

#### UNIVERSITY<sup>of</sup> BIRMINGHAM

# QUALITY OF STEPPED WEDGE TRIAL REPORTING CAN BE RELIABLY ASSESSED USING A NEW REPORTING GUIDELINE AND A CROWD-SOURCING REVIEW

NOVEMBER 2018

KARLA HEMMING INSTITUTE OF APPLIED HEALTH RESEARCH, UNIVERSITY OF BIRMINGHAM, UK.



#### UNIVERSITY<sup>OF</sup> BIRMINGHAM

# BACKGROUND

- The Consolidated Standards Of Reporting Trials (CONSORT) extension for the Stepped-Wedge Cluster Randomised Trial (SW-CRT) is a recently published reporting guideline.
- The SW-CRT includes several design characteristics which make it different from the conventional cluster trial.
- Assessments of the quality of reporting of SW-CRTs against the CONSORT statement for CRTs have demonstrated poor reporting.
- Understandably, there have been no assessments of the quality of reporting according to the new CONSORT for SW-CRTs.

Grayling MJ, Wason JM, Mander AP. Stepped wedge cluster randomized controlled trial designs: a review of reporting quality and design features. Trials. 2017 Jan 21;18(1):33.

# **OBJECTIVES**

- Assess the quality of reporting of a recent sample of SW-CRTs according to the newly developed reporting guideline.
  - This assessment provides both a document of current reporting and will serve as a baseline assessment for any future study to identify any improvements over time.
- Determine if a crowd-sourcing type review is feasible to implement and reliable.
  - If the crowd-sourcing methodology is reliable it has the potential to be used to our benefit at the *current developments in cluster randomised trials and stepped wedge designs meeting.*

Hemming K, Taljaard M, McKenzie JE, Hooper R, Copas A, Thompson JA, Dixon-Woods M, Aldcroft A, Doussau A, Grayling M, Kristunas C, Goldstein CE, Campbell MK, Girling A, Eldridge S, Campbell MJ, Lilford RJ, Weijer C, Forbes AB, Grimshaw JM. Reporting of stepped wedge cluster randomised trials: extension of the CONSORT 2010 statement with explanation and elaboration. BMJ. 2018 Nov 9;363:k1614. doi: 10.1136/bmj.k1614.

### **OVERVIEW OF METHODS**

- Identified the most recently published SW-CRT trial reports.
- Randomly allocated for quality assessment by participants who were attending current developments in cluster randomised trials and stepped wedge designs meeting, London, UK during November 2017.
- We expected approximately 50 participants to attend this workshop.
- To allow for independent extraction by two (with content expertise) reviewers per report, and to allow for the possibility of more participants than expected attending the workshop we sought to identify approximately 30 studies.

Citizen science: crowdsourcing for systematic reviews Lucy Strang and Rebecca K. Simmons; Published by: The Healthcare Improvement Studies Institute 2018 ISBN: 978-1-9996539-1-0

# **SCOPE OF REVIEW**

- Included SW-CRTs:
  - Minimum of 3 sequences of allocations to periods spent in the control condition followed by periods in the intervention condition.
  - Two treatment conditions.
  - Cluster randomisation.
- Publication types:
  - Primary reports of SW-CRTs, i.e., protocols and reports of secondary analyses of a
    previously reported trial were excluded.
  - Open access or viewable from either the University of Birmingham or University of Ottawa libraries.
  - Published in English.

### **SEARCH PROCESS**

- Objective was to identify the 30 most recently published SW-CRTs up to November 2017.
- We identified eligible studies in PubMed using a previously published search strategy [Martin 2016] run on the 21<sup>st</sup> November 2017.
- We identified and ordered studies by date listed in Medline.
- To allow for exclusion of ineligible studies, titles and abstracts of the first 50 studies were screened in duplicate and independently by KH and MT and discussed to obtain a consensus on eligibility.
- Full copies of the reports were then obtained and assessed against the inclusion criteria, again in duplicate, identifying the required 30 full study reports.

Martin J, Taljaard M, Girling A, Hemming K. Systematic review finds major deficiencies in sample size methodology and reporting for stepped-wedge cluster randomised trials. BMJ Open. 2016 Feb 4;6(2):e010166.

# **CONSENT PROCESS**

- All registered delegates invited by email to participate in the review.
- Participants were informed that attending the workshop would require undertaking a small amount of work in advance and during the event, with an invitation to contribute to the resulting manuscript as a *group author*.
- Anyone not wishing to participate was invited to opt out.
- Participants were also informed that data on inter-rater reliability as well as the quality of reporting would be evaluated (anonymously).

