BMJ Open

Evaluation of the effect of multidisciplinary simulation-based team training on patients, staff and organisations: protocol for a stepped-wedge cluster-mixed methods study of a national, insurer-funded initiative for surgical teams in New Zealand public hospitals

Jennifer Weller ,[1] Jennifer Anne Long,[1] Peter Beaver,[1] David Cumin,[2] Chris Frampton,[3] Alexander L Garden,[4,5] Matthew Moore,[2] Craig S Webster ,[1,2] Alan Merry[2]

To cite: Weller J, Long JA, Beaver P, *et al*. Evaluation of the effect of multidisciplinary simulation-based team training on patients, staff and organisations: protocol for a stepped-wedge cluster-mixed methods study of a national, insurer-funded initiative for surgical teams in New Zealand public hospitals. *BMJ Open* 2020;**10**:e032997. doi:10.1136/bmjopen-2019-032997

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2019-032997).

Received 16 July 2019
Revised 15 January 2020
Accepted 27 January 2020

Check for updates

For numbered affiliations see end of article.

Correspondence to
Professor Jennifer Weller; j.weller@auckland.ac.nz

ABSTRACT

Introduction NetworkZ is a national, insurer-funded multidisciplinary simulation-based team-training programme for all New Zealand surgical teams. NetworkZ is delivered in situ, using full-body commercial simulators integrated with bespoke surgical models. Rolled out nationally over 4 years, the programme builds local capacity through instructor training and provision of simulation resources. We aim to improve surgical patient outcomes by improving teamwork through regular simulation-based multidisciplinary training in all New Zealand hospitals.

Methods and analysis Our primary hypothesis is that surgical patient outcomes will improve following NetworkZ. Our secondary hypotheses are that teamwork processes will improve, and treatment injury claims will decline. In addition, we will explore factors that influence implementation and sustainability of NetworkZ and identify organisational changes following its introduction. The study uses a stepped-wedge cluster design. The intervention will roll out at yearly intervals to four cohorts of five District Health Boards. Allocation to cohort was purposive for year 1, and subsequently randomised. The primary outcome measure is Days Alive and Out of Hospital at 90 days using patient data from an existing national administrative database. Secondary outcomes measures will include analysis of postoperative complications and treatment injury claims, surveys of teamwork and safety culture, in-theatre observations and stakeholder interviews.

Ethics and dissemination We believe this is the first surgical team training intervention to be implemented on a national scale, and a unique opportunity to evaluate a nation-wide team-training intervention for healthcare teams. By using a pre-existing large administrative data set, we have the potential to demonstrate a difference to surgical patient outcomes. This will be of interest to those working in the field of healthcare teamwork, quality improvement and patient safety. New Zealand Health and Disability Ethic Committee approval (#16/NTB/143).

Trial registration number Australian and New Zealand Clinical Trials Registry ID ACTRN12617000017325 and the Universal Trial Number is U1111-1189-3992.

Strengths and limitations of this study

- ► The study uses a mixed methods approach to test the effectiveness of the NetworkZ training intervention in the context of real-world implementation across an entire nation.
- ► Our primary outcomes measure, Days Alive and Out of Hospital at 90 Days, is an holistic measure of surgical outcomes using routinely collected health data from a national administrative database.
- ► Programme rollout follows a stepped wedge cluster design in which cohorts of sites move progressively from baseline to post-implementation states in random order.
- ► Running the training programmes is subject to willingness of hospitals to participate, and their own constraints around dates and numbers of courses. This could compromise the purity of the stepped wedge design.
- ► Possible confounding by another factor operating progressively over time is a limitation of our design.

BACKGROUND

Unintended harm to patients due to medical treatment is a major contributor to the global burden of disease.[1] In a systematic review of studies assessing inpatient outcomes[2] surgical procedures were associated with 39.6% of all adverse events and were commonly associated with unintended injury during hospitalisation. Communication failures contribute to many medical mishaps[3] and sentinel events. In one study communication failures affected about 30% of team interactions in the operating theatre,[4] and in another they contributed to 43% of surgical errors.[5] Even when operating theatre staff work hard to maintain good relationships and minimise tensions,[6] disciplinary silos,[7] professional rivalries[8] and hierarchy[3] create barriers to effective communication.

The Agency for Healthcare Research and Quality recommends team training as a patient safety strategy.[9] Team training has been shown to improve communication.[10 11] There is some evidence that it can improve team processes, clinical processes and potentially, patient outcomes.[12 13] Team training has been shown to improve safety culture and encourage attitudes and behaviours commonly found in high reliability organisations.[14] Simulation is frequently used in team training initiatives.[15 16] This study protocol focuses on the evaluation of a national simulation-based team training initiative.

