Protocol for –
A randomized controlled evaluation a training intervention to increase the use of statistical process control charts for hospitals in England

Kelly Ann Schmidtke,[1] Laura Kudrna,[2] Laura Quinn,[2] Paul Bird,[2,3] Karla Hemming,[2] and Richard Lilford[2]

[1] University of Warwick

[2] University of Birmingham

[3]West Midlands Academic Health Science Network

Kelly Ann Schmidtke https://orcid.org/0000-0001-5993-0358

Laura Kudrna https://orcid.org/0000-0002-8163-7112

Laura Quinn https://orcid.org/0000-0001-9660-4631

Paul Bird https://orcid.org/0000-0003-2044-8956

Karla Hemming https://orcid.org/0000-0002-2226-6550

Richard Lilford https://orcid.org/0000-0002-0634-984X

* Kelly Ann Schmidtke is the corresponding author and can be contacted at Kelly.A.Schmidtke@warwick.ac.uk

This research protocol was prepared in anticipation of using the SQUIRE 2.0 guidelines.[1]

Update Notification

Update Date	Change
02-Mar-2021	Updated Data Retrieval duration. In February 2022 it was realized that two trust in the waitlist control group could not schedule training in August but could schedule in July. To accommodate this scheduling conflict that would negatively affect our retention rate, a choice was made through consensus discussions with RL, KH, LK and KAS to change the waitlist control from a six-month waitlist control to a five-month waitlist control. All post-observation measurements (in the control and intervention groups) were moved forward 1 month.

Abstract

Background. Hospitals collect copious amounts of data to share with their board for quality assurance and improvement purposes. The way these data are presented can influence board members' decisions. For example, time-series charts highlight the highest and lowest data but do not clarify whether those data lie outside expected or 'common cause' variation. Statistical process control charts make this clarification and, in so doing, guide quality assurance and improvements in a more targeted fashion.

Local problem/Intervention. A previous study showed that data suitable for presentation as a control chart are seldom presented in that format. A training intervention called 'Making Data Count' was created to improve the uptake of statistical process control charts by hospitals in England. The current study will use a randomized design to evaluate whether the intervention increases control charts use for hospitals that were low performers and non-early adopters of the training intervention.

Methods. A parallel cluster randomized trial (with baseline-line measurements) across 20 National Health Service (NHS) hospitals in England. The hospitals will be randomly split into two groups. One group will be scheduled to experience the training intervention, and the other group will be placed on a waiting list to experience the training later. The primary analysis will compare the difference in the use of indicated control charts between waitlist control and intervention hospitals (adjusting for pre-intervention use) reported with 95% confidence intervals. A qualitative thematic analysis of feedback forms will be conducted.

Discussion. The present research will evaluate the impact of the training intervention on the use of control charts. The results will apply to institutions that are non-early adopters of this training intervention.

Keywords: Quality Improvement, Data Visualisation, Change Management, Learning Health System, Inservice Training

INTRODUCTION/BACKGROUND

Problem description

The way data are presented can influence board members' decisions.[2] Statistical process control (SPC) charts provide a data-driven approach to guide decisions related to quality improvement.[3,4] A previous study demonstrated that SPC charts are seldom used in National Health Service (NHS) hospital board papers in England,[5,6] and inspired an NHS-Improvement/England (NHS I/E) (2019) initiative called 'Making Data Count.' Making Data Count aims to help NHS staff create and use SPC charts effectively.[7] Starting in late 2017 this training was offered on a first-come, first-serve basis. As of April 2021, approximately 120 organizations have taken part. A non-randomized retrospective evaluation suggests that the training may be effective for early adopters,[8,9] but whether these benefits extend to hospitals less eager to take part is uncertain.

NHS-Improvement's team reviewed published board papers for every hospital in England in 2020 (at the time 217) to determine what presentation format was predominately used in their board papers to present quality and safety metrics. Seventy-five hospital board papers lacked SPC charts. Rather, these board papers were predominantly composed of Red-Amber-Green (R-A-G) data presentations or two-point comparisons, e.g., year 1 vs year 2, neither of which are data-driven presentation methods for quality improvement.[10]

Available knowledge

The theoretical argument for SPC charts is based on the concept of signal-to-noise ratios, and the need for a statistical method to distinguish between the signal (called special cause variations/unexpected) and the noise (called common cause variations/expected). The use of SPC charts in quality improvement methodology was pioneered by Walter Shewhart in 1920 while working for the Western Electric Company.[11] William Edwards Deming extended their use to new industries, e.g., Toyota. The use of SPC methods in healthcare emerged in

the 1960s mainly for laboratory processes and then for direct patient care processes.[12] A literature review located 40 studies about SPC chart usage in healthcare published between 1996 and 2017.[13] These studies evidence benefits of SPC chart use, e.g., to reduce surgical site infections,[14] to monitor mortality rates,[15] and to optimize daily staffing.[16]