# DATA ABSTRACTION PROCESS

- Participants not opting out were *randomly* allocated to one of the selected studies using computer-generated numbers, so that two participants were allocated to each study.
- One week in advance of the meeting, a full PDF copy of their allocated study was e-mailed to participants, along with a simple quality assessment tool.
- Participants were asked to *independently* assess their allocated study using this quality assessment tool.
- Participants were kept *blind* to the other allocated assessor of the same report until the morning of the workshop.
- After a 30 minute summary of the background to the project, the two participants assigned to each report met over a 30 minute period, discussed their discrepancies and reached a consensus.
- Data were therefore abstracted independently and in *duplicate*.

# DATA ABSTRACTED

- Participants were asked to assess the quality of reporting for each of the 26 items according to a four-point scale:
  - Clearly reported in full / clearly but partially reported / unclearly reported / not reported.
  - Feedback after the independent assessment resulted in a change to a five-point scale with the addition of "not applicable" for the joint assessment.
- The data were entered into an Excel database by one person (KH).

# **STATISTICAL ANALYSIS**

- The inter-rater reliability:
  - Percentage agreement (*within item across pairs*) and the Gwet A1 statistic using the *kappaetc* command in Stata 14.
  - Using the four-point scale and by dichotomising the four-point scale into a twopoint scale( clearly reported versus not clearly reported).
- The joint assessment of the quality of reporting for all 26 items:
  - Average number of items clearly or fully reported.

Excluded:

Independent reviews if only one of the pair submitted their independent assessments (n=3).

Any individual level items which were missing in either one or both independent assessments.

Any assessments of not applicable because this was not included in the independent assessment tool.



**FLOW OF STUDIES** 

11

PARTICIPANTS 0 ×0



### **RESULTS: RELIABILITY**



Blue: Objective items? Red: Subjective items?

#### **RESULTS: JOINT ASSESSMENT OF REPORTING**



# **CONCLUSIONS: REPORTING OF SW-CRTS**

- To improve quality of reporting, authors of SW-CRTs should carefully ensure reporting of all the minimal items as described in the CONSORT extension for SW-CRTs.
- Particular attention should be given to ensure clear reporting on:
  - The exact format of the design with justification;
  - How the clusters and individuals were identified for inclusion in the study,
  - Whether this was done before or after randomisation of the clusters both of which are crucial to the assessment of risks of bias.

Caille A, Kerry S, Tavernier E, Leyrat C, Eldridge S, Giraudeau B. Timeline cluster: a graphical tool to identify risk of bias in cluster randomised trials. BMJ. 2016 Aug 16;354:i4291.

#### **CONCLUSIONS: THE FEASIBILITY OF USING CROWD-SOURCING**

- Feasible and rated by last years participants as the best session at the meeting.
- Agreement between reviewers was low for some items.
- Possible explanations:
  - Wording in extension statement might be not so clear or *irrelevant item* (i.e. why the trial stopped)?
  - More training for consistency in reviewing or simply *difficult to assess* (i.e. "interpretation")?
  - Or, inconsistency in reviewing common place just not so often assessed?

Robson RC, Pham B, Hwee J, Thomas SM, Rios P, Page MJ, Tricco AC. Few studies exist examining methods for selecting studies, abstracting data, and appraising quality in a systematic review. J Clin Epidemiol. 2018 Oct 9. pii: S0895-4356(18)30165-3.6

# LIMITATIONS

- Convenience sample vs random sample of a well defined cohort (i.e. one calendar year)?
- Abstract screen was probably specific but lacked sensitivity:
  - For example, missed anything that was self defined as a wait list design.
- Small sample size (limited to 25 studies).
- Basic implementation.