Healthcare for New Zealand's population of 4.8 million people is provided through 20 District Health Boards (DHBs) of varying size, with the smallest servicing a population of around 32 000 and the largest servicing a population to over 600 000. Each DHB includes one or more publicly funded hospital. New Zealand has a national no-fault insurer, the Accident Compensation Corporation (ACC), which compensates accident victims who sustain injuries, including injuries following treatment (known as 'treatment injuries'). The number and costs of treatment injuries has increased over the last 5 years. In 2015/2016, ACC accepted 8881 medical and surgical treatment injury claims, with incurred costs of NZ$418 million. In an effort to reduce the human and fiscal toll of these events ACC is investing NZ$45 million over a 5-year period into initiatives to reduce treatment injuries, targeting surgery, surgical site infections, pressure injuries and maternity care.[17] Ten million dollars of that investment is funding a multidisciplinary, simulation-based team training programme, NetworkZ.

The NetworkZ programme is run in situ in operating theatres with full surgical teams who normally work together, to be implemented in all public hospitals in New Zealand. The aim is to establish regular team simulations in each hospital operating department, supplemented by additional 30–60 min workshops on speaking up, actively listening, structured recap and handover, and closed loop communication.

It is supported by the NetworkZ Instructor course which trains DHB staff to deliver the programme in their DHB. A full description of the programme can be found at www.networkz.ac.nz and in online supplementary material 1.

In a pilot study, 20 surgical teams from two large Auckland hospitals participated in NetworkZ at the University of Auckland Simulation Centre for Patient Safety with positive participant evaluations and improved scores for observed teamwork.[18–20] The national roll-out now offers an opportunity to test the impact of team training on surgical outcomes at a national level.

We propose that, for a team-training programme to have a widespread and lasting impact on teamwork and communication, it must meet a number of conditions. It should involve teams that normally work together, and take place in the workplace, and in their own environment. It should be developed and delivered by a multidisciplinary team. It should have a sound theoretical framework for teamwork, educational approach and implementation. It should be locally 'owned' and led (as opposed to being externally imposed) following the initial implementation phase and the training should be provided in a 'sufficient dose' to involve the majority of staff in the training.

The educational framework for NetworkZ draws on the work of Kolb[21] and Boud *et al*[22] on experiential learning and reflective practice, and embraces interprofessional learning, gaining insights into the beliefs and roles of others, challenging assumptions and strengthening social bonds between group members. It draws on an evidence-based model of the features of effective teamwork conceptualised by Salas and colleagues.[23] The key features of this model are leadership, adaptability, mutual performance monitoring, back-up behaviour and team orientation. The underpinning elements are a shared mental model, clear concise communication and mutual trust and respect.

Our implementation strategy builds on the work of Bate,[24] and Dixon-Woods *et al*.[25] Factors believed to foster successful implementation of initiatives such as ours include organisational support and a shared commitment to the programme, leadership engagement, staff empowerment and networking, engaging clinical champions, together with building intrinsic staff motivation.[24 26] Hence implementation of the NetworkZ programme will provide extensive initial support to local project and instructor teams, senior management engagement and resourcing and follow a train-the-trainers model for local embedding and long-term sustainability. Instructors attend a 2-day instructor workshop, complete online learning package are provided with on-site feedback and assessment as they take over responsibilities for running the course. Initial DHB courses are led by NetworkZ faculty, mentoring local instructors to the point where they can run the courses independently. Ongoing monitoring of the quality of the simulation training is provided through centralised review of participant evaluations and end of course reports, instructor accreditation visits by NetworkZ faculty and regular advanced instructor courses. The implementation strategy is described in more detail in online supplementary material 1.

The aim of NetworkZ is to improve outcomes for surgical patients by improving teamwork and communication in operating theatre teams through embedding regular simulation-based multidisciplinary team training in New Zealand hospitals.

METHODS

In this study we will address the following hypotheses and research questions.

Primary hypothesis

Our primary hypothesis is that surgical patient outcomes will be improved following the implementation of NetworkZ as manifested in Days Alive and Out of Hospital over 90 days ($DAOH_{90}$).

Secondary hypotheses

Our two secondary hypotheses are: (1) teamwork and communication processes will be improved and (2) the number and cost of ACC treatment injury claims will decline following the introduction of our training programme.

Qualitative research questions

We aim to: explore the factors that facilitate or impede implementation of NetworkZ; understand requirements for long-term sustainability; and identify the types of changes occurring in the workplace following the implementation of NetworkZ.

Trial design

This is a mixed methods study, using quantitative and qualitative data in a stepped wedge, cluster randomised quality improvement design (see figure 1). There are four cohorts of five DHBs each. Randomisation process: The 20 DHBs were divided into five groups of similar population size, calculated from publicly available Ministry of Health data on population sized served by each DHB. The first cohort comprised one DHB purposively selected from each of these five groups on the basis of capacity to start implementation. Following this initial selection, a random DHB from each size group was then selected for cohort 2 and then for cohort 3. The remaining DHBs comprised cohort 4. In other words, one DHB from each size group was randomised to start the programme at the beginning of intervention period 2, period 3 and period 4 (cohorts 2, 3 and 4).

Comparison of outcomes will be undertaken between the pre-intervention period (control) starting 1 year before the first NetworkZ course in cohort 1 DHBs, and the post-intervention period ending 1 year after the end of the transition period in cohort 4. The transition period runs from the start of the DHB training for a cohort, and lasts for 15 months, during which at least one third of staff in that cohort should have attended a course. We believe that the effects from the intervention will not clearly manifest until at least 15 months, by which time a significant proportion of the staff will be trained and experienced with the intervention. No quantitative data will be collected during the transition period. We have chosen the stepped wedge cluster design because it lends itself to a staged quality improvement project in which individual participants cannot be randomised to intervention or control[27] (see figure 1).

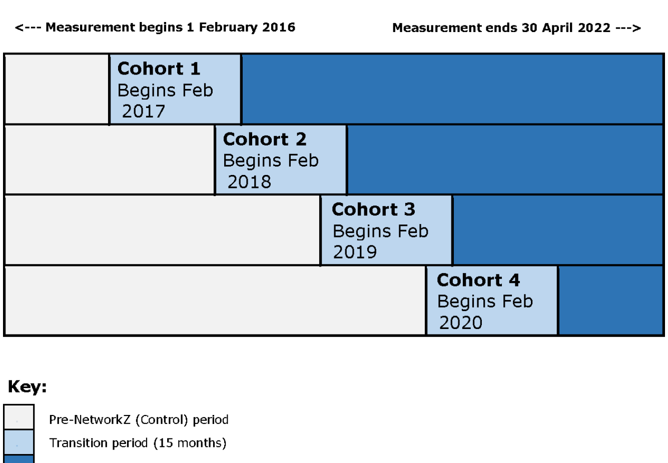

Figure 1 The stepped-wedge cluster cohort design.

Study population

There are two distinct populations included in the study protocol: the population of surgical patients and the population of surgical staff.

Patient population

We will include patients undergoing surgery in public hospitals in New Zealand between 1 February 2016 and 30 April 2022 for patient outcome measures. We will obtain outcome data for our primary hypothesis from the National Minimum Dataset (NMDS), a reliable administrative dataset collected by the New Zealand Ministry of Health. This data set provides a confidential and comprehensive record of all public hospital admissions and discharges in New Zealand, including demographics, International Classification of Diseases (ICD) codes for diagnoses, procedures, length of stay and mortality associated with each admission. In 2017, the NMDS included 330 353 cases of publicly funded surgery.[28] The funding for NetworkZ allowed us to provide the team training for five major surgical specialities: general, orthopaedics, urology, otorhinolaryngology and plastics.

For inclusion and exclusion criteria, see table 1.

Patient sample size and power analysis

The sample will be all patients undergoing surgery in public hospitals during the study period who stay at least one night in hospital and are undergoing procedures in the included specialty groups. This number has been estimated from NMDS historical data to be in the order of 550 000 cases over the 6-year study period.

Our sample size is primarily pragmatic. Both our design (ie, stepped-wedge) and primary outcome variable (ie, DAOH) present significant obstacles to conventional sample size calculations. Nonetheless, we have undertaken an empirical power analysis on simulated data. We generated smooth sampling distributions from historical

Table 1 Inclusion and exclusion criteria for patients and hospital staff

Study inclusion criteria	Study exclusion criteria
Surgical patient admissions	
Patient admissions that meet *all* of the following criteria: ► Receive surgery in the five targeted surgical specialities. ► Receive surgical intervention involving a full surgical team. ► Requiring an overnight stay. ► Receive regional or general anaesthesia. ► Admitted for surgery to a publicly funded hospital that has been, or will be, involved in the NetworkZ programme. All subsequent readmissions to any hospital facility within the following 90 days will also be included.	Patient admissions that meet *any* of the following criteria: ► Admitted and discharged on the day of surgery. ► Admitted to private hospitals. ► Receiving endoscopy using operator supervised sedation. ► Procedures requiring local infiltration only. ► Patients admitted to stand-alone paediatric hospitals.
Hospital staff	
Public hospital staff working in one of the following roles: ► Surgical specialists in the included five specialty groups. ► Anaesthetic specialists. ► Nursing staff working in operating theatres. ► Anaesthetic assistants (referred to as Anaesthetic Technicians in New Zealand).	Hospital staff working as: ► Trainees in surgery or anaesthesia. ► Solely specialised paediatric hospital employees. ► House surgeons (junior doctors not in specialist training programmes). ► Healthcare assistants. ► Solely private hospital employees.

data for each DHB. This historic data feature the cohort structure of DHBs, and the within-DHB clustering of patients likely to be present in our future study data. We also generated versions of each distribution with a difference imposed at given quantiles by either flattening or accentuating the peaks in the data (ie, a larger late peak means more patients with better DAOH). Using the recruitment parameters of our study, we generated 1000 synthetic datasets with a simulated intervention effect. We compared the control and intervention groups using the Wilcoxon-Mann-Whitney U test, and assessed the significance of the calculated statistic by randomly permuting sites between cohorts 1000 times. Our simulations indicate a greater than 95% power at α=0.05 to detect an intervention effect of one or more DAOH at the median, and two or more DAOH at the 0.25 quantile. We may apply quantile regression risk-adjustment to statistically evaluate differences of less than one DAOH (ie, the minimum resolution of unadjusted DAOH data).

Operating theatre staff population

NetworkZ aims to train the majority of eligible staff working in the operating theatres of New Zealand public hospitals. Because there are four cohorts with a staged annual roll out from 2017 to 2020 inclusive, the expected proportion of eligible operating theatre staff attending training over the study period are: 100% of cohort 1, 75% of cohort 2, 50% of cohort 3, 25% of cohort 4 and this is the basis of our sample size estimate. For surveys the unit of analysis will be the individual, while for observations, the unit of analysis will be the theatre team. See table 1 for inclusion and exclusion criteria.

Total surgical staff population size

From specialist registration boards and hospital staffing lists we estimate the following: specialist surgeons ~730; specialist anaesthetists ~750; operating theatre nurses ~1800; anaesthetic technicians ~480. (Total =~3800 staff)

Operating theatre staff sample size estimate

For surveys, we will include all staff meeting the inclusion criteria. We will rate between 5 and 10 cases in each DHB in the 3-month period prior to the introduction of NetworkZ in that DHB and immediately following the transition period in that DHB. Over the course of the study, this will generate between 100 and 200 in the pre-NetworkZ period and a similar number in the post-NetworkZ period. Previous work indicates that fewer than 20 cases before and 20 after an intervention can detect a difference of one point on the seven-point WHOBARS scale[29] in 80% of cases. We are undertaking more observations to obtain a representative sample across DHBs.

Outcome measures

Primary outcome measure (addressing primary hypothesis)

The primary outcome measure for the study is Days Alive and out of Hospital after 90 days ($DAOH_{90}$), recently validated for the surgical context.[30]

DAOH provides a holistic approach to measuring patient harm that incorporates mortality and length of time out of hospital for a defined number of days (in this case 90 days) following an index event (in this case, a surgical procedure).[31] Days in hospital include those until discharge following surgery, and those during any readmissions to public hospitals within the specified 90 day postoperative period. $DAOH_{90}$ scores will be affected by mortality and any complications that either prolong hospitalisation or result in readmission. $DAOH_{90}$ score will be zero for patients who die without ever leaving hospital, but any day alive and out of hospital during the 90-day period will be counted.

Weller J, *et al*. *BMJ Open* 2020;**10**:e032997. doi:10.1136/bmjopen-2019-032997

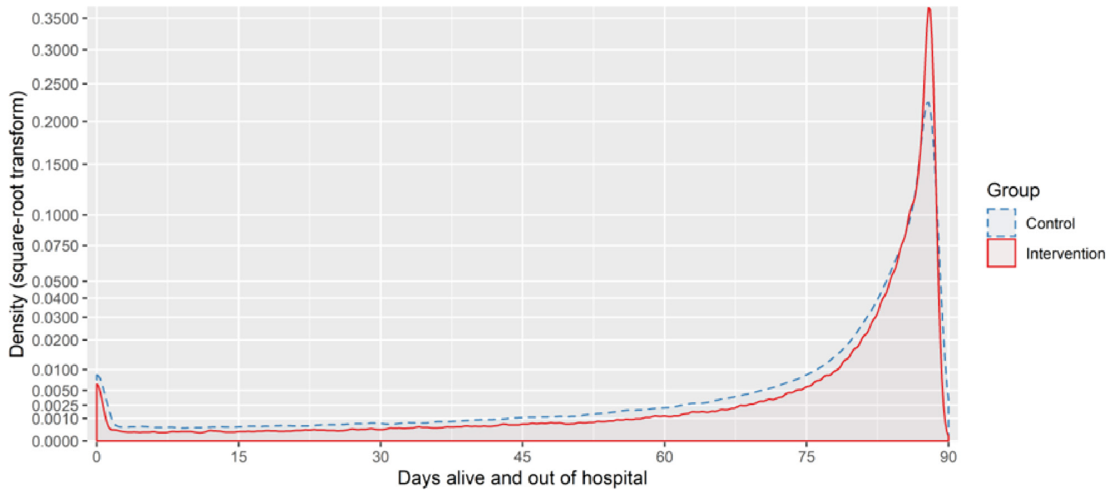

Figure 2 Probability density graph for the distribution of DAOH based on historical data from 1 July 2011 to 30 June 2016 (blue dotted) and the same with a simulated difference of 1 day imposed at the median (red solid). note square root transform on y-axis.

Index procedures for $DAOH_{90}$ will be limited to patients who stay in hospital for at least one night. The rationale for this is to restrict the patient population to those at higher risk of adverse events, and to increase the likelihood that when a patient is admitted for a diagnostic procedure prior to definitive surgery, the 90-day period begins with this definitive surgery. A surgical admission within 90 days will be subtracted from the 90 days when calculating $DAOH_{90}$. Additionally, we will explore the influence of the proportion of staff in each DHB who have attended a NetworkZ course. In a secondary analysis, we will also calculate Days out of Hospital and Mortality separately to tease out the individual effects of these on the primary outcome. The analysis will follow the procedure for the analysis of the primary outcome measure.

Statistical analysis

Using the Stepped Wedge cluster design we will analyse the differences between pre-NetworkZ and post-NetworkZ cohorts on each of the patient outcome measures. Data will be classified into pre-NetworkZ (control), transition and post-intervention intervention categories (see figure 1).

Due to the highly skewed and bimodal nature of the $DAOH_{90}$ scores (see figure 2), we will use the non-parametric Wilcoxon-Mann-Whitney rank sum test as our primary test to derive a statistic for the overall difference between distributions before and after implementation of NetworkZ. To assess the significance of the difference, we will compare the statistic to the values derived from 10 000 random permutations of cluster (ie, DHB) to sequences (ie, cohorts). Our permutation test will differ from that demonstrated by Thompson, *et al*[32] in that we will not use within-period comparisons. The distribution of the test statistic generated by this procedure may not be centred at zero, particularly if there is confounding of the control vs intervention comparison with an overall temporal effect. For this reason, the observed test statistic will be compared with the centre of the generated distribution. We have implemented this in R.[33] The permutation test does not vary the timing of an individual patient's admission, rather an individual patient will be permuted with all other patients in their cohort, as they will be assigned to the intervention based on the interplay between their actual time and the cohort sequence to which they are permuted. Thus, the null distribution of the Mann Whitney U test will account for any overall temporal effects within the dataset.

We will then perform quantile regression at the 0.1, 0,25, 0.75, 0.9 quantile to characterise any difference between the two distributions. This quantile regression will include time (month and year of procedure) in order to adjust for time effects, cohort, DHB (within cohort) and intervention as factors in the model. To assess the impact of the purposive sampling of cohort 1 DHBs, we will conduct a sensitivity analysis by re-running the above analysis using cohorts 2, 3 and 4 data.

Secondary outcome measures (addressing secondary hypotheses)

Teamwork perceptions survey

The Teamwork Perceptions Survey was designed to align with the intended learning outcomes of NetworkZ (online supplementary material 2). Items were drawn from behaviours of effective teams described by Salas *et al*[23] and refined through review and consensus with an expert group. The survey will be administered to all operating theatre staff who are regularly working in operating theatres in the specialities targeted by the course.

This survey will be administered in the 3-month period prior to a DHB commencing training and repeated at the end of the transition period (at 15 months). Staff in DHBs around the country will be assigned unique confidential identifiers. These identifiers will be used in the surveys to enable matching of participants' preintervention and postintervention survey responses and determine if they have participated in NetworkZ.

Analysis

Pre-Teamwork and post-Teamwork Perceptions Survey total scores and subscale scores will be assessed using a Generalized Linear Mixed Model (GLMM). The analysis will include terms for cohort, *DHB (random effect nested within cohort), intervention, month and year,* with the primary comparison using the cohort as the replicate of the intervention effects. Small sample correlations will be considered. Additionally, we will explore the influence of the proportion of staff in each DHB who have been trained.

Surgical Safety Culture Survey

The Surgical Safety Culture Survey measures attitudes and behaviours relevant to surgical safety[34] and survey scores have been shown to predict 30-day risk of postoperative death.[35] The Health Quality and Safety Commission (HQSC) has made minor changes to the survey terminology for the New Zealand context. HQSC administers this national survey to all operating theatre staff on a bi-annual basis. The survey is anonymous but includes demographics such as DHB and professional role. HQSC administered the survey in February 2017 with subsequent administrations planned for February 2019 and February 2021.

Analysis

Pre-SCSS and post-SCSS total scores and subscale scores will be assessed using a GLMM. The analysis will include terms for cohort, *DHB (random effect nested within cohort), intervention, month and year,* with the primary comparison using the cohort as the replicate of the intervention effects. Small sample correlations will be considered.

Observations of WHO Surgical Safety Checklist administration

We will use the World Health Organisation Behaviourally Anchored Rating Scale (WHOBARS) observation tool previously developed and validated by members of our research group[29 36] (see online supplementary material 3). WHOBARS measures the quality of administration of the Surgical Safety Checklist (SSC). The SSC is used almost universally in New Zealand and provides a standardised opportunity to share important information with other members of the surgical team. Changes in teamwork and communication should be reflected in this measure. Case selection for the WHOBARS observations will be a convenience sample of available cases over a 3-day rating period in each DHB and will include any surgical case that falls within the eligibility criteria for NetworkZ. We will record unique staff identifiers for each team observed to enable post-NetworkZ influence of the number and role of staff in a team who have participated in a NetworkZ course. In order to generate reliable scores, we will train raters using a series of videos and actual theatre cases.

Analysis

The analysis will compare mean and subscale WHOBARS scores for surgical teams pre-NetworkZ and post-NetworkZ intervention using a GLMM. The analysis will include terms for cohort, *DHB (random effect nested within cohort), intervention, month and year,* with the primary comparison using the cohort as the replicate of the intervention effects. Small sample correlations will be considered. We will also undertake an exploratory analysis to investigate if the proportion of staff trained moderates the difference between pretraining and post-training.

Number and costs of ACC treatment injury claims

To provide an indication of the fiscal outcomes of the intervention, surgical treatment injury claims submitted to ACC will be analysed using claims that are relevant to the surgical specialties targeted by NetworkZ. We will exclude claims that relate to specialties not included in NetworkZ. We will analyse ACC data using the same patient population groups over the same time-periods as established for the stepped wedge design for the primary analysis. The analysis will compare claim data pre-NetworkZ and post-NetworkZ implementation using a GLMM. The analysis will include terms for cohort, *DHB (random effect nested within cohort), intervention, month and year,* with the primary comparison using the cohort as the replicate of the NetworkZ effects. Small sample correlations will be considered. We will conduct separate analyses for the number and cost of claims.

Postoperative complications captured in the NMDS

To provide secondary data around patient safety, we intend to examine postoperative complications available in the NMDS. The NMDS captures information about new diagnoses that emerged as a result of surgical complications and instances of medical misadventures during surgery that have been captured in medical notes. This data will be analysed using the same patient population groups over the same time periods as the primary analysis.

Sick leave records

Improved teamwork has the potential to reduce staff retention and the use of sick-leave. We will assess monthly OR staff time off from illness by DHB to assess if there are reductions in sick leave following introduction to the programme.

Qualitative measures (addressing qualitative research questions)

We will undertake interviews with each cohort to address the three qualitative research questions. We will conduct semi-structured interviews with DHB staff involved in NetworkZ on the project team or as instructors. We will identify potential participants through the NetworkZ database and sampling will be purposive to include a range of views from large and small DHBs, from different professional groups, and those with different roles. Sampling will continue to the point of data sufficiency, when no new themes are emerging and existing themes are sufficiently described. We anticipate a sample size of 20–30 interviews per cohort. Timing of interviews will be during the transition period, during the period of implementation of NetworkZ. Interviews will be transcribed and coded using NVivo prior to framework analysis or general thematic analysis.

Patient and public involvement

There was no patient or public involvement in the development of this study protocol.

ETHICS AND DISSEMINATION

The protocol was registered with the Australian and New Zealand Clinical Trials Registry on 5 January 2017 and last updated on 7 August 2018. The protocol title is 'Evaluation of NetworkZ: Can multidisciplinary team-training improve the safety of surgery in New Zealand?'

DISCUSSION

NetworkZ is a nationwide response to the global need to improve the safety of surgical care. To our knowledge, this will be the first occasion when a surgical team training initiative has been implemented on a national scale. Access to routinely collected health data at a national scale is key to the feasibility of our evaluation.

The unique features of NetworkZ team training include: interactive surgical models integrated with computerised full-body computerised manikins that allow all members of the surgical team to participate; in situ training using real teams in real operating theatres; providing simulation resources and instructor training to build local capacity to independently run the course.

Strengths of the study

This is a national study, involving multiple sites and large numbers. It is adequately powered to show a difference in DAOH if one exists. Access to high-quality data routinely collected nationally on hospital events is a major strength, without which the study would not be feasible.

Our primary outcome measure, DAOH, is a relatively new measure for surgical patient outcomes, recently validated for the surgical context.[30] $DAOH_{90}$, provides a holistic evaluation of surgical outcomes that is sensitive to any cause of death, prolonged hospital stay or readmission. This measure can be reliably and cost-effectively derived from the data in the NMDS and is objective.

Stepped wedge cluster design is very suitable for quality improvement initiatives in which it is impossible (or at least highly impracticable) to randomise individual cases or participants.[27 37] It is uniquely suited to initiatives that are rolled out over time.

Another strength of our study is the mixed methods approach. The secondary outcome measures and qualitative data can provide supporting evidence of change attributed to NetworkZ and provide complementary insights into whether the programme is working as intended and why the programme is or is not working.

Weaknesses of the study

Team training interventions may take time to have an effect and we may thus fail to detect an effect due to sampling too early. Due to the limited timeframe for implementation of training in the initial cohort as dictated by our funding contract, cohort 1 DHBs were not selected randomly, but were purposively selected as we had an existing relationship with those DHBs. There are anticipated limitations of exploring the effect of quality improvement initiatives in the real world, including implications for standardisation, randomisation and control groups. For example, implementation of training depends on the willingness of DHBs to support NetworkZ, local staff to undertake the instructor training and DHB staff to attend the courses.