Two multi-center cluster randomized controlled trials evaluating SPC chart interventions to improve patient safety have been conducted. In both trials, the control charts provided to the intervention groups were produced by external organizations. The first trial was conducted in England between 2001 and 2006, and it focused on hospital-acquired infections.[17] In this trial, 75 wards in 24 hospitals were randomized to an intervention group or a control group. Significant differences were not found across groups, plausibly due to contamination effects as nurses working in intervention wards could be redeployed to control wards within the same hospital. The second trial was conducted in France between 2014 and 2018, and it focused on adverse surgical events.[18] In this trial, the potential for spill-over effects were mitigated as entire hospitals were randomized to the intervention group or the control group. Here, hospitals in the intervention group experienced a significant decrease in adverse events, while hospitals in the control group did not.

Rationale

The current evaluation does not seek to further demonstrate the effectiveness of control charts themselves. Rather, we seek to evaluate the effectiveness of a training intervention to increase the use of control charts in NHS hospitals. While the previous evaluation suggests that the training intervention may be effective for early adopters,[19] the non-experimental nature of that evaluation means that we cannot be certain of cause and effect. Those early adopter hospitals that benefitted from the training could have been more motivated generally. In line with self-determination theory, the training intervention may be effective only or mainly for early adopters that self-select to take advantage of training when

offered.[20,21] On the other hand, the hospitals that have not self-selected to take up training may have greater room for improvement, and targeted invitations may increase such hospitals' motivation to engage and ultimately improve.[22]

Specific aims

Our main objective is to evaluate the effectiveness of a training intervention aiming to increase the use of SPC charts for hospitals in England identified as low performers and that did not self-select to take part in the training earlier, i.e., the non-early adopters. We also examine staff evaluations to improve future training sessions.

METHODS

Context

In England's NHS system, hospital trusts may be composed of single or multiple hospitals. Within the current paper, we refer to a single hospital trust as a "hospital" and multiple hospital trusts as "hospitals" to align with a more international nomenclature. As hospitals are hierarchically arranged organizations, successful implementation of an intervention to increase the use of control charts requires building bottom-up capacity (e.g., from the data analysts to produce control charts) and generating top-down support (e.g., line managers and board members).[23] At the time of this present evaluation, England is recovering from the COVID-19 pandemic, and hospitals are struggling with long waitlists.[24] In addition, all hospitals in England are set to become parts of statutory Integrated Care Systems by April 2022.[25]

Intervention[1]

The Making Data Count training intervention was designed by NHS-Improvement to improve knowledge about SPC charts and to increase their uptake. Training sessions are

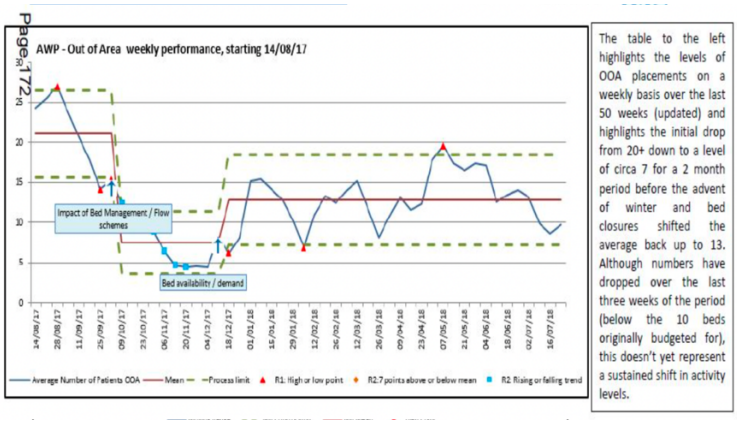
[1] The intervention will be described more briefly within the final manuscript and more fully in an appendix according to the TiDIER framework.[1]

tailored for two sets of attendees: board members and data analysts. Board member and analyst training sessions are delivered as close as possible in time, typically within the same month. Training sessions for board members are usually delivered over about one-and-a-half hours and focus more heavily on the benefits of control charts compared to other charts. Training sessions for analysts are usually delivered over three hours and focus more heavily on the structure and interpretation of the individual and moving range charts (X-mR charts).

The training sessions are delivered by one of two trainers with academic accreditation and work experience in data analytics. The training was originally delivered face-to-face at the hospitals, but since the COVID-19 pandemic has been adapted to an online environment and can now be offered in either format. The training sessions are tailored for each hospital, by creating, presenting, and discussing charts constructed from that hospital's recent data. Anonymized examples of the slides used during the training during this study will be provided as appendixes in the final manuscript. Guidebooks supplementing the training are available online.[26] Any modifications made to the training regime and any unexpected events that interfere with their planned delivery will be reported in the final manuscript.