#### More sophisticated systems

Inclusion decisions	-	2018-1	2018-11-12: Conflicts BIRGOFF			
Undecided 0		Showing 1 to	Showing 1 to 10 of 95 unique entries			
Included	22	Date	Title	≜ Aut		
Excluded	68			,		
Conflict	5	2017-01-0	te wrong study design Conference proceedings Can the use of computer decision support system prevent complicated ulcer among	patie <sub>Petersen</sub> , J. M.		
Decision by	-	2017-01-0	khemming wrong outcome Effects of stepped psychooncological care on referral to psychosocial services and emotional well-bei	ng in Singer, S.; Da		
<u>Mr. John Gallis</u> <u>k.hemming@bham.ac.uk</u> <u>Liz Turner</u>		2017-01-0	khemming Liz Primary outcome not binary Primary binary outcome: primary outcomes: tobacco (ever) Effectiveness of a pragmatic school-ba	ased Hodder, R. K.;		
Search methods (Add new)	č -	2017-01-0	khemming wrong outcome Effects of Enhancing School-Based Body Mass Index Screening Reports with Parent Education on Rep	ort Ut		
Uploaded References [Conflict.txt]	95	2017-01-0	khemming wrong outcome Antenatal nutrition behaviour change communication: change in birthweight in rural Bangladesh	Chowdhury, M		
Keywords for include [Add new]	-	2017-01-0	khemming wrong outcome Effect of participatory women's groups and counselling through home visits on children's linear grow	<b>th in r</b> <sub>Nair,</sub> N.; Tripa		
<u>trial</u> randomized	84 û 71 û	2017-01-0	khemming Factorial CRT The Effect of Nursing Quality Improvement and Mobile Health Interventions on Infant Sleep Practices:	a <b>Ran</b> <sub>Moon</sub> , R. Y.; H		
controlled trial randomized controlled trial	50 U 33 Ū	2017-01-0	khemming Conference proceedings Effects of discontinuing long-term opioid therapy in patients with chronic pain	Husain, J. M.;		
randomly	32 0					
randomised compared with	2/  23					
randomised controlled trial	150	4				
randomly assigned	14 Û	Include	Plactide PExclude Reason Label Vightights ON D Upload PDF full-texts			
randomly allocated	10 û	<b>F</b> ((		the state of the s		
More >>		trial	or stepped psychooncological care on referral to psychosocial services and emotional well-being in cancer patiel	nts: a cluster- <mark>ra</mark>		
Keywords for exclude [Add new]	-	METHOD need for	3: In a cluster-randomized trial, wards were randomly allocated to stepped versus standard care. Stepped care comprised screening for distress, consult CL services, and provision of CL service. Primary outcomes were referral to psychosocial services and emotional well-being half a year after baseline, meas	tation between doctor a sured with the Hospital		
prevalence	18 0	secondar	rendpoint was uptake of outpatient health care. Analysis employed mixed effects multivariate regression modeling. RESULTS: Thirteen wards were rando	mized ; 1012 patients		
survey	12 0	(N = 570	(N = 570; 7 wards), 22% of the patients were referred to CL services and 3% with standard care (N = 442; 6 wards; odds ratio [OR] 10.0; P < .001). Well-being 6 months after baseline			
rross-sectional	5 =	341) and	9:4 alter statituaru care (in = 2.34, p -0.3) Y = ./1). Arter scepped care, padents with psychiatric comorbidity went more often to psychotherapists (UK 4.0, nationite without comorbidity used neurohistricts less often (OR 0.4, P = .04) than in standard care. CONCLUSIONS: Stepped care resulted in better referval	, r = .05) and to psych		
regression analysis	30	not impr	not improved, but uptake of outpatient psychiatric help was increased in patients with psychiatric comorbidity and decreased in patients without. OBJECTIVE: Emotional distress in cancer			
observational	3 🗊	daily rou	daily routine; therefore, distress screening is now recommended in many national guidelines. However, screening alone does not necessarily translate into better well-being. We examines			

https://rayyan.qcri.org/

# DISCUSSION

- Was the ethical oversight and consent process sufficient?
- What should we call the group authorship?
- Should we continue to use this approach in future meetings?
  - How to prioritise topics?
- Can the methodology be improved?
  - Assessment of reliability of joint assessment?

#### THE GOOD NEWS...

Tentatively accepted by JCE (Journal of Clinical Epidemiology)

JCCE       Mobile         JOURNAL OF CLINICAL EPIDEMIOLOGY       Subscription						
Articles and Issues - Collections For Authors - Journal Info - Subscribe More Periodicals -						
All Content     Search         Advanced Search						
On the Cover	Current Issue December 2018 Volume 104	Access this journal on ScienceDirect				
<page-header><image/><image/><image/><image/><image/><image/><section-header></section-header></page-header>	Featured Article       Image: Second Se	Twitter Follow Journal of Clinical Epidemiology on <u>Twitter</u> ! Tweet us at @JClinEpi.				
	Articles in Press Most Read Most Cited	David Sackett Young Investigator Award				
Subscribe to Journal	The methodological quality of dose-response meta-analyses needed substantial improvement: A cross-sectional survey and proposed recommendations Chang Xu, Yu Liu, Peng-Li Jia, Ling Li, Tong-Zu Liu, Liang- Liang Cheng, Ke Deng, A.S.M. Borhan, Lehana Thabane, Xin Sun	JCE's annual David Sackett Young Investigator Award is in the spirit of the late David L. Sackett, who over many decades and in numerous ways continuously inspired and educated generations of young investigators in the fields of clinical epidemiology and evidence-based medicine. We				