Despite numerous measures to achieve consistent, high-quality training, there are challenges in achieving standardisation of NetworkZ delivery in individual DHBs where the local staff take on the responsibility for implementing the programme in their own environments.

Due to the potential for other factors progressively influencing patient outcomes over time, we will not be able to prove that the changes can be attributed to NetworkZ. This will be mitigated to some extent through the stepped wedge design and triangulation with measures of teamwork through surveys and observations.

CONCLUSION

With the financial support of our publicly funded ACC, the unique opportunity exists to implement and evaluate a nation-wide team-training initiative for full operating theatre teams in all public hospitals in New Zealand. By using a pre-existing large administrative data set, we have the potential to demonstrate a difference to surgical patient outcomes from this quality improvement team-training initiative.

Author affiliations

[1]Centre for Medical and Health Sciences Education, The University of Auckland, Auckland, New Zealand

[2]Department of Anaesthesiology, The University of Auckland, Auckland, New Zealand

[3]Department of Medicine, Christchurch School of Medicine and Health Sciences, University of Otago, Christchurch, New Zealand

[4]Anaesthesia, Wellington Hospital, Wellington, New Zealand

[5]School of Biological Sciences, Victoria University of Wellington, Wellington, New Zealand

Acknowledgements The authors would like to express special gratitude to Kaylene Henderson, Jane Torrie, Richard Hamblin and Ian Civil for their tireless efforts in the development and delivery of the NetworkZ programme, and the ongoing support of its evaluation.

Contributors JMW and AFM conceived the original proposal and drafted the original manuscript. JMW, AFM, JAL, PB, DC, CF, ALG, MRM, CSW contributed to the development and refinement and statistical analysis of the protocol. All authors critically appraised the drafted manuscript and made important intellectual contributions to the writing. All authors read and approved the final submitted manuscript.

Funding This delivery and evaluation of NetworkZ is funded by the Accident Compensation Corporation of New Zealand, and supported by grants from the Australian and New Zealand College of Anaesthetists (2017 Douglas Joseph Professorship Grant, Project Number DJ17/001), and the Lottery Health Research Fund (Project Number: R-LHR-2017-49141).

Competing interests AM is Chair of the Health Quality and Safety Commission NZ and a director of SaferSleep LLC. Both AM and CSW hold shares in SaferSleep LLC.

Patient consent for publication Not required.

Ethics approval This study has been approved by the New Zealand Health and Disability Ethic Committee (HDEC) (#16/NTB/143).

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs
Jennifer Weller http://orcid.org/0000-0002-8752-9286
Craig S Webster http://orcid.org/0000-0002-6997-4263

REFERENCES

1 Jha AK, Larizgoitia I, Audera-Lopez C, *et al*. The global burden of unsafe medical care: analytic modelling of observational studies. *BMJ Qual Saf* 2013;22:809–15.
2 de Vries EN, Ramrattan MA, Smorenburg SM, *et al*. The incidence and nature of in-hospital adverse events: a systematic review. *Qual Saf Health Care* 2008;17:216–23.
3 Sutcliffe KM, Lewton E, Rosenthal MM. Communication failures: an insidious contributor to medical mishaps. *Acad Med* 2004;79:186–94.
4 Lingard L, Espin S, Whyte S, *et al*. Communication failures in the operating room: an observational classification of recurrent types and effects. *Qual Saf Health Care* 2004;13:330–4.
5 Gawande AA, Zinner MJ, Studdert DM, *et al*. Analysis of errors reported by surgeons at three teaching hospitals. *Surgery* 2003;133:614–21.
6 Lingard L, Reznick R, Espin S, *et al*. Team communications in the operating room: talk patterns, sites of tension, and implications for novices. *Acad Med* 2002;77:232–7.
7 Gillespie BM, Chaboyer W, Wallis M, *et al*. Why isn't 'time out' being implemented? An exploratory study. *Qual Saf Health Care* 2010;19:103–6.
8 Lingard L, Reznick R, DeVito I, *et al*. Forming professional identities on the health care team: discursive constructions of the 'other' in the operating room. *Med Educ* 2002;36:728–34.
9 Shekelle P, Wachter R, Pronovost P, *et al*. *Making health care safer II: an updated critical analysis of the evidence for patient safety practices*. Rockville: MD: Agency for Healthcare Research and Quality, 2013.
10 Salas E, Rosen MA. Building high reliability teams: progress and some reflections on teamwork training. *BMJ Qual Saf* 2013;22:369–73.
11 Buljac-Samardzic M, Dekker-van Doorn CM, van Wijngaarden JDH, *et al*. Interventions to improve team effectiveness: a systematic review. *Health Policy* 2010;94:183–95.
12 Weaver SJ, Dy SM, Rosen MA. Team-training in healthcare: a narrative synthesis of the literature. *BMJ Qual Saf* 2014;23:359–72.
13 Hughes AM, Gregory ME, Joseph DL, *et al*. Saving lives: a meta-analysis of team training in healthcare. *J Appl Psychol* 2016;101:1266–304.
14 Jones KJ, Skinner AM, High R, *et al*. A theory-driven, longitudinal evaluation of the impact of team training on safety culture in 24 hospitals. *BMJ Qual Saf* 2013;22:394–404.
15 Low XM, Horrigan D, Brewster DJ. The effects of team-training in intensive care medicine: a narrative review. *J Crit Care* 2018;48:283–9.
16 Salas E, Zajac S, Marlow SL. Transforming health care one team at a time: ten observations and the TRAIL ahead. *Group Organ Manag* 2018;43:357–81.
17 Accident Compensation Corporation. *Supporting patient safety: treatment injury information*. Wellington: ACC, 2017.
18 Weller J, Cumin D, Torrie J, *et al*. Multidisciplinary operating room simulation-based team training to reduce treatment errors: a feasibility study in New Zealand hospitals. *N Z Med J* 2015;128:40–51.
19 Weller J, Civil I, Torrie J, *et al*. Can team training make surgery safer? Lessons for national implementation of a simulation-based programme. *N Z Med J* 2016;129:9–17.
20 Weller JM, Cumin D, Civil ID, *et al*. Improved scores for observed teamwork in the clinical environment following a multidisciplinary operating room simulation intervention. *N Z Med J* 2016;129:59–67.
21 Kolb DA. *Experiential learning: experience as the source of learning and development*. Englewood Cliffs: FT press, 2014.
22 Boud D, Keogh R, Walker D. Promoting reflection in learning: a model. In: Edwards R, Hanson A, Raggatt P, eds. *Boundaries of adult learning*. New York: Routelage, 1996: 32–56.
23 Salas E, Sims DE, Burke CS. Is there a "Big Five" in Teamwork? *Small Group Res* 2005;36:555–99.
24 Bate P, Mendel P, Robert G. *Organizing for quality: the improvement journeys of leading hospitals in Europe and the United States*. Oxford and New York: Radcliffe Publishing, 2008.
25 Dixon-Woods M, McNicol S, Martin G. Ten challenges in improving quality in healthcare: lessons from the Health Foundation's programme evaluations and relevant literature. *BMJ Qual Saf* 2012;21:876–84.
26 Dixon-Woods M, Bosk CL, Aveling EL, *et al*. Explaining Michigan: developing an ex post theory of a quality improvement program. *Milbank Q* 2011;89:167–205.
27 Hemming K, Haines TP, Chilton PJ, *et al*. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. *BMJ* 2015;350:h391.
28 Ministry of Health. Services delivered: acute and elective patient discharge volumes, 2018. Available: https://www.health.govt.nz/publication/services-delivered-acute-and-elective-patient-discharge-volumes [Accessed 8 May 2018].
29 Devcich DA, Weller J, Mitchell SJ, *et al*. A behaviourally anchored rating scale for evaluating the use of the WHO surgical safety checklist: development and initial evaluation of the WHOBARS. *BMJ Qual Saf* 2016;25:778–86.
30 Jerath A, Austin PC, Wijeysundera DN. Days alive and out of hospital: validation of a patient-centered outcome for perioperative medicine. *Anesthesiology* 2019;131:84–93.
31 Ariti CA, Cleland JGF, Pocock SJ, *et al*. Days alive and out of hospital and the patient journey in patients with heart failure: insights from the candesartan in heart failure: assessment of reduction in mortality and morbidity (CHARM) program. *Am Heart J* 2011;162:900–6.
32 Thompson JA, Davey C, Fielding K, *et al*. Robust analysis of stepped wedge trials using cluster-level summaries within periods. *Stat Med* 2018;37:2487–500.
33 Moore MR. *SteppedWedgeAnalysis: Simulation and Analysis of Stepped Wedge Data [program]*, 2020.
34 Singer SJ, Jiang W, Huang LC, *et al*. Surgical team member assessment of the safety of surgery practice in 38 South Carolina hospitals. *Med Care Res Rev* 2015;72:298–323.
35 Molina G, Berry WR, Lipsitz SR, *et al*. Perception of safety of surgical practice among operating room personnel from survey data is associated with all-cause 30-day postoperative death rate in South Carolina. *Ann Surg* 2017;266:658–66.
36 Medvedev ON, Merry AF, Skilton C, *et al*. Examining reliability of WHOBARS: a tool to measure the quality of administration of WHO surgical safety checklist using generalisability theory with surgical teams from three New Zealand hospitals. *BMJ Open* 2019;9:e022625.
37 Webster CS. Evidence and efficacy: time to think beyond the traditional randomised controlled trial in patient safety studies. *Br J Anaesth* 2019;122:723–5.