During training sessions, trainees explore the strengths and weaknesses of different presentation methods. SPC charts are discussed in-depth to explain what they are, how to construct them, and why they are recommended. Data analysts are told how to prepare control charts to present in their board papers and to place the control limits at three-sigma. If analysts deviate from the 3-sigma rule they are asked to explain how and why in supporting text in a box near the control chart. The supporting text might also include explanations for performance variations and recommendations for improvements, see Figure 1.

Figure 1. Example provided for how supporting text should accompany charts.



A summary SPC icon system is introduced to help analysts quickly depict whether variations across many quality and safety measures are changing or whether targets are consistently met, see Figure 2a.[27] These icons allow analysts to be more selective about which control charts they ultimately present in their board papers, as the information contained in control charts not presented can be summarized on large dashboards, see Figure 2b. An example with supporting text is also included, see Figure 2c.

Figure 2. SPC icons

A) Legend describing what each SPC icon represents

Variation			Assurance		
Common cause – no significant change	Special cause of concerning nature or higher pressure due to (H)igher or (L)ower values	Special cause of improving nature or lower pressure due to (H)igher or (L)ower values	Variation indicates inconsistently hitting passing and falling short of the target	Variation indicates consistently (P)assing the target	Variation indicates consistently (F)alling short of the target

B) Example of SPC icons presented on a dashboard

SPC Summary Dashboard

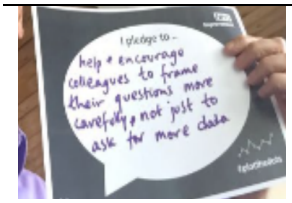
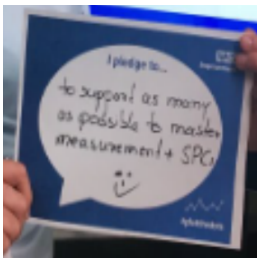
Key Performance Indicator	Target & Assurance	August 2018 & Variance	Average Performance
Sickness Rate (%)	4.6	4.8	5
Staff Turnover (%)	8.0	12.0	15
Appraisal (%)	95.0	96.0	90
Supervision (%)	85.0	96.0	92
Mandatory Training (%)	90.0	96.0	95
Staff FFT positive feedback (%)	90.0	84.0	86
Staff FFT response rate (%)	15.0	12.0	16
Records Management (%)	75.0	85.0	90

C) Example of SPC icons presented on dashboard with supporting text

	Jun-18	Target	Variation	Target Capability	Comment
Staff Sickness absence	4.4%	3.5%		F	Shift change in August 2017 showing increase in sickness - staff survey review indicated.....

Near the end of the first module of the data analysts training, analysts write a short commitment expressing how they intend to use their new skills. Such commitment devices are a behavior change technique that sharpens vague behavioral intentions and imposes costs on one's future self for failing to follow through.[28,29] Examples of previous commitments appear in Table 1.

Table 1. Two examples of commitments analysts made.

Picture	Hand-written pledge transcribed for clarity
	"I pledge to… help and encourage colleagues to frame their questions more carefully and not just to ask for more data"
	"I pledge to… to support as manage as possible to master measurement and SPC. ☺"

After the training session, trainees to fill out an anonymous feedback form. The feedback questions will be structured according to Kirkpatrick's four levels of evaluation: reaction, learning, behavior, and results.[30] A draft of the feedback form questions is presented in Table 2.

Table 2. Training feedback form questions.

Item	Response options
1. [Reaction] What is your overall reaction to the course? It was:	(a) very good (b) good (c) average (d) poor (e) very poor
2. [Learning] Did you learn anything new?	(a) yes (b) no
If A – then: Please describe the most valuable thing you learned:	[free text]
If B – then: Please tell us anything you think we should be covering in future sessions to improve the usage of data in hospitals: [free text]	[free text]
3. [Behavior] Do you intend to use any of the training over the next three months?	(a) yes (b) no
If A – then: Wonderful, please tell us one way you intend to use the training over the next three months.	[free text]
If B – then: We are sorry to hear that. Could you briefly explain why you do not intend to use the training? Your feedback will help us improve the session for future attendees.	[free text]
4. [Results] Do you think that adopting a SPC approach to data management can improve your organisations quality and safety measures?	(a) yes (b) no
5. Any other comments about today?	[free text]

Study of the intervention

Design

The current study is a quantitative evaluation of the effect of the Making Data Count training intervention on SPC charts indicated in hospital board papers, comparing those hospitals which are randomly allocated to either an intervention group or a waitlist control group. An analysis of the feedback forms will also be conducted to improve future renditions of the training.

Sample size calculation

We plan to include 20 hospitals, of which 10 will be randomly allocated to the intervention group and 10 to the waitlist control group. The sample size is based on pragmatic

reasons, as it needs to accommodate the number of hospitals that the Making Data Count team can schedule and train in a year. Delaying training longer than a year was deemed unreasonable by the Making Data Count team for hospitals planning to take up the training.