# ACKNOWLEDGEMENTS

 K Carroll, J Thompson, A Forbes and M Taljaard and the SW-CRT review group

# **ACKNOWLEDGEMENTS GROUP AUTHORS**

Susan Dutton OCTRU, University of Oxford susan.dutton@csm.ox.ac.uk; Vichithranie Madurasinghe QMUL v.madurasinghe@gmul.ac.uk; Katy Morgan LSHTM katy.morgan@lshtm.ac.uk; Beth Stuart University of Southampton bls1@soton.ac.uk; Katherine Fielding LSHTM katherine.fielding@lshtm.ac.uk; Victoria Cornelius Imperial College London v.cornelius@imperial.ac.uk; Liz Turner Duke University liz.turner@duke.edu; Richard Hooper QMUL r.l.hooper@gmul.ac.uk; Bruno Giraudeau INSERM U1246, France bruno.giraudeau@univ-tours.fr; Catey Bunce King's College London ediz.hussein@kcl.ac.uk; Paul Seed King's College London paul.seed@kcl.ac.uk; Alecia Nickless University of Oxford alecia.nickless@phc.ox.ac.uk; Michael Grayling MRC Biostatistics Unit mjg211@cam.ac.uk; Mélanie Pragu Bordeaux University, Inria, Inserm U1219, France melanie.prague@inria.fr; Sally Kerry QMUL s.m.kerry@qmul.ac.uk; Lauren Bell LSHTM Lauren.Bell@Ishtm.ac.uk; Eila Watson Oxford Brookes University ewatson@brookes.ac.uk; Rafael Gafoor King's College London rafael.gafoor@kcl.ac.uk; Nadine Marlin QMUL n.marlin@qmul.ac.uk; Emel Yorganci King's College London Cicely Saunders Institute emel.yorganci@kcl.ac.uk; Lesley Smith Oxford Brookes University lesleysmith@brookes.ac.uk; Murielle Mbekwe INSERM U1246, Tours, France muriellembekwe@yahoo.fr; Steven Teerenstra Radboud University Nijmegen z824116@umcn.nl; Claire Chan QMUL c.l.chan@qmul.ac.uk; Mirjam Moerbeek Utrecht University m.moerbeek@uu.nl; Pamela Jacobsen King's College London pamela.jacobsen@kcl.ac.uk; Simon Bond Cambridge Ćlinical trials Unit simon.bond@addenbrookes.nhs.uk; Agnès Caille INSERM U1246, France agnes.caille@univ-tours.fr; Ben Jones Plymouth University ben.jones@plymouth.ac.uk; John Preisser University of North Carolina, USA jpreisse@bios.unc.edu; Mona Kanaan University of York mona.kanaan@york.ac.uk; Catherine Hewitt University of York catherine.hewitt@york.ac.uk; Tracy Pellatt-Higgins University of Kent t.pellatt-higgins@kent.ac.uk: Laura Pankhurst NHS Blood and Transplant laura.pankhurst@nhsbt.nhs.uk; Schadrac Agbla LSHTM schadrac.agbla@lshtm.ac.uk; Sandra Eldridge QMUL s.eldridge@qmul.ac.uk; Robin Lerner University of Oxford robin.lerner@ndorms.ox.ac.uk; Clémence Leyrat LSHTM clemence.leyrat@lshtm.ac.uk; Mark Pilling University of Cambridge mark.pilling@medschl.cam.ac.uk; Julia Forman King's College London julia.forman@kcl.ac.uk; Indrani Bhattacharya ICR-CTSU indrani.bhattacharya@icr.ac.uk; Lesley Smith Oxford Brookes University lesleysmith@brookes.ac.uk; Nicholas Magill King's College London nicholas.magill@kcl.ac.uk; Jane Candlish University of Sheffield jane.candlish@gmail.com; Cliona McDowell NICTU cliona.mcdowell@nictu.hscni.net; James Martin Birmingham University J.T.Martin@bham.ac.uk; Louise Linsell National Perinatal Epidemiology Unit louise.linsell@npeu.ox.ac.uk; Caroline Kristunas University of Leicester cak21@le.ac.uk; Elizabeth Allen LSHTM elizabeth.allen@lshtm.ac.uk; Nadine Seward LSHTM nadine.seward@lshtm.ac.uk; Elaine Nicholls Keele University e.nicholls@keele.ac.uk; Bryony Franklin UCL School of Pharmacy bryony.deanfranklin@ucl.ac.uk;

