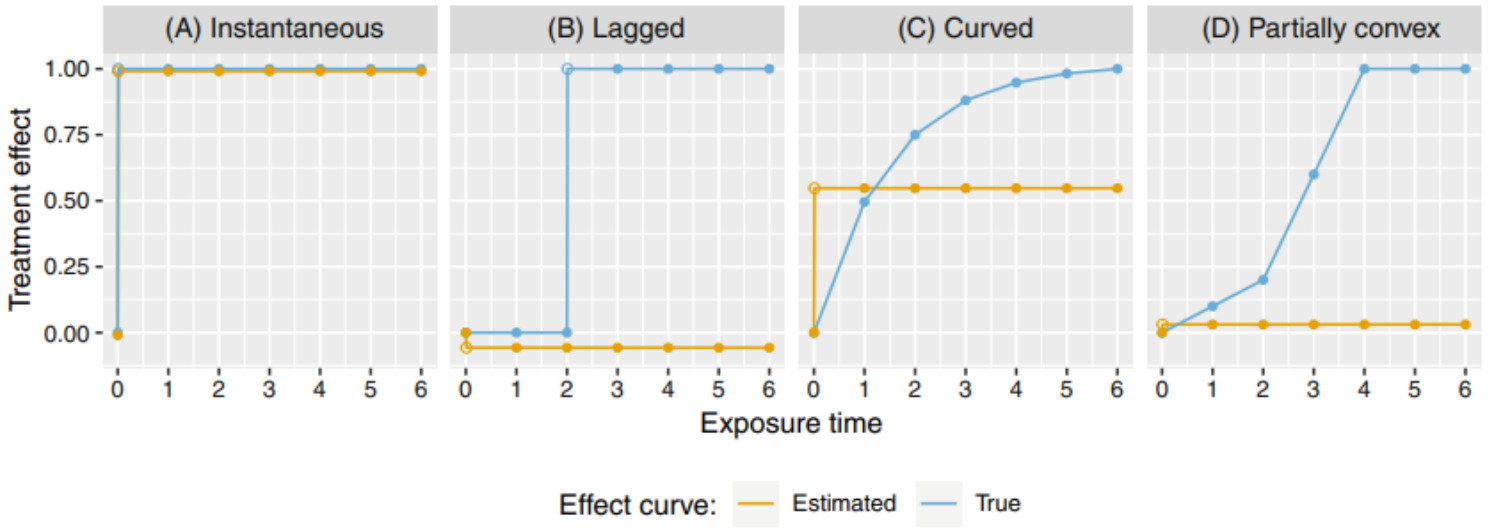


FIGURE 2 Four possible true effect curves plotted against the expected effect curves estimated from an IT model, for a design with $Q = 6$ sequences and $\phi = 0.5$

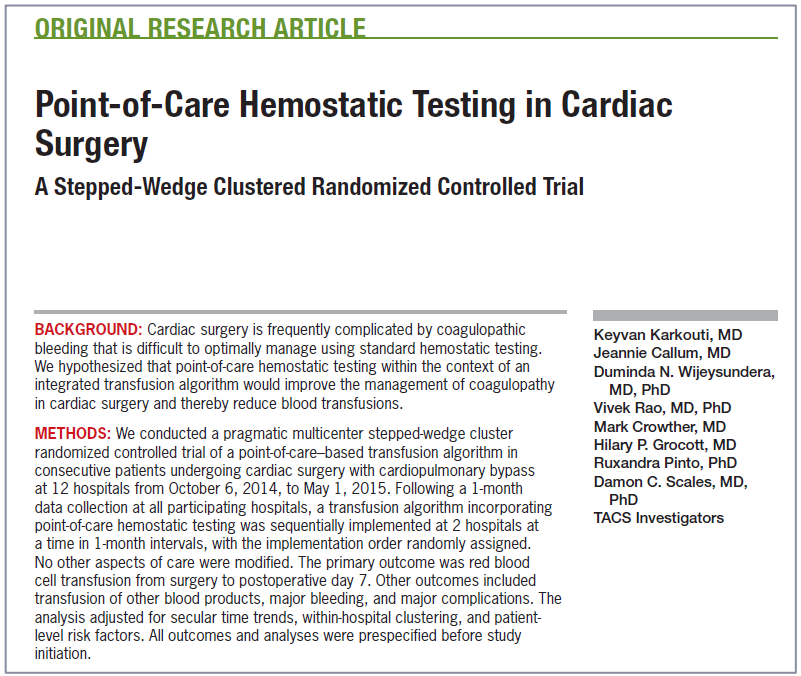
Kenny A, Voldal E, Xia F, Heagerty PJ, Hughes JP. Analysis of stepped wedge cluster randomized trials in the presence of a time-varying treatment effect. *Statistics in Medicine*. 2022; 41(22): 4311– 4339

(4) COMPLEX CORRELATIONS

- Intracluster correlations are complex!
  - (1) Within clusters in the same period
  - (2) Within clusters in different periods
- Design stage
  - Correlation parameters usually unknown
- Analysis stage
  - Can get different answers by specifying different types of correlation structures
  - Computational challenges

EXAMPLE

- Did not allow for correlation decay



ORIGINAL RESEARCH ARTICLE

Point-of-Care Hemostatic Testing in Cardiac Surgery

A Stepped-Wedge Clustered Randomized Controlled Trial

BACKGROUND: Cardiac surgery is frequently complicated by coagulopathic bleeding that is difficult to optimally manage using standard hemostatic testing. We hypothesized that point-of-care hemostatic testing within the context of an integrated transfusion algorithm would improve the management of coagulopathy in cardiac surgery and thereby reduce blood transfusions.