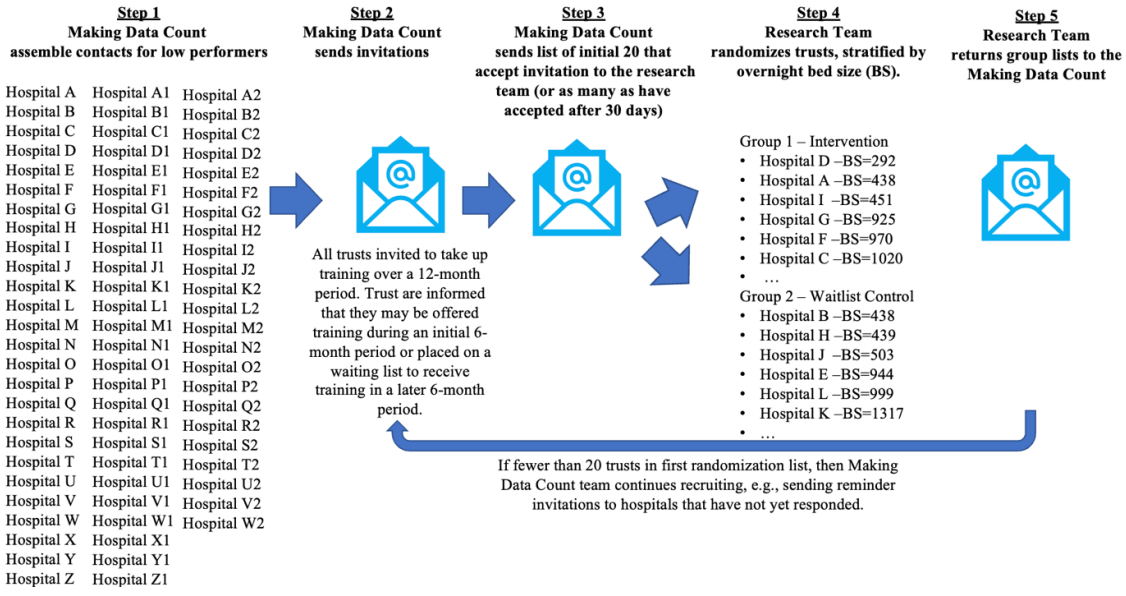
Based on our sample size calculation, a minimum of 16 hospitals with pre- and post-intervention measures is required to detect a 30-percentage point increase in the proportion of indicated SPC charts from 10% to 40% between control and intervention measures. The sample size was calculated with an alpha of 0.05 and a power of 0.80. Due to the study design (measurements taken of the outcome in a baseline period), adjustment for the correlation between pre- and post-intervention measures were made, estimated at r=0.90.[31] The primary analysis will be conducted based on intention-to-treat. While only 16 hospitals are required, we have planned to include 20 on an 'insurance principle" in case some hospitals drop out.

Selection and randomization of hospitals

At the present time, SPC training cannot be mandated. The hospitals that take part will ultimately depend on their willingness and availability to experience training. The five-step selection and randomization process for hospitals is presented in Figure 3. **Step 1** involves obtaining contact information for all the hospitals identified as relying on R-A-G charts or two-point comparisons to present quality and safety performance metrics in their board papers (75 at the time the protocol is being written). **Step 2** involves the Making Data Count team inviting the hospitals to take up training over a specified 12-month period. As part of their informed consent, hospitals will be made aware that they may be offered training over either the initial 6-months or the later 6-months in a randomized fashion. **Step 3** involves selecting the 20 hospitals that first express an interest in taking up the training or a smaller number which do so after 30 days have passed (note that recruitment will continue until 20 express an interest – see Step 5). The Making Data Count team will email the list of

selected hospitals to the research team. **Step 4** involves the research team randomizing the available hospitals to either the intervention group or waitlist control group. Randomization will be stratified based on the number of overnight beds from the most recently available count available from NHS England (dichotomized at the medium, to create "large" and "small" hospitals).[32] **Step 5** involves the research team emailing the randomized lists back to the Making Data Count team to commence scheduling. If the original list contains fewer than 20 hospitals, then the Making Data Count team will continue recruitment and another list of hospitals will be sent to the research team to randomize in the same stratified manner each month. The purpose of this is pragmatic – to ensure trainings can commence in a timely fashion, alongside the research being conducted.

Figure 3. Five steps in the selection and randomization process



Scheduling training

Hospitals allocated to the intervention group will be invited to schedule training over an initial five-month period. Hospitals allocated to the waitlist control group will be invited to schedule a training over a later five-month period. A hypothetical training schedule appears

in Figure 4. Approximately two hospitals may receive training each month, but the schedule will ultimately depend on hospital availability and training capacity.

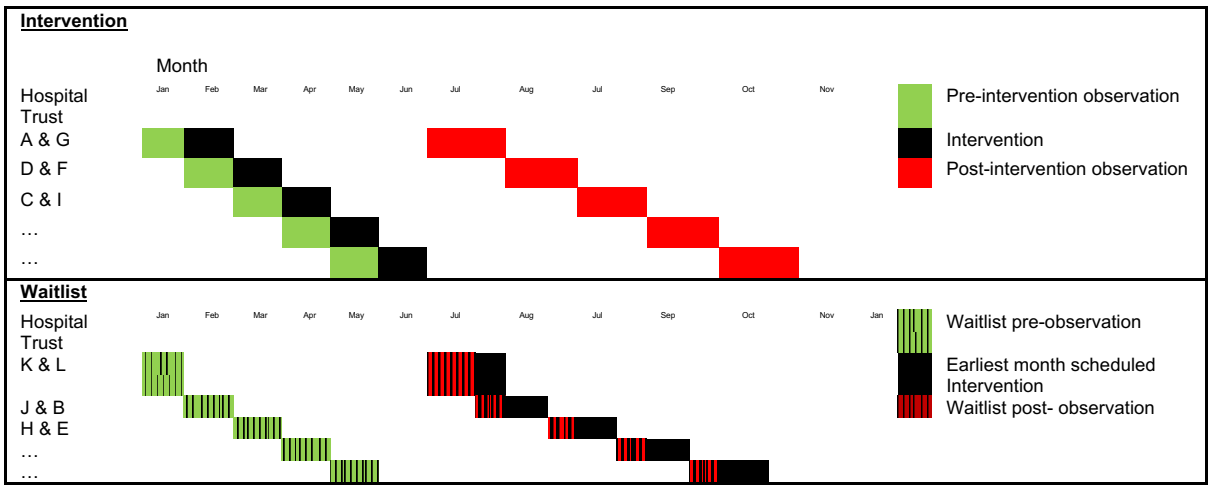
Figure 4. Example of potential schedule

Scheduled Training Order	Group 1 - Intervention
Month 1	Hospital A
Month 1	Hospital G
Month 2	Hospital D
Month 2	Hospital F
Month 3	Hospital C
Month 3	Hospital I
…	….
	Group 2 – Waitlist
Month 6	Hospital K
Month 6	Hospital L
Month 7	Hospital J
Month 7	Hospital B
Month 8	Hospital H
Month 8	Hospital E
…	…

Selection of board papers from hospitals

From each hospital, two board papers will be retrieved for a total sample of 40 papers. For the hospitals in the intervention group, we will retrieve the papers published in the nearest month before the intervention was delivered (pre-intervention observation) and approximately five months after the intervention was delivered (post-intervention observation). This selection process is represented in the top half of Figure 5. As boards do not publish their papers every month, it is not always possible to sample precisely one month pre-intervention or five months post-intervention.

Figure 5. Selected board papers retrieved from each hospital during the pre-intervention intervention and post intervention observation periods



Notes: The top box depicts the trusts that will experience the intervention (the intervention group), and the bottom box depicts trusts that will be placed on a waiting list to experience the intervention later (the control group). The black cells show the month of the training intervention. Green cells represent the month from which the pre-intervention papers will be retrieved, and the red cells represent the month from which the post-intervention papers will be retrieved.

For hospitals in the waitlist control group, we will form a matched list based on the order in which each is scheduled to receive training (earliest month) under the waitlist design. For example, the hospital in the intervention group scheduled to receive training first will be matched with the hospital in the waitlist group scheduled to receive training first, see the bottom half of Figure 5. This ordering is planned to minimize the possibility for waitlist hospitals to experience the training in the timespan planned for the current evaluation the waitlist control group's "post-observation" board papers will be selected from a month before each waitlist control hospital experiences training.

Outcome measures[2]

[2] The outcome measures were informed by the previous non-randomized studies discussed in the problem section and through conversations with trainers who designed the Making Data Count intervention.

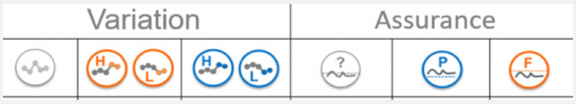
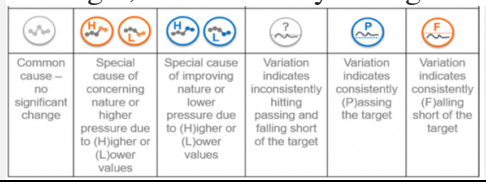
Our data will be made publicly available for replication and further analyses. The main outcome measure will be the proportion of SPC charts indicated about quality and safety measures (the numerator) out of all quality and safety charts indicated in the board papers (the denominator). SPC icons will be added to the numerator and the denominator where they indicate that a unique control chart has been used but not included in the paper. Additional secondary outcome measures will be extracted from the board papers and feedback forms, as described below.

Data extraction from board papers

Quantitative measures for data contained in the board papers will be created by coding information within each board paper. Four independent reviewers (R1, R2, R3, R4) will conduct the coding over three steps.

Step 1: R1 will download the papers. From each paper, R1 and R2 will independently extract the following three types of items: (A) the charts that appear, (B) SPC icons that appear indicating a control chart has been used, and (C) any supporting text aside charts that appear as instructed by the Making Data Count training, see Figure 1. They will then independently code aspects of the extracted charts and icons as described in Table 3.

Table 3. Step 1 coded elements.

Extracted item to be coded	Code Name	Explanation
Charts	Quality and Safety-Chart	whether the chart is about a quality or safety measure (yes/no)
SPC Icons	Quality and Safety-Icon	whether the icon is about a quality or safety measure (yes/no)
→If "yes" to Quality and Safety-Icon, then the following element also coded		
Quality and Safety SPC Icons	Icon indicating unique SPC chart	whether the SPC icons indicates a unique chart (yes/no) – a "yes" response indicates that this icon is added to the total numbers of control charts (numerator) and charts (denominator)
Quality and Safety SPC Icons	Icon – Purpose	whether the SPC icon is being presented for "assurance" purposes or to convey information about process "variation" 
Quality and Safety SPC Icons	Icon – Information	whether the SPC icon indicates common cause, special cause of a concerning nature, special cause of an improving nature, inconsistently hitting the target, consistently passing the target, or consistently failing the target 

SPC icons will not be extracted that simply inform the reader what they stand for, e.g., the legend in Figure 2a. The coded element named "Indicating unique SPC chart" reflects that when multiple icons refer to the same SPC chart only one SPC chart will be counted as unique, see Figure 6. Where an icon appears along with the SPC chart it represents, it will not be counted as unique, see Figure 7.

Figure 6. Example of two icons appearing together about the same SPC chart.

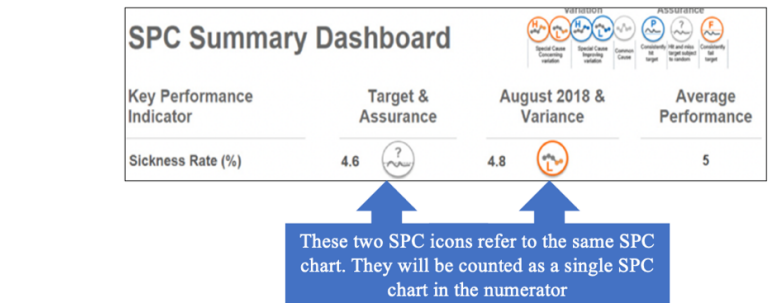
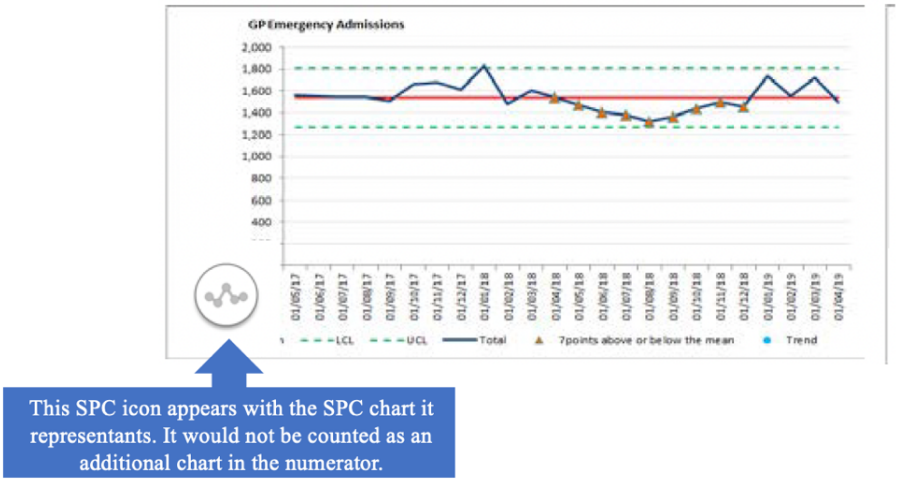


Figure 7. Example of an SPC icon appearing with the SPC chart it represents.



Step 2: R1 will take a screenshot of each quality and safety chart and the supporting text, removing identifying features, e.g., the name of the hospital and/or calendar dates, and send the redacted charts to R3 and R4. A screen shot of the supporting text will be taken, also removing identifying features.

Step 3: R3 and R4 will independently code features of the charts and supporting texts, which are described in Table 4.

Table 4. Step 4 coded elements

Extracted item to be coded	Code Name	Explanation
Quality and Safety Charts	Chart type	whether the chart is a time series chart, between subjects chart, time and between chart, or other (e.g., pie). The categories will be exclusive, such that only one may be selected.
Quality and Safety Charts	SPC?	whether the chart is a SPC chart (yes/no)
→If "yes" to SPC? then the following elements are also coded		
Quality and Safety Charts SPC Charts	Recalculations	whether the control limits are recalculated (yes/no)
Quality and Safety Charts SPC Charts	Special Cause Highlights	whether any data are highlighted as special cause (yes/no)
	…Special Cause Highlights-as recommended	and if "yes" whether the recommended blue/orange colors are used (yes/no)
Quality and Safety Charts SPC Charts	R-A-G[3]	whether any data are highlighted as R-A-G colors (yes/no)
	…R-A-G as NOT recommended	and if "yes" whether those R-A-G colors follow the former performance-based target model of data monitoring (yes/no)
Quality and Safety Charts SPC Charts	Labels	whether there are labels on the chart describing where process labels are set (yes/no)
	…Labels	and if "yes" AND are the labels different from "3-sigma" (yes/no)
Supporting text	Explaining where control lines are set	whether the text explains where the control lines are set (yes/no)
Supporting text	Reasons for variations	whether the text posits reasons for variation (yes/no)
Supporting text	Suggestions for improvements	whether the text suggests interventions (yes/no)

[3] Making Data Count trainers believe it is valuable to extract data about whether RAG colorings continue to be used for at least two reasons. First, RAG colors are predominantly used by the trust being invited to attend the training. Second, the training intervention explicitly informs attenders that RAG presentations are not well suited to quality improvement methodologies and urges attenders to move away from RAG presentations towards SPC presentations.

Blinding and agreement

Initially, codes will be independently assigned. Disagreements will be resolved through consensus discussions. Any disagreements that cannot be resolved will be referred to the chief investigator. It is not possible to blind reviewers R1 and R2 to the board papers group or time-period, as they are extracting charts from the published board papers. R3 and R4 will be blinded and will be instructed to inform R1 if they are unblinded to experimental group at any point, and whether/how often they are unblinded will be reported in the final manuscript.

Data extraction from feedback forms

Data extracted from feedback forms will include trainees' Likert ratings and free-text responses, see Table 2. These data will be compiled into an Excel file. R1 and R2 will independently code whether each free-text response is positive (e.g., "The training was practically useful."), negative (e.g., "The training went so fast that I couldn't keep up."), or neutral (e.g., "Handouts provided.").

Planned Analyses

Analysis of data from board papers

The hospitals will be described according to how many days it took them to accept the training invitation, along with their size (number of beds), and local deprivation status, as indicated on NHS Digitals Peer Finder Tool.[33] The inter-rater reliability of the data extracted from the board papers will be calculated using Kappa statistics and percentage agreement to quantity the level of agreement between reviewers. Descriptive statistics will enumerate the proportion of control charts that show the full detail and those appearing as icons.

The main analysis will examine the effect of the training intervention. For each hospital, we will observe the proportion of indicated SPC charts, including SPC icons, about quality and safety measures (the numerator) out of all quality and safety charts indicated in

the board papers (the denominator), for pre-intervention and post-intervention periods. To determine the absolute effect of the intervention we will compare the proportions of charts presented as SPC charts (between intervention and control arms) for each hospital and adjust for the proportion of SPC charts in the pre-intervention period. These differences will be compared between waitlist and intervention hospitals using a t-test and reported with 95% confidence intervals (a log transformation will be used if appropriate). To determine the relative effect (risk ratio) of the intervention, we will fit a zero-inflated negative Binomial regression model (outcome data are likely to be over-dispersed with a high number of zero counts), with the outcome of the number of SPC in the post-intervention period, fixed categorical effects for the intervention and the proportion of charts which were SPC charts in the pre-intervention period, and an exposure of the number of charts (in the post-intervention period). The offset here acts much like a denominator representing the number of opportunities each chart could have been depicted in an SPC format.

A sensitivity analysis based on the per-protocol group allocations will be performed if necessary. The remaining outcomes will be reported using descriptive statistics, e.g., counts, percentages, medians, and interquartile ranges. We will describe the proportion of control charts displayed as icons versus the full control chart.

Analysis of data from feedback forms.

Likert scale responses to each question on the feedback form will be described using medians and interquartile ranges. For the free-text responses, the number of positive, negative, and neutral comments will be reported. In addition, the available free-text responses will be thematically analyzed according to Braun and Clarke's method to reveal ways to improve future training or reasons others may (or may not) want to take it up.[34]

Ethical considerations

This research has received ethical approval to conduct from the University of Warwick Biomedical and Scientific Research Ethics committee (BSREC 100/20-21). The research team declares no conflicts of interest to disclose. The research team will work closely with NHS Improvement's Making Data Count team to design and conduct the study, and their support will be acknowledged in the final manuscript. NHS Improvement will have no role in the analysis, interpretation of data, or in writing the manuscript.

DISCUSSION

Summary

The summary will highlight whether the Making Data Count training intervention significantly increased control charts usage. Then it will compare the present results with those found in the previous, non-randomized evaluation.[35] We will explore how the responses to the feedback forms help us understand the effectiveness of the training. The literature will be reviewed for more recent publications nearer to the time this manuscript is written. Implications for clinical practice across types of care (e.g., health and social care) will be considered, along with the potential for the intervention's effectiveness to generalize to organizations outside England.

Limitations

Limitations are acknowledged within this protocol having to do with methods and the potential for our results to generalize. Regarding study methods, we are unable to blind the first set of reviewers to whether a hospital belongs to the intervention or control group. This is not possible because they will be looking directly at the board papers as they extract the initial information, and the board papers will inevitably contain identifiable information. The second set of reviewers will be blinded, as they will only be looking at extracted charts. Second, while board papers provide an objective measure of whether control charts are used, they are silent on how control charts are used. While the existence of control charts is a

necessary pre-condition for their being used well, it is possible that they are not being used as part of a broader quality improvement methodology that could positively impact patient care. In that sense, our study only seeks to establish whether the necessary conditions for SPC-based improvement have been met. A future study may investigate how SPCs are used.

Regarding generalization, one limitation is the amount of data we consider. Including additional hospitals may increase statistical precision but would increase the time and resources needed to deliver the study. As we have powered the study to locate large effects, we suspect that further precision would not change our interpretations. Further, a large effect is necessary on an upstream variable (in this case use of SPC methodology) that affects patient outcomes (such as patient safety) downstream.[36] Finally, the study focuses on hospitals, and that it cannot be assumed that the other diverse organizations have the necessary 'absorptive capacity' to respond to the training. NHS-Improvement plans to extend their training to primary care. A future study could be prospectively designed to assess whether and how Making Data Counts impacts quality improvement initiatives in primary care.

Declarations

Ethics approval and consent to participate: This research has been approved by the University of Warwick Ethics committee (BSREC 100/20-21). All participants who completed the feedback form will provide their informed consent to use their reactions in the present analyses.

Consent for publication: Not applicable.

Availability of data and materials: Any datasets used and/or analysed during the current study will be available from the corresponding author on reasonable request.

Competing interests: The authors declare that they have no competing interests.

Funding: This research is supported by the National Institute for Health Research (NIHR) Applied Research Centre (ARC) West Midlands. The views expressed are those of the author(s) and not necessarily those of the NIHR, ARC, or the Department of Health and Social Care. The funders had no role in the design of the study, and will have no role in the data collection, analysis, and interpretation of data and in writing the manuscript.

Acknowledgements: Mohammad Mohammad has read and approves of this protocol. The research team extends their gratitude to NHS Improvement, and in particular Samantha Riley and Karen Hayllar, for their support.

UNFORMATTED REFERENCES

[1] Ogrinc G, Davies L, Goodman D, Batalden P, Davidoff F, Stevens D. SQUIRE 2.0 (Standards for QUality Improvement Reporting Excellence): revised publication guidelines from a detailed consensus process. BMJ Qual Saf. 2016 Dec;25(12):986-992. doi: 10.1136/bmjqs-2015-004411. Epub 2015 Sep 14. PMID: 26369893; PMCID: PMC5256233.

[2] Marshall T, Mohammed MA, Rouse A. A randomized controlled trial of league tables and control charts as aids to health service decision-making. Int J Qual Health Care. 2004 Aug;16(4):309-15. doi: 10.1093/intqhc/mzh054. PMID: 15252005.

[3] Schmidtke KA, Watson DG, Vlaev I. The use of Control Charts by Laypeople and Hospital Decision-Makers for Guiding Decision Making. Quarterly Journal of Experimental Psychology. 2017;70(7):1114-1128. doi:10.1080/17470218.2016.1172096

[4] Mohammed MA, Cheng KK, Rouse A, Marshall T. Bristol, shipman, and clinical governance: Shewhart's forgotten lessons. Lancet. 2001 Feb 10;357(9254):463–7.

[5] https://www.hsj.co.uk/quality-and-performance/understanding-the-right-and-wrong-time-for-intervention/7015608.article

[6] Schmidtke KA, Poots AJ, Carpio J, et al Considering chance in quality and safety performance measures: an analysis of performance reports by boards in English NHS trusts BMJ Quality & Safety 2017;26:61-69.

[7] NHS Improvement. Making Data Count website. Retrieved 14 August 2019 from https://improvement.nhs.uk/resources/making-data-count/

[8] (Kudrna, et al in review at BMJ Quality and Safety).

[9] Riley S, Burhouse A, Nicholas T National Health Service (NHS) trust boards adopt statistical process control reporting: the impact of the Making Data Count Training

Programme BMJ Leader Published Online First: 30 April 2021. doi: 10.1136/leader-2020-000357

[10] Anhøj J, Hellesøe AM. The problem with red, amber, green: the need to avoid distraction by random variation in organisational performance measures. BMJ Qual Saf. 2017 Jan 1;26(1):81-4.

[11] Best, M., & Neuhauