

RESEARCH

Open Access



Strategies to minimise and monitor biases and imbalances by arm in surgical cluster randomised trials: evidence from ChEETAh, a trial in seven low- and middle-income countries

NIHR Global Research Health Unit on Global Surgery^{1*} and The Hospital Principle Investigator

Abstract

Background Cluster randomised controlled trials (cRCT) present challenges regarding risks of bias and chance imbalances by arm. This paper reports strategies to minimise and monitor biases and imbalances in the ChEETAh cRCT.

Methods ChEETAh was an international cRCT (hospitals as clusters) evaluating whether changing sterile gloves and instruments prior to abdominal wound closure reduces surgical site infection at 30 days postoperative. ChEETAh planned to recruit 12,800 consecutive patients from 64 hospitals in seven low-middle income countries. Eight strategies to minimise and monitor bias were pre-specified: (1) minimum of 4 hospitals per country; (2) pre-randomisation identification of units of exposure (operating theatres, lists, teams or sessions) within clusters; (3) minimisation of randomisation by country and hospital type; (4) site training delivered *after* randomisation; (5) dedicated 'warm-up week' to train teams; (6) trial specific sticker and patient register to monitor consecutive patient identification; (7) monitoring characteristics of patients and units of exposure; and (8) low-burden outcome-assessment.

Results This analysis includes 10,686 patients from 70 clusters. The results aligned to the eight strategies were (1) 6 out of 7 countries included ≥ 4 hospitals; (2) 87.1% (61/70) of hospitals maintained their planned operating theatres (82% [27/33] and 92% [34/37] in the intervention and control arms); (3) minimisation maintained balance of key factors in both arms; (4) post-randomisation training was conducted for all hospitals; (5) the 'warm-up week' was conducted at all sites, and feedback used to refine processes; (6) the sticker and trial register were maintained, with an overall inclusion of 98.1% (10,686/10,894) of eligible patients; (7) monitoring allowed swift identification of problems in patient inclusion and key patient characteristics were reported: malignancy (20.3% intervention vs 12.6% control), midline incisions (68.4% vs 58.9%) and elective surgery (52.4% vs 42.6%); and (8) 0.4% (41/9187) of patients refused consent for outcome assessment.

Collaborators are listed in Additional file 1.

*Correspondence:

NIHR Global Research Health Unit on Global Surgery
o.omar@bham.ac.uk

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Conclusion cRCTs in surgery have several potential sources of bias that include varying units of exposure and the need for consecutive inclusion of all eligible patients across complex settings. We report a system that monitored and minimised the risks of bias and imbalances by arm, with important lessons for future cRCTs within hospitals.

Keywords Surgical site infection, Abdominal surgery, Global surgery, Global health, Cluster randomised controlled trial, Bias, Research methodology, Quality assurance, Trial management

Introduction

Surgical innovations often involve complex, across-team interventions that require behavioural change [1–3]. Evaluation of these innovations in a randomised trial requires a cluster randomised design because of the high risk of contamination between intervention and control arms with individual patient randomisation, as well as the logistical and practical issues around the delivery of the intervention [4–6]. However, few cluster randomised controlled trials (cRCTs) have been conducted to date in surgery [7]. As a result, methods for high-quality delivery of cRCTs in surgical settings are still evolving [8, 9].

A major methodological challenge in cRCTs is minimising bias and arm imbalance. In cluster randomisation, chance imbalances in patient characteristics are likely to occur [10]. This can lead to a risk confounding bias. Selection bias occurs where there is incomplete identification and recruitment of eligible patients within a cluster overall. This can lead to a sample that is unrepresentative of the target population compromising external validity and/or an unfair comparison of the trial arms due to the differences in known and unknown confounders compromising internal validity. In addition, the lack of allocation concealment can also impact cluster size variability resulting from selection bias, causing further imbalances.

In surgical trials with cluster randomisation, there are specific challenges. The risk of bias is raised when it is necessary to unmask the clusters to their randomised group prior to the recruitment of eligible participants (i.e. where allocation concealment is not possible). In the context of a surgical cRCT, recruitment of participants occurs dynamically over a period of time and in different settings (e.g. preoperative clinic, intraoperatively in theatre), when undergoing an eligible procedure. In this context, unmasking of the cluster to their randomised group prior to participant recruitment is required for the purposes of training and delivery of the randomised allocation. In addition, there may be multiple units of exposure (e.g. operating rooms, teams, theatre lists) within a cluster representing a major source of variability, and many interdisciplinary team members are involved in the perioperative care pathway.