METHODS: We conducted a pragmatic multicenter stepped-wedge cluster randomized controlled trial of a point-of-care–based transfusion algorithm in consecutive patients undergoing cardiac surgery with cardiopulmonary bypass at 12 hospitals from October 6, 2014, to May 1, 2015. Following a 1-month data collection at all participating hospitals, a transfusion algorithm incorporating point-of-care hemostatic testing was sequentially implemented at 2 hospitals at a time in 1-month intervals, with the implementation order randomly assigned. No other aspects of care were modified. The primary outcome was red blood cell transfusion from surgery to postoperative day 7. Other outcomes included transfusion of other blood products, major bleeding, and major complications. The analysis adjusted for secular time trends, within-hospital clustering, and patient-level risk factors. All outcomes and analyses were prespecified before study initiation.

Keyvan Karkouti, MD
Jeannie Callum, MD
Duminda N. Wijeysundera, MD, PhD
Vivek Rao, MD, PhD
Mark Crowther, MD
Hilary P. Grocott, MD
Ruxandra Pinto, PhD
Damon C. Scales, MD, PhD
TACS Investigators

- Assuming a constant ICC of 0.095 (no decay)
  - Need 12 hospitals, 7402 patients
- Allowing for a 20% correlation decay per month
  - Need 24 hospitals, 14280 patients

METHODO-LOGICAL ARGUMENTS FOR SW-CRTs

(and why they may not work)

1. “To improve rigour”
2. “To facilitate recruitment”
3. “To reduce the required sample size”
4. “To simplify logistics”
5. “To reduce bias”

- Hemming K, Taljaard M. Reflection on modern methods: when is a stepped-wedge cluster randomized trial a good study design choice? *Int J Epidemiol*. 2020 Jun 1;49(3):1043-1052.
- Hooper R, Eldridge SM. Cutting edge or blunt instrument: how to decide if a stepped wedge design is right for you. *BMJ Quality & Safety* 2021;**30:**245-250.

METHODO-LOGICAL ARGUMENTS FOR SW-CRTs

(and why they may not work)

1. “To improve rigour”
2. “To facilitate recruitment”
3. “To reduce required sample size”
4. “To simplify logistics”
5. “To reduce bias”
7. ~~“I have always wanted to try a stepped wedge”~~
8. “~~It will make my grant application more attractive to the funder~~”

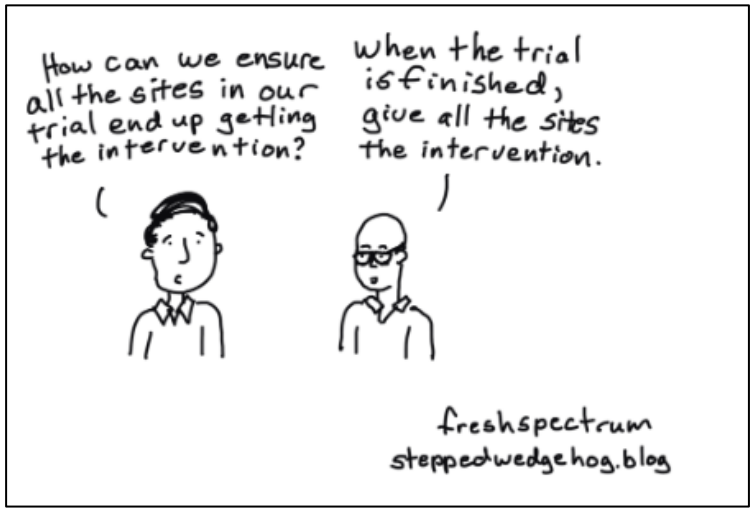
- Hemming K, Taljaard M. Reflection on modern methods: when is a stepped-wedge cluster randomized trial a good study design choice? *Int J Epidemiol*. 2020 Jun 1;49(3):1043-1052.
- Hooper R, Eldridge SM. Cutting edge or blunt instrument: how to decide if a stepped wedge design is right for you. *BMJ Quality & Safety* 2021;**30:**245-250.

1: TO IMPROVE RIGOUR

- YES
  - Decision has been made by a stakeholder to implement a program / intervention
  - SW-CRT design allows more rigorous evaluation than a before-and-after design
- NO
  - Will have to convince stakeholder of the importance of randomization
  - Will have to reconcile need for adherence to allocated schedule with stakeholder preferences

2: TO FACILITATE RECRUITMENT

- YES
  - Easier to recruit clusters to the trial when all are guaranteed to receive something new



https://steppedwedgehog.blog/

- NO
  - Consider parallel arm design with control clusters offered intervention at the end of the trial
  - In a SW-CRT, some clusters may have to wait even longer to receive the intervention

ELSEVIER

Journal of Clinical Epidemiology 107 (2019) 89–100

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

Systematic review showed that stepped-wedge cluster randomized trials often did not reach their planned sample size

Felizitas A. Eichner[a,*,1], Rolf H.H. Groenwold[a,b], Diederick E. Grobbee[a], Katrien Oude Rengerink[a]

[a]*Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, the Netherlands*
[b]*Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, the Netherlands*

Accepted 14 November 2018; Published online 17 November 2018

Of N=46 SW-CRTs, 20% could not recruit their target number of clusters

3: TO REDUCE REQUIRED SAMPLE SIZE

- YES
  - The SW design *usually* requires fewer clusters than parallel arm design
- NO
  - Sample size parameters reliable?
  - Consider more efficient parallel arm (e.g., before-and-after) or cross-over design
  - A trial with very few clusters is probably not advisable

Median (Q1-Q3) number of clusters in 160 SW-CRTs published 2016-2022 was 11 (8-18)

RETURNING TO THE EXAMPLE

ORIGINAL RESEARCH ARTICLE

Point-of-Care Hemostatic Testing in Cardiac Surgery

A Stepped-Wedge Clustered Randomized Controlled Trial

BACKGROUND: Cardiac surgery is frequently complicated by coagulopathic bleeding that is difficult to optimally manage using standard hemostatic testing. We hypothesized that point-of-care hemostatic testing within the context of an integrated transfusion algorithm would improve the management of coagulopathy in cardiac surgery and thereby reduce blood transfusions.

METHODS: We conducted a pragmatic multicenter stepped-wedge cluster randomized controlled trial of a point-of-care–based transfusion algorithm in consecutive patients undergoing cardiac surgery with cardiopulmonary bypass at 12 hospitals from October 6, 2014, to May 1, 2015. Following a 1-month data collection at all participating hospitals, a transfusion algorithm incorporating point-of-care hemostatic testing was sequentially implemented at 2 hospitals at a time in 1-month intervals, with the implementation order randomly assigned. No other aspects of care were modified. The primary outcome was red blood cell transfusion from surgery to postoperative day 7. Other outcomes included transfusion of other blood products, major bleeding, and major complications. The analysis adjusted for secular time trends, within-hospital clustering, and patient-level risk factors. All outcomes and analyses were prespecified before study initiation.

Keyvan Karkouti, MD
Jeannie Callum, MD
Duminda N. Wijeysundera, MD, PhD
Vivek Rao, MD, PhD
Mark Crowther, MD
Hilary P. Grocott, MD
Ruxandra Pinto, PhD
Damon C. Scales, MD, PhD
TACS Investigators