Despite these risks, strategies to minimise bias and imbalance have been poorly described and inconsistently

reported to date [4, 11]. This paper describes strategies to minimise bias and imbalance by arm used in a global surgery cRCT and transparently reports their implementation and effectiveness.

Methods

Trial aims, design, and setting

ChEETAH was an international, multicentre, 2-arm, cRCT with an internal pilot [12]. It evaluated the use of separate sterile gloves and instruments before closing the abdominal wall to reduce the rate of surgical site infection in the 30 days after surgery. Patients undergoing any abdominal surgery, for any indication with an abdominal incision ≥ 5 cm were eligible, except for caesarean sections. The primary outcome was surgical site infection (SSI) at 30 days, based on a US Centers for Disease Control definition of SSI [13]. Overall, ChEETAH planned to recruit 12,800 patients in at least 64 clusters. UK ethical approval was obtained from the University of Birmingham International Research Ethics Committee. All individual participating countries obtained local or national ethical approval for ChEETAH in accordance with local requirements (available upon request). Individual patient-level consent for exposure to the intervention or control (routine practice) was deemed not to be required, so patients confirmed their consent (written or fingerprint) prior to discharge for inclusion in data collection at 30 days postoperatively.

Clusters were defined at the level of the hospital, from seven low- and middle-income countries (Benin, Ghana, India, Mexico, Nigeria, Rwanda, and South Africa). Hospitals were randomised (1:1) between (i) intervention (change of gloves and use of separate, sterile instruments) and (ii) current routine hospital practice (no change of gloves or use of separate, sterile instruments) *before* closing the abdominal wall. Low- and middle-income status was defined by the Development Assistance Committee (DAC) Official Development Assistance (ODA) list. Any cluster (hospital) in those LMICs that performed elective and/or emergency abdominal surgery and where glove and instrument change was not routine practice were eligible to participate. As the intervention and control require whole-team implementation, site investigators were not blinded, but patients were blinded to their randomisation status. This study was a pre-planned analysis

of ChEETAH trial data to describe the strategies to minimise and monitor biases and chance imbalances and inform future cRCTs in surgery. DMC and TSC approval was obtained for the publication of this data.

Trial structure and processes

In each randomised hospital and their pre-specified operating rooms, a bespoke local pathway was developed to recruit all eligible patients. Potentially eligible patients could be identified by any member of the surgical team (research nurse, clinical officer, surgeon in training, operating surgeon), either before, during or after surgery but before discharge.

Pre-defined strategies to minimise bias and chance imbalances

This cRCT protocol adopted eight strategies to monitor and minimise bias and imbalance. The strategies are ordered and aligned to the relevant stage of the trial pathway (Fig. 1). These focused on each of the three potential sources of bias based on case-mix variability: (1) hospitals; (2) units of exposure (operating theatres, surgical teams, and/or theatre lists); and (3) patients.

Strategy 1

Hospital-level: The protocol required a minimum of 4 hospitals randomised per participating LMIC to ensure balance in a number of clusters within each country.

Strategy 2

Unit of exposure level: Hospitals (clusters) were required to pre-specify their predicted participating units of exposure *prior to* randomisation (elective-only operating theatre/emergency-only operating theatre/elective and emergency (mixed) theatre). This was prospectively monitored, and any deviation in actual units of exposure post-randomisation were queried and recorded, with a direct intervention by the Trial Management Group (TMG) where feasible. This strategy aimed to prevent sites from modifying their planned case mix after knowledge of their randomised allocation. We report the proportion of hospitals that maintained their pre-specified operating theatre case mix overall and by randomised allocation.

Strategy 3

Hospital-level: Randomisation was minimised by country and by hospital type (i.e. hospital that accepts pre-operative referrals from other surgical teams (referral hospital)

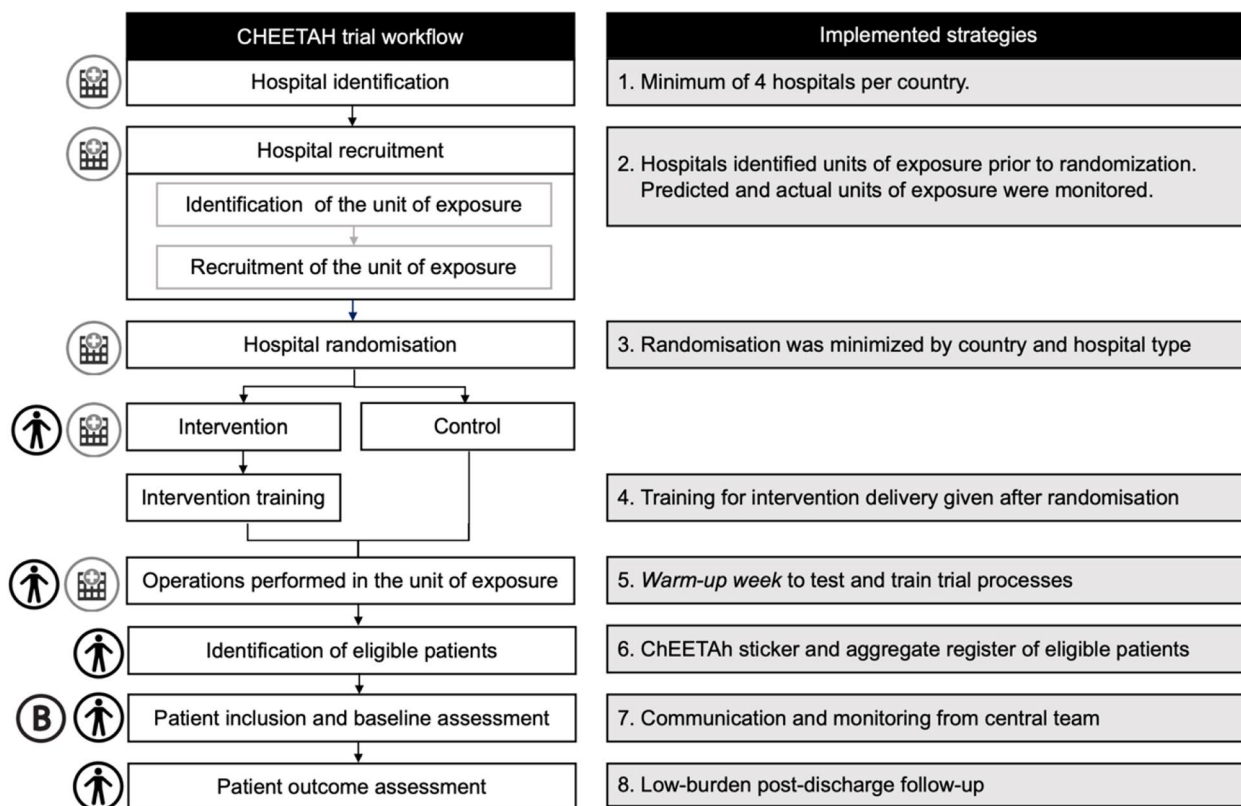


Fig. 1 Flowchart of the trial processes and strategies to minimise bias and imbalance by arm

or not (non-referral hospital)) to force balance in randomisation allocation for these key characteristics. We reported the balance of intervention to control groups across the minimisation characteristics.

Strategy 4

Hospital-level: Hospitals and investigators were given a training on the intervention or control group allocation *after* randomisation to minimise contamination between the arms. Delivering training to hospitals prior to randomisation would have required the inclusion of information about the delivery of both the intervention and control groups. Training was conducted with multiple stakeholders across research hubs (national lead sites) and spokes (linked participating hospital), often across multiple theatre teams within each cluster. We reported the proportion of participating hospitals that received training after randomisation.

Strategy 5

Hospital-level: A dedicated 'warm-up week' was conducted by each hospital investigator team to establish and test processes for patient identification, including the use of the trial-specific register and the ChEETAH trial operation sticker (Additional file 1: Appendix C, D). We reported the number of clusters that completed a warm-up week.

Strategy 6

Patient-level: The identification of patients within clusters was ensured using a ChEETAH trial sticker in the clinical notes of all patients undergoing abdominal surgery in participating units of exposure and monitored with a dedicated in-theatre trial-specific register where patient eligibility and inclusion were both recorded. This information was held at sites and aggregated periodically, for both site monitoring against theatre logbooks and periodic reporting to the central co-ordinating team. The central team would then review the aggregate register (Additional file 1: Appendix E) and the cases uploaded onto REDCap, to identify any inappropriate patient exclusions and mitigate against resulting bias and imbalances. The number of eligible patients identified and included in the cRCT divided by the number of potentially eligible patients recorded in the aggregate register was used to calculate a case ascertainment rate summarised as a percentage with 95% confidence intervals.

Strategy 7

Hospital-level: Regular communication between the sites and the central coordinating team was maintained throughout the trial. This focused on monitoring key characteristics of patients: Patient level—age

(\geq / $<$ 18 years old), sex (male/female), urgency of operation (elective/emergency), operative approach (open midline/open non-midline), indication for surgery (malignant/benign).

These were reported monthly by the hospital to the TMG meeting to prospectively monitor and intervene in any hospitals where the quality assurance rules were not being followed and to the Data Monitoring Committee (DMC) and Trial Steering Committee (TSC) in regular reports. This strategy aimed to identify early any indication that units of exposure and patients may have been recruited selectively (i.e. unrepresentative sample, unusual or unexplained imbalance between the randomisation arms). We report the impact of our communication strategy in identifying and intervening on participating sites. This included routine central monitoring of the aggregate register, sharing lessons learned across the network and corrective and preventative measures around training and provision of additional guidance to all sites reinforcing several key trial processes (more details in Additional file 1: Appendix A).

Strategy 8

Patient-level: A pragmatic, low-burden post-discharge follow-up schedule was designed with support from patient and public representatives, to maximise feasibility and minimise selective outcome reporting in the trial. Refusal of consent for outcome assessment was reported overall across key risk subgroups and by trial arm.

Data management and governance

No patient-level outcome data (e.g. SSI rates) was seen by the TMG during trial conduct, nor is it included in this publication. Clusters were pseudo-anonymised by the use of a hospital ID for presentation in this analysis. Reporting of this process was pre-defined in the published study protocol [12]. Summary data were described using summary statistics in Stata V17.0 (Stata Corporation).

Results

Data were included in this analysis from 10,686 patients undergoing surgery in 70 hospitals in four countries. Summary data is presented below for each of the eight strategies to monitor and minimise bias and imbalances (Fig. 1).

Strategy 1

Six countries had met their minimum requirement for a number of participating centres (India, 21 hospitals; Nigeria, 16 hospitals; Rwanda, 12 hospitals; Ghana, 10 hospitals; Benin, 5 hospitals; Mexico, 4 hospitals) and one was below this target (South Africa, 2 hospitals).

Strategy 2

Of all hospitals included in the analysis, 61 of 70 (87.1%, 95% CI 77.0 to 93.9%) collected data from patients in the same number of elective-only, emergency-only and mixed elective-emergency units of exposure that they had predicted prior to randomisation (Fig. S4). There were no occasions where these hospitals changed their specified elective or emergency units of exposure (i.e. swapping one elective theatre for another in the same hospital when examined by randomisation allocation (Fig. 2)); there was appropriate balance between the intervention and control arm in the number of elective (37.3%, 95% CI 28.2 to 47.0% [41/110] versus 49.3%, 95% CI 40.7 to 57.9% [69/140]), emergency (18.2%, 95% CI 11.5 to 26.7% [20/110] versus 20.0%, 95% CI 13.7 to 27.6% [28/140]) and mixed elective and emergency theatres (44.5%, 95% CI 35.1 to 54.3% [49/110] versus 30.7%, 95% CI 23.2 to 39.1% [43/140]).

Strategy 3

Randomisation was generally well-balanced at a hospital level in Benin (4 interventions versus 1 control), Ghana (4 interventions versus 6 controls), India (11 interventions versus 10 controls), Mexico (2 interventions versus 2 controls), Nigeria (8 interventions versus 8 controls), Rwanda (7 interventions versus 5 controls) and South

Africa (1 intervention versus 1 control) in line with the minimisation criteria. Equally, randomisation was well-balanced between referral (30 interventions versus 27 controls) and non-referral hospitals (7 interventions versus 6 controls).

Strategy 4

All sites completed investigator training during virtual site initiation, with a median of 10 (IQR: 5–19) investigators present per virtual training session. All sites completed the online training modules with a mean of 9 investigators per site.

Strategy 5

The warm-up week was successfully conducted by all sites (70/70), and feedback was used to refine trial processes.

Strategy 6

Training on the use and relevance of the ChEETAH sticker was delivered to all participating hospitals (70/70). The aggregate register was maintained at all sites (70/70), with an overall inclusion of 98.1% (10,686/10,894, 95% CI 97.8 to 98.3%) of the eligible patients which was balanced by randomised allocation (98.6%, 95% CI 98.3 to 98.9% control vs 97.5%, 95% CI 97.1 to 97.9% intervention).

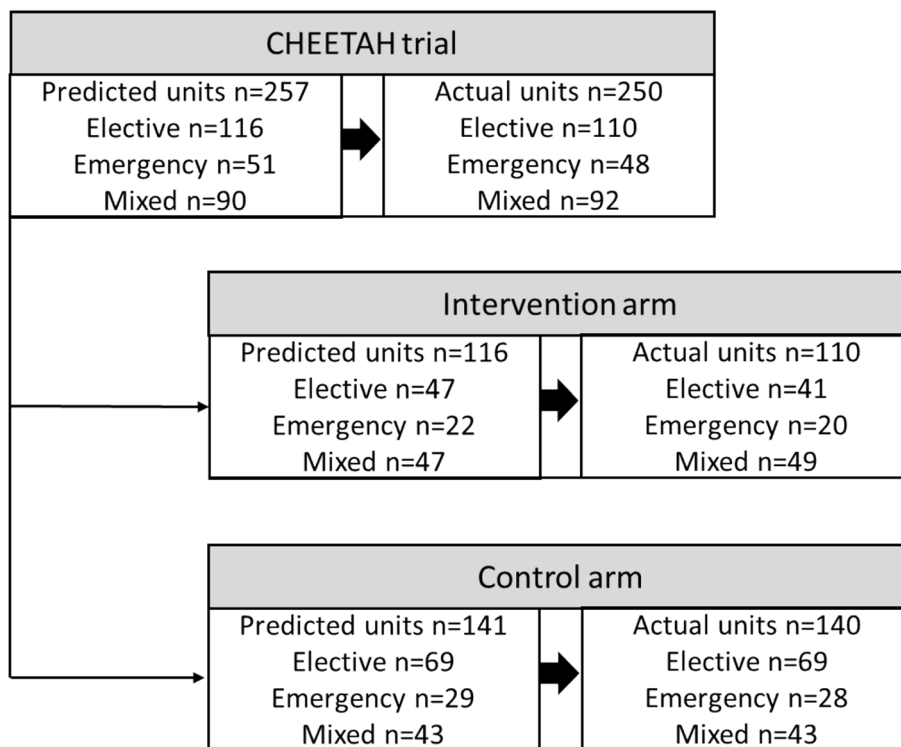


Fig. 2 Flowchart describing the predicted and actual theatres participating in the ChEETAH cRCT

A summary is presented in Additional file 1: Table S1. A median of 0 eligible patients (IQR: 0–1) was missed from inclusion. Routine monitoring of the aggregate register against the patient-level data uploaded on RED-Cap was conducted, allowing prompt identification of missed patients and further site training on recruitment processes.

Strategy 7

A summary of patient-level monitoring measures by cluster is shown in Additional file 1: Table 2, demonstrating the between-cluster variability in patient characteristics. A summary of monitoring measure by trial arm is presented in Table 1. Chance imbalances arose between the intervention and control groups in the proportions of patients who were < 18 years (8.9% versus 5.9%), ASA grade IV or V (7.2% versus 4.2%), emergency surgery (56.8% versus 46.6%, Fig. S1), benign disease (80.4% versus 59.4%), no WHO Surgical Safety Checklist (3.4% versus 0.2%) and open non-midline surgery (30.5% versus 19.1%). Chance imbalances reduced over time; an

Table 1 Baseline key characteristics of the patients included in the intervention and control arms

Factor	Routine practice, n = 5640	Change of gloves and instruments, n = 5046
Age		
< 18 years	484 (8.6)	522 (10.3)
≥ 18 years	5155 (91.4)	4524 (89.7)
ASA grade		
Grade I	2539 (45.0)	2600 (51.5)
Grade II	2086 (37.0)	1571 (31.1)
Grade III	757 (13.4)	735 (14.6)
Grade IV	158 (2.8)	121 (2.4)
Grade V	99 (1.8)	19 (0.4)
Timing of surgery		
Elective	2486 (44.1)	2603 (51.6)
Emergency	3153 (55.9)	2443 (48.4)
Indication		
Malignant disease	860 (15.2)	955 (18.9)
Benign disease	4506 (79.9)	3812 (75.5)
Trauma	274 (4.9)	279 (5.5)
WHO surgical safety checklist		
Yes	5387 (95.5)	4882 (96.8)
No	252 (4.5)	163 (3.2)
Operative approach		
Open—midline	3344 (59.3)	3325 (65.9)
Open—non-midline	2172 (38.5)	1605 (31.8)
Laparoscopic	38 (0.7)	56 (1.1)
Laparoscopic converted to open	86 (1.5)	60 (1.2)

example is provided for the urgency of surgery in Fig. S2. Figure S3 demonstrates the interplay between the urgency of surgery, the type of incision and surgical indication overall. Non-midline incisions were more frequent in elective surgery (27.2% [416/1635] vs 21.7% [352/1527] in emergency surgery) and were more often performed for benign surgical indications. These key patient characteristics were monitored centrally and reported at the TMG meetings, with swift action from the trials unit whenever biases were suspected. As an example, one site was enquired about inconsistencies in the aggregate register, uploaded data and theatre logbooks, admitting that some patients were missed. Lessons learned were disseminated across all hubs and spokes to avoid similar situations (see Additional file 1: Appendix F for more details on communication with the central management team).

Strategy 8

Overall, 0.4% patients (41/9187, 95% CI 0.3 to 0.6) refused consent for outcome assessment. This was balanced by trial arm (0.5% control vs 0.4% intervention) and across key risk subgroups (Additional file 1: Table S3).

Discussion

This study evaluated eight strategies to monitor and minimise bias and imbalances within an international cluster randomised controlled trial in surgery. Hospital-level strategies contributed to balance in unit of exposure level strategies, which in turn contributed to balance in patient-level strategies. Although all the strategies were delivered with success, chance imbalances in monitoring measures such as elective versus emergency operations, midline versus non-midline incisions and surgery for benign versus malignant indications were observed. Monitoring data were regularly reviewed by the TMG, with intervention at a site level where required. The diverse delivery network demonstrates the generalisability of future analyses of the primary outcome measure. Our data reflect that cRCTs are prone to chance imbalances at multiple stages and a transparent reporting of mitigation strategies is crucial for an adequate interpretation of the trial results. Pre-planning for chance imbalance in cRCT statistical analysis plans should be considered paramount.

cRCTs have increasingly been recognised in surgery and interventional specialties as a methodology for evaluating complex and/or behavioural change interventions in operating theatres [1–3, 14]. Operating theatres are multi-professional, multi-specialty environments where culture change is often required to empower practice change and uptake of evidence-based practice is slow [15]. Cluster methods, learning from concepts in

implementation and behavioural science, have significant potential in both evaluating late-phase interventions and encouraging sustainable adoption where benefit is observed. This may be particularly relevant to resource-constrained settings, where contextually sensitive interventions require deep co-development, and the ability to scale across networks is key [16]. However, cluster trials in surgery are challenging. First, clusters are often defined at a hospital level but significant within-cluster heterogeneity exists both in the unit of exposure (operating theatre, surgical team or theatre list) and patient-level case mix (urgency and types of operation). Although patients undergoing elective and emergency surgery are expected to have distinct characteristics and outcomes [17–19], the arm imbalances observed in the ChEETAH trial are expected as with any cluster trials. With appropriate adjustments for key confounders in statistical analysis, this will limit the compromise of the internal validity of the results. Second, patients cannot be recruited upfront before randomisation as in the case chronic disease management, so recruitment is performed after randomisation [20]. In the ChEETAH trial, 97% of the hospitals maintained their case mix of emergency and elective theatres, demonstrating very little recruitment bias after randomisation. The main model of this trial will adjust for minimisation factors (country and type of hospital) as well as key confounders which include both hospital-level and patient-level covariates.

Previous authors have described similar methods for cRCTs but have described application to outpatient settings or chronic disease and have limited applicability to the surgical settings (i.e. one-off intervention exposure, multi-level within and between cluster variability) [20]. Chance imbalances are expected in cRCTs and adjustment for the key cluster- and patient-level variables whilst accounting for clustering is essential during data analysis [8]. These statistical techniques however cannot account for unmeasured confounding so should be seen as part of a whole-trial approach to bias mitigation, rather than standing alone [11]. We observed here that chance imbalances between key risk variables decreased over time (i.e. as sample size increased); this has been previously observed in meta-analyses of cRCTs [21] which suggests that patient-level imbalance is expected in cRCTs.

As this trial did not require patient-level consent for exposure to the intervention, but did for outcome assessment, the risk of bias from the refusal of consent was low. We worked with patient and community partners to design a pragmatic follow-up schedule [22] that was low burden (completed at a single time point) and collected outcome data from electronic health records where feasible (e.g. reoperation) [23]. Equally, we had no

centres that recruited no patients so would be excluded from an intention-to-treat analysis [20], nor was there a risk of cluster ‘migration’ (where participants move out or into one cluster to/from another) as with some trials in chronic disease. Patients were the only blinded party in this trial, as the operating teams were delivering the intervention and outcome assessors were likely to be aware of the hospital allocation. Cluster randomised trials in surgery are unique, and designing them requires multifaceted considerations of different issues that may warrant such strategies to minimise risk of bias as compared to other fields. For instance, there are weekend and night operations, and decisions on patient inclusion into trials are difficult. Further blinding in surgery is not entirely feasible. This trial has been co-developed with stakeholders from LMICs, through face-to-face meetings and a formal Delphi process, to ensure its relevance. Both the interventions and design of research were judged to be applicable to settings in lower-resource countries, as judged by frontline surgeons. Finally, we had many clusters ($N=70$) and a randomisation-minimisation algorithm, so allocation concealment was easily maintained (i.e. unlikely that new clusters would be able to anticipate their randomised allocation). Future trials where there would be a risk of these alternative sources of bias should consider including methods to monitor and minimise these during planning and implementation [24].

This study had limitations. First, we were unable to fully account for the risk of recruitment bias using the methods described [25]. However, by using consecutive sampling of eligible patients and monitoring for refusal of consent are actively aware of this, and with interpretation of the full trial results with this caveat. Second, imbalances persisted in certain patient-level variables between randomisation arms. These decreased as sample size increased (suggesting chance imbalance only). Future meta-analyses of cRCTs in this area are warranted to fully assess the effectiveness of these strategies in mitigating against selection bias in cRCTs. Third, we specifically highlight chance imbalance and attrition bias as key causes of differential misclassification in cRCTs but other sources of bias exist (performance bias, measurement bias, dilution bias) that we have not considered here. Fourth, whilst we have focussed on surgical cRCTs, the findings may be applicable to other interventional and procedural specialties where differences in outcomes can occur related to the procedural performer, environment and technique in addition to patient-related factors. We encourage other investigators to explore these concepts across other specialty areas.

Future guidance on the delivery and reporting of cRCTs in surgery and other areas of knowledge should incorporate strategies similar to these to mitigate against

bias and imbalance by arm. The CHEETAh trial provides a structured approach which was implemented successfully across diverse settings and can be flexibly applied to future cRCTs.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13063-022-06852-2>.

Additional file 1: Fig. S1. Balance of elective and emergency surgery by trial arm. **Fig. S2.** Changes in imbalance by arm of urgency of surgery over time. **Fig. S3.** Flowchart of key surgery characteristics overall. **Fig. S4.** Number of actual and predicted units of exposure per hospital. **Table S1.** Number of eligible patients included (patient-level analysis). **Table S2.** Baseline characteristics of included patients (N=, %) (hospital-level analysis). **Table S3.** Refusal of consent for outcome assessment (patient-level analysis). **Table S5.** Mapping strategy domains to Cochrane risk of bias tool. **Appendix A.** Author list. **Appendix B.** Strategies following identification of potential sources of bias. **Appendix C.** CHEETAh trial-specific register. **Appendix D.** CHEETAh trial operation sticker. **Appendix E.** CHEETAh aggregate register. **Appendix F.** CHEETAh Patient Pathway Flowchart.

Data sharing

Anonymised data available upon the request to the CHEETAh Study Management Group, upon successful completion of a Data Sharing Agreement.

Authors' contributions

The author read and approved the final manuscript.

Funding

CHEETAh was funded through an advanced clinician scientist award from the NIHR Academy and supported by a hub and spoke network funded through a National Institute for Health Research (NIHR) Global Health Research Unit Grant (NIHR 16.136.79). The funder and sponsor had no role in the study design or writing of this report. The funder has approved the submission of this report for publication. The views expressed are those of the authors and not necessarily those of the National Health Service, the NIHR or the UK Department of Health and Social Care.

Declarations

Ethics approval and consent to participate

UK ethical approval was obtained from the University of Birmingham International Research Ethics Committee. All individual participating countries obtained local or national ethical approval for CHEETAh in accordance with local requirements (available upon request).

Competing interests

The author declares that there are no competing interests.

Author details

¹NIHR Global Health Research Unit on Global Surgery, Institute of Applied Health Research, University of Birmingham, B15 2TH Birmingham, UK.

Received: 27 May 2022 Accepted: 19 October 2022

Published online: 05 April 2023

References

1. Haugen AS, et al. Effect of the World Health Organization checklist on patient outcomes: a stepped wedge cluster randomized controlled trial. *Ann Surg.* 2015;261(5):821–8.

2. Collaborative EESA. ESCP Safe Anastomosis ProGramme in CoLorectal SurgEry (EAGLE): study protocol for an international cluster randomised trial of a quality improvement intervention to reduce anastomotic leak following right colectomy. *Colorectal Dis.* 2021;23(10):2761–71.
3. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implement Sci.* 2011;6:42.
4. Campbell MK, et al. Consort 2010 statement: extension to cluster randomised trials. *BMJ.* 2012;345:e5661.
5. Gonzales R, et al. A cluster randomized trial of decision support strategies for reducing antibiotic use in acute bronchitis. *JAMA Intern Med.* 2013;173(4):267–73.
6. Investigators A. Enhanced postoperative surveillance versus standard of care to reduce mortality among adult surgical patients in Africa (ASOS-2): a cluster-randomised controlled trial. *Lancet Glob Health.* 2021;9(10):e1391–401.
7. Conroy EJ, et al. Randomized trials involving surgery did not routinely report considerations of learning and clustering effects. *J Clin Epidemiol.* 2019;107:27–35.
8. Hemming K, et al. How to design efficient cluster randomised trials. *BMJ.* 2017;358:j3064.
9. Crespi CM. Improved designs for cluster randomized trials. *Annu Rev Public Health.* 2016;37:1–16.
10. Bouwsma EVA, et al. Effectiveness of an internet-based perioperative care programme to enhance postoperative recovery in gynaecological patients: cluster controlled trial with randomised stepped-wedge implementation. *BMJ Open.* 2018;8(1):e017781.
11. Eldridge S, et al. Internal and external validity of cluster randomised trials: systematic review of recent trials. *BMJ.* 2008;336(7649):876–80.
12. Surgery, N.G.H.R.U.o.G. Study protocol for a cluster randomised trial of sterile glove and instrument change at the time of wound closure to reduce surgical site infection in low- and middle-income countries (CHEETAH). *Trials.* 2022;23(1):204.
13. Berrios-Torres SJ, et al. Centers for Disease Control and Prevention Guideline for the Prevention of Surgical Site Infection, 2017. *JAMA Surg.* 2017;152(8):784–91.
14. Skivington K, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. *BMJ.* 2021;374:n2061.
15. Ergina PL, et al. Challenges in evaluating surgical innovation. *Lancet.* 2009;374(9695):1097–104.
16. Dron L, et al. The role and challenges of cluster randomised trials for global health. *Lancet Glob Health.* 2021;9(5):e701–10.
17. GlobalSurg C. Surgical site infection after gastrointestinal surgery in high-income, middle-income, and low-income countries: a prospective, international, multicentre cohort study. *Lancet Infect Dis.* 2018;18(5):516–25.
18. GlobalSurg C. Mortality of emergency abdominal surgery in high-, middle- and low-income countries. *Br J Surg.* 2016;103(8):971–88.
19. Biccard BM, et al. Perioperative patient outcomes in the African Surgical Outcomes Study: a 7-day prospective observational cohort study. *Lancet.* 2018;391(10130):1589–98.
20. Giraudeau B, Ravaud P. Preventing bias in cluster randomised trials. *PLoS Med.* 2009;6(5):e1000065.
21. Bolzern J, Mnyama N, Bosanquet K, Torgerson D. Comparing evidence of selection bias between cluster-randomised and individually randomised controlled trials: a systematic review and meta-analysis. *Lancet.* 2018;392:S21.
22. Tallon D, Chard J, Dieppe P. Relation between agendas of the research community and the research consumer. *Lancet.* 2000;355(9220):2037–40.
23. Nepogodiev D, et al. Differences in outcome between patients readmitted to index vs non-index hospital trusts after colorectal resection. *Colorectal Dis.* 2019;21(8):943–52.
24. Hahn S, et al. Methodological bias in cluster randomised trials. *BMC Med Res Methodol.* 2005;5:10.
25. Torgerson DJ. Diabetes education: selection bias in cluster trial. *BMJ.* 2008;336(7644):573.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.