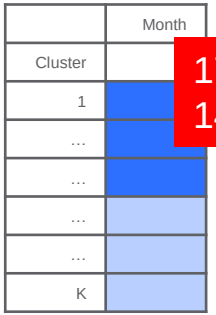
Stepped wedge

	Month						
Step	1	2	3	4	5	6	
1							
2							
3							
4							
5							
6							

12 hospitals
7140 patients

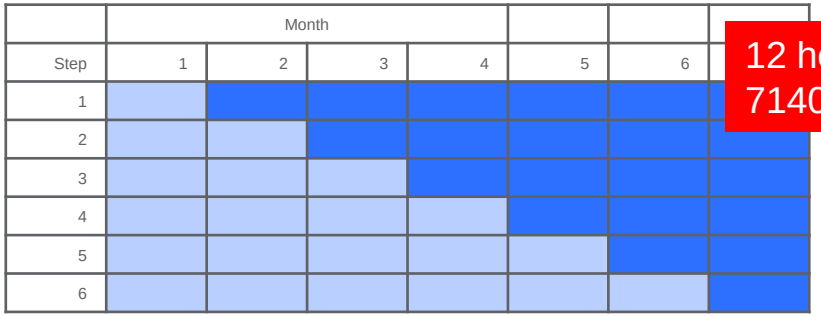
RETURNING TO THE EXAMPLE

Parallel arm



172 hospitals
14,620 patients

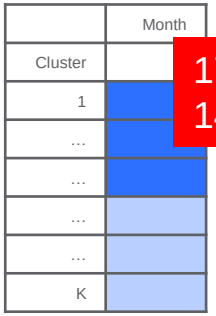
Stepped wedge



12 hospitals
7140 patients

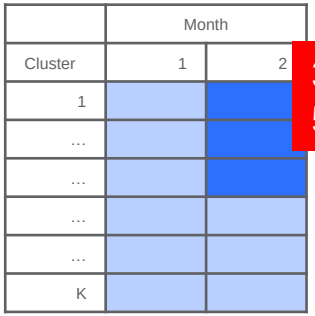
RETURNING TO THE EXAMPLE

Parallel arm



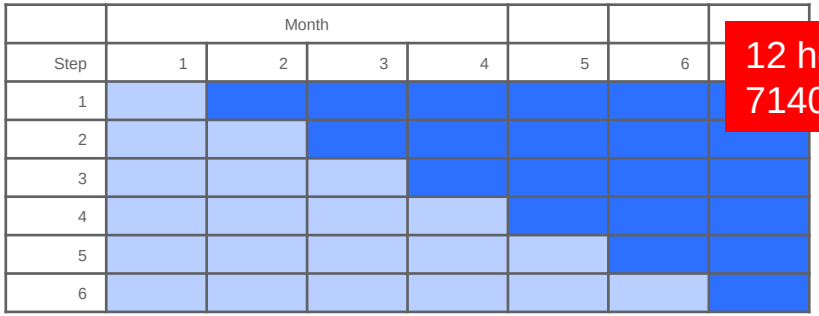
172 hospitals
14,620 patients

Parallel arm before and after



34 hospitals
5780 patients

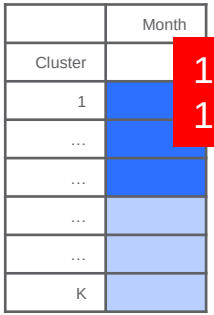
Stepped wedge



12 hospitals
7140 patients

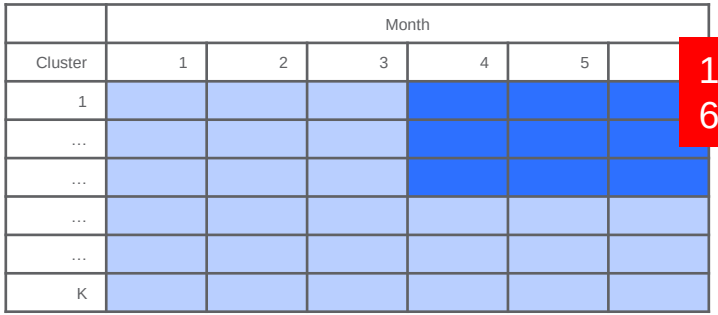
RETURNING TO THE EXAMPLE

Parallel arm



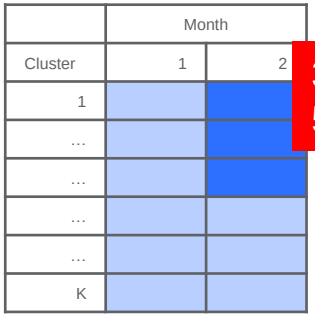
172 hospitals
14,620 patients

Parallel before and after



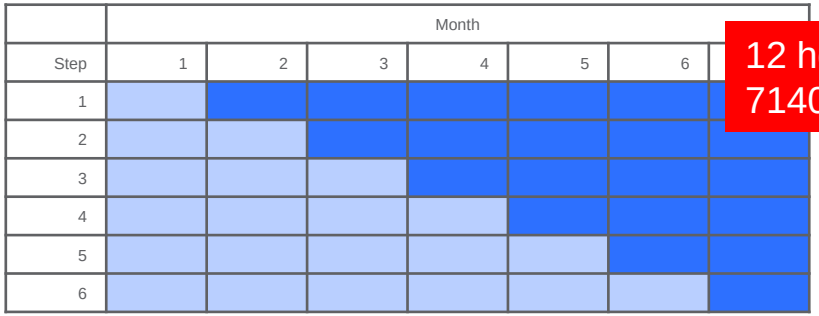
12 hospitals
6120 patients

Parallel arm before and after



34 hospitals
5780 patients

Stepped wedge



12 hospitals
7140 patients

4: TO SIMPLIFY LOGISTICS

- YES
  - Logistically challenging to implement intervention at many clusters at the same time
- NO
  - SW design has many logistical challenges
  - Consider parallel arm design with staggered implementation

RETURNING TO THE EXAMPLE

ORIGINAL RESEARCH ARTICLE

Point-of-Care Hemostatic Testing in Cardiac Surgery

A Stepped-Wedge Clustered Randomized Controlled Trial

BACKGROUND: Cardiac surgery is frequently complicated by coagulopathic bleeding that is difficult to optimally manage using standard hemostatic testing. We hypothesized that point-of-care hemostatic testing within the context of an integrated transfusion algorithm would improve the management of coagulopathy in cardiac surgery and thereby reduce blood transfusions.

METHODS: We conducted a pragmatic multicenter stepped-wedge cluster randomized controlled trial of a point-of-care–based transfusion algorithm in consecutive patients undergoing cardiac surgery with cardiopulmonary bypass at 12 hospitals from October 6, 2014, to May 1, 2015. Following a 1-month data collection at all participating hospitals, a transfusion algorithm incorporating point-of-care hemostatic testing was sequentially implemented at 2 hospitals at a time in 1-month intervals, with the implementation order randomly assigned. No other aspects of care were modified. The primary outcome was red blood cell transfusion from surgery to postoperative day 7. Other outcomes included transfusion of other blood products, major bleeding, and major complications. The analysis adjusted for secular time trends, within-hospital clustering, and patient-level risk factors. All outcomes and analyses were prespecified before study initiation.

Keyvan Karkouti, MD
Jeannie Callum, MD
Duminda N. Wijeysundera, MD, PhD
Vivek Rao, MD, PhD
Mark Crowther, MD
Hilary P. Grocott, MD
Ruxandra Pinto, PhD
Damon C. Scales, MD, PhD
TACS Investigators

Parallel CRT with staggered implementation

	Months			
Waves	3	6	9	12
1				
2				
3				

12 hospitals
6120 patients

5: TO REDUCE BIAS

- YES
  - It is *partially* true that each cluster serves as their own control
- NO
  - Intervention is confounded with time by design
  - SW-CRT brings many additional risks of bias

RETURNING TO THE EXAMPLE

ORIGINAL RESEARCH ARTICLE

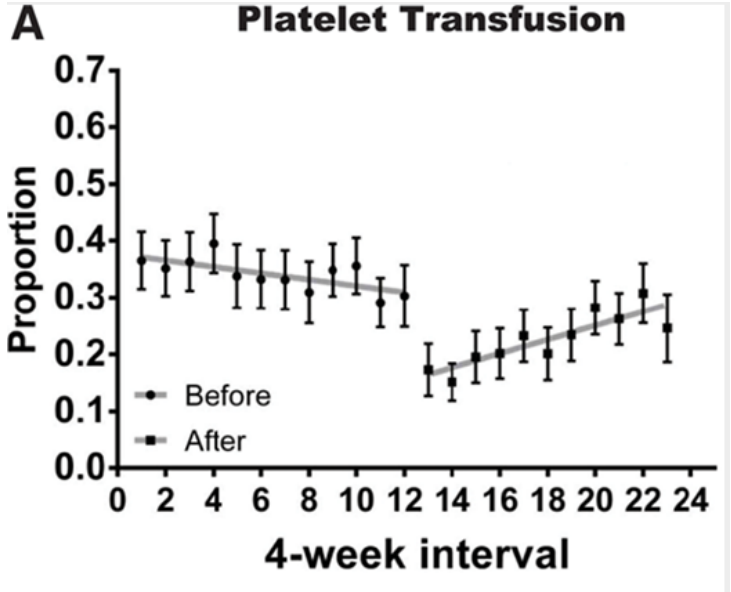
Point-of-Care Hemostatic Testing in Cardiac Surgery

A Stepped-Wedge Clustered Randomized Controlled Trial

BACKGROUND: Cardiac surgery is frequently complicated by coagulopathic bleeding that is difficult to optimally manage using standard hemostatic testing. We hypothesized that point-of-care hemostatic testing within the context of an integrated transfusion algorithm would improve the management of coagulopathy in cardiac surgery and thereby reduce blood transfusions.

METHODS: We conducted a pragmatic multicenter stepped-wedge cluster randomized controlled trial of a point-of-care–based transfusion algorithm in consecutive patients undergoing cardiac surgery with cardiopulmonary bypass at 12 hospitals from October 6, 2014, to May 1, 2015. Following a 1-month data collection at all participating hospitals, a transfusion algorithm incorporating point-of-care hemostatic testing was sequentially implemented at 2 hospitals at a time in 1-month intervals, with the implementation order randomly assigned. No other aspects of care were modified. The primary outcome was red blood cell transfusion from surgery to postoperative day 7. Other outcomes included transfusion of other blood products, major bleeding, and major complications. The analysis adjusted for secular time trends, within-hospital clustering, and patient-level risk factors. All outcomes and analyses were prespecified before study initiation.

Keyvan Karkouti, MD
Jeannie Callum, MD
Duminda N. Wijeysundera, MD, PhD
Vivek Rao, MD, PhD
Mark Crowther, MD
Hilary P. Grocott, MD
Ruxandra Pinto, PhD
Damon C. Scales, MD, PhD
TACS Investigators



Karkouti K, McCluskey SA, Callum J, Freedman J, Selby R, Timoumi T, Roy D, Rao V. Evaluation of a novel transfusion algorithm employing point-of-care coagulation assays in cardiac surgery: a retrospective cohort study with interrupted time-series analysis. Anesthesiology. 2015 Mar;122(3):560-70.

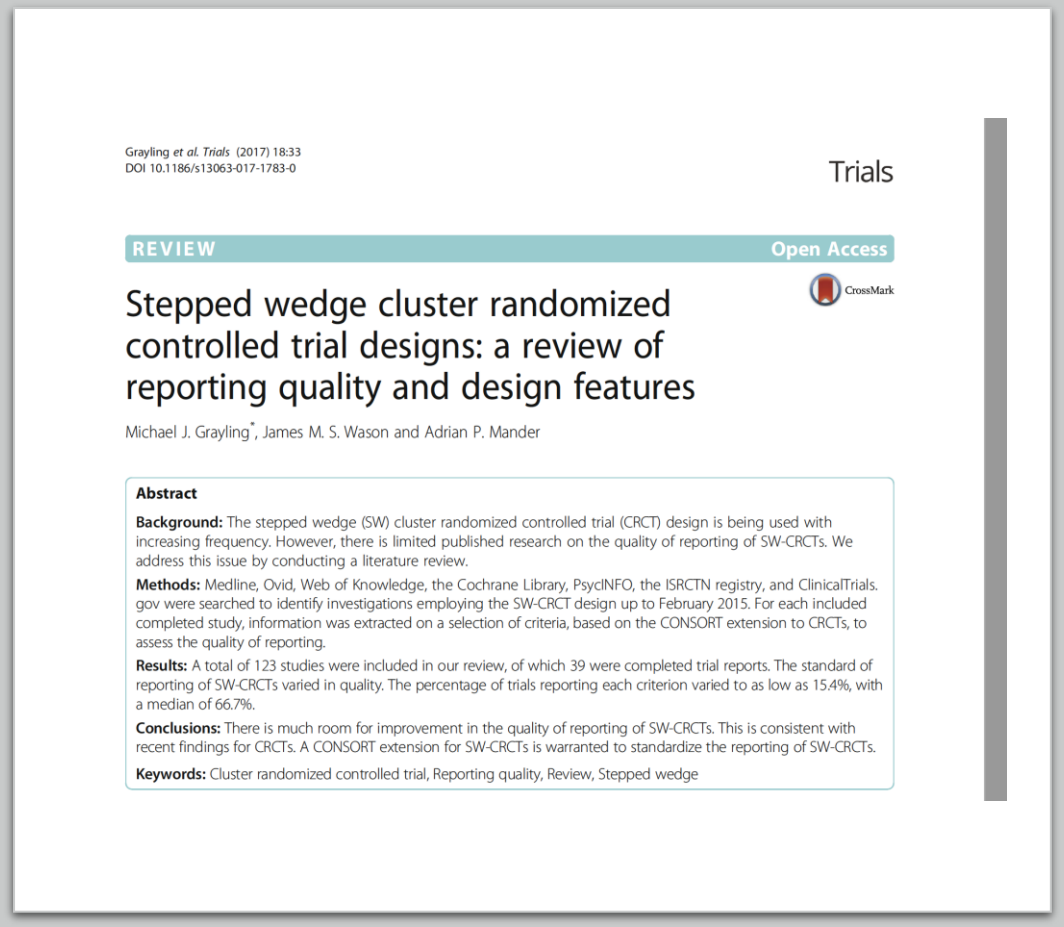
IN SUMMARY

- Most methodological arguments in favour of the SW-CRT have good counter-arguments
- Recommend working with an experienced trial statistician to empirically examine implications of alternative designs
- Choose the most scientifically robust design given the practical constraints of the study

Ethical justifications for SW-CRT

- David Magnus, PhD
- Thomas A Raffin Professor of Medicine and Biomedical Ethics and Professor of Pediatrics
- Director, Stanford Center for Biomedical Ethics
- Associate Dean of Research

"the most common reason given for the use of the SW-CRCT design was a desire for the intervention to be made available to all clusters by the end of the trial, on ethical or equity grounds (51/123, 41.5%)."



Grayling *et al. Trials* (2017) 18:33
DOI 10.1186/s13063-017-1783-0

Trials

REVIEW Open Access

Stepped wedge cluster randomized controlled trial designs: a review of reporting quality and design features

Michael J. Grayling*, James M. S. Wason and Adrian P. Mander

Abstract

Background: The stepped wedge (SW) cluster randomized controlled trial (CRCT) design is being used with increasing frequency. However, there is limited published research on the quality of reporting of SW-CRCTs. We address this issue by conducting a literature review.

Methods: Medline, Ovid, Web of Knowledge, the Cochrane Library, PsycINFO, the ISRCTN registry, and ClinicalTrials.gov were searched to identify investigations employing the SW-CRCT design up to February 2015. For each included completed study, information was extracted on a selection of criteria, based on the CONSORT extension to CRCTs, to assess the quality of reporting.

Results: A total of 123 studies were included in our review, of which 39 were completed trial reports. The standard of reporting of SW-CRCTs varied in quality. The percentage of trials reporting each criterion varied to as low as 15.4%, with a median of 66.7%.

Conclusions: There is much room for improvement in the quality of reporting of SW-CRCTs. This is consistent with recent findings for CRCTs. A CONSORT extension for SW-CRCTs is warranted to standardize the reporting of SW-CRCTs.

Keywords: Cluster randomized controlled trial, Reporting quality, Review, Stepped wedge

Brief ethical history

- SW-CR Trials were often used to evaluate implementation or evaluate efficacy of an intervention being rolled out
- The ethical argument for SW-CR became a major focus of debate during the Ebola crisis roughly a decade ago
- Still largely used in pragmatic trials, but echoes of ethics debate from Ebola remain

The ethical argument for SW-CR trials

Ebola context—ethical arguments against use of placebo in clinical trials—Caplan:

Comment

Morality in a time of Ebola

The first true epidemic of Ebola led to widespread panic. The virus appeared in so many countries in 2014—including Guinea, Liberia, Mali, Nigeria, Senegal, Sierra Leone, Spain, and the USA—that WHO officials at the

engineered vaccines from recombinant vesicular stomatitis virus, modified rabies vaccine; and use of antibodies obtained from the serum of Ebola survivors.[10]

The proliferation of possible solutions to the Ebola

Published Online
February 20, 2015
http://dx.doi.org/10.1016/
S0140-6736(14)61653-6

Clement Adebamowo, Oumou Baj-Sow, Fred Binka, et al, “Randomised controlled trials for Ebola: practical and ethical issues,” the Lancet, 384: 1423-4; 2014.

- *“When conventional care means such a high probability of death, it is problematic to insist on randomising patients to it when the intervention arm holds out at least the possibility of benefit.”*

- *“Populations who are terrified by the progress of the epidemic, and who lack trust in health-care and aid workers, and in public authorities in the aftermath of civil wars, cannot be expected to offer informed consent to* [placebo controlled] *randomized trials.”*

Emergency Ebola response: a new approach to the rapid design and development of vaccines against emerging diseases

Claire M Tully, Teresa Lambe, Sarah C Gilbert, Adrian V S Hill

Lancet Infect Dis 2015; 15: 356–59
Published Online January 14, 2015
http://dx.doi.org/10.1016/S1473-3099(14)71071-0

The epidemic of Ebola virus disease has spread at an alarming rate despite containment efforts. As a result, unprecedented large-scale international response efforts have been made in an attempt to gain control of the outbreak and reduce transmission. Several international consortia have been formed in a remarkable worldwide collaborative effort to expedite trials of two candidate Ebola virus vaccines: cAd3-EBOZ and rVSV-EBOV. In parallel, both vaccines are being manufactured in large amounts to enable future rapid deployment for management of the crisis.

"administration of placebo vaccine during a viral outbreak with a case-fatality rate of greater than 70% has not been done before, and raises serious ethical questions. An alternative trial design is a stepped wedge, which would compare rates of infection in vaccinated and unvaccinated groups."

In a stepped-wedge design trial, such as the Sierra Leone Trial to Introduce a vaccine against Ebola (STRIVE) trial of health-care workers in Sierra Leone, the Ebola vaccine is sequentially rolled out to participants in clusters, such as clinics or hospitals, throughout several time periods. By the end of the study, all participants will have received the intervention. When the intervention is expected to confer a large benefit, a stepped-wedge design mitigates the ethical dilemma of nontreatment, such as in the case of a parallel control group, or withdrawal of treatment as would occur in a standard crossover study.

Jolanta Piszczek, Eric Partlow, "Stepped-Wedge Trial Design to Evaluate Ebola Treatments," the Lancet Infect. Dis. 15: 762-2; 2015.

Arguments in favor of SW-CRT

- Belief that intervention provides more benefit than harm
- Therefore, there is no equipoise for a trial against placebo or standard care that is known to be ineffective
- Therefore, no standard RCT can ethically be conducted
- In addition, SW-CRT makes it possible for all participants to get the intervention
- Finally, may be required when resource limitations make parallel implementation impossible

To cite: Doussau A, Grady C. *J Med Ethics* 2016;**42**:797–804.

PAPER

Deciphering assumptions about stepped wedge designs: the case of Ebola vaccine research

Adélaïde Doussau, Christine Grady

Department of Bioethics, Clinical Center, National Institutes of Health, Bethesda, Maryland, USA

Correspondence to
Dr Adélaïde Doussau, Department of Bioethics, Clinical Center, National Institutes of Health, 10 Center Drive, Building 10/1C118, Bethesda, MD 20892-1156 USA; doussau2.adelaide@gmail.com

Received 2 December 2015
Revised 1 August 2016
Accepted 19 September 2016

ABSTRACT
Ethical concerns about randomising persons to a no-treatment arm in the context of Ebola epidemic led to consideration of alternative designs. The stepped wedge (SW) design, in which participants or clusters are randomised to receive an intervention at different time points, gained popularity. Common arguments in favour of using this design are (1) when an intervention is likely to do more good than harm, (2) all participants should receive the experimental intervention at some time point during the study and (3) the design might be preferable for practical reasons. We examine these assumptions when considering Ebola vaccine research. First, based on the claim that a stepped wedge design is indicated

systematic review of SW designs described their advantages as follows:

> "by the end of the study, all participants will have received the intervention, although the order in which participants receive the intervention is determined at random. The design is particularly relevant where it is predicted that the intervention will do more good than harm (making a parallel design, in which certain participants do not receive the intervention unethical) and/or where, for logistical, practical or financial reasons, it is impossible to deliver the intervention simultaneously to all participants".[8]

These major assumptions about SW designs—that

Counter arguments

- In Ebola, no evidence that benefits outweighed the harms of proposed therapeutics and vaccines.
- SW-CRT rarely used as a design for novel therapeutics or vaccines
- Confusion that all participants get the intervention in SW-CRT
  - Each cluster may get intervention, but depending upon research design, individual participants may not get the intervention (depending on when they receive supportive treatment)
- Even if resources limited, logistics may make SW-CRT in resource poor setting impractical

Extended essay

Delaying and withholding interventions: ethics and the stepped wedge trial

Ariella Binik[1,2]

[1]Department of Philosophy, McMaster University, Hamilton, Ontario, Canada
[2]Institute on Ethics & Policy for Innovation, McMaster University, Hamilton, Ontario, Canada

ABSTRACT
Ethics has been identified as a central reason for choosing the stepped wedge trial over other kinds of trial designs. The potential advantage of the stepped wedge design is that it provides all arms of the trial with the active intervention over the course of dire. But despite the increased use of the stepped wedge trial and appeals to its ethical superiority, there has been limited critical attention to this study design.[i]

In what follows, I examine whether there are persuasive reasons to think that the stepped wedge

To cite: Binik A. *J Med Ethics* 2019;**45**:662–667.

Arguments against ethical requirement for SW-CRT

- Arguments in favor often confuses individual *belief* in benefit of intervention with equipoise (which requires consensus in the field that the intervention is of benefit)
- If truly not in equipoise, delay in providing intervention no more justified than placebo—depending on strength of conviction of benefit, could undermine validation of trial at all
- True protection is clinical equipoise (properly understood) and it is neutral between SW-CRT and parallel research designs

Prost *et al. Trials* (2015) 16:351
DOI 10.1186/s13063-015-0837-4

TRIALS

RESEARCH Open Access

Logistic, ethical, and political dimensions of stepped wedge trials: critical review and case studies

Audrey Prost[1*], Ariella Binik[2], Ibrahim Abubakar[3], Anjana Roy[3], Manuela De Allegri[4], Christelle Mouchoux[5], Tobias Dreischulte[6], Helen Ayles[7], James J. Lewis[7] and David Osrin[1]

- *Evaluated 6 SW-CR public health trials in high, middle, and lower-income countries*

- *"the strongest arguments for a stepped wedge design are logistic and political rather than ethical."*
- *Equipoise still ethically required*
- *Largely justified when simultaneous roll out not possible*

Contemporary Clinical Trials 115 (2022) 106703

Contents lists available at ScienceDirect

Contemporary Clinical Trials

journal homepage: www.elsevier.com/locate/conclintrial

Ethical and epistemic issues in the design and conduct of pragmatic stepped-wedge cluster randomized clinical trials

Carole A. Federico [a], Patrick J. Heagerty [b], John Lantos [c], Pearl O'Rourke [d], Vasiliki Rahimzadeh [a], Jeremy Sugarman [e], Kevin Weinfurt [f], David Wendler [g], Benjamin S. Wilfond [h], David Magnus [a,*]

[a] Stanford Center for Biomedical Ethics, Stanford University, Stanford, CA 94305, USA
[b] Department of Biostatistics, University of Washington, Seattle, WA 98185, USA
[c] Children's Mercy Hospital Bioethics Center, University of Missouri-Kansas City, Kansas City, MO 64108, USA
[d] Harvard Medical School, Boston, MA 02115, / USA
[e] Berman Institute of Bioethics, Johns Hopkins University, Baltimore, MD 21218, USA
[f] Center for Health Measurement, Duke University, Durham, NC 27701, USA
[g] Department of Bioethics, NIH Clinical Center, Bethesda, MD 20892, USA
[h] Treuman Katz Center for Pediatric Bioethics, Seattle Children's Research Institute, Seattle, WA 98185, USA

- "We found that decisions to use SW-CRT design were justified by practical and epistemic reasons rather than ethical ones."

- Evaluation of 5 SW-CRT pragmatic demonstration trials funded by the NIH Collaboratory
- Concerns about feasibility of simultaneous roll out a major factor
- Hope that SW-CRT design could overcome cluster heterogeneity that would likely doom parallel CRT
- Desire to combine research with implementation (on assumption that intervention going to happen anyway
    - This last reason arguably makes equipoise mistake

Conclusion

- The primary ethical arguments in favor of SW-CRT fail to justify why such trials are necessary or superior to other designs
- Just as CRT's sometimes seen as desirable for avoiding informed consent requirements, there is potential that this could be used as justification for SW-CRT (perhaps to a greater extent)
- Just as this is an inadequate rationale for parallel CRT, it is also inadequate for SW-CRT
- Even if regulatory requirements for waiver of consent are met in a SW-CRT, there may be obligations to be transparent and to disclose information to participants about the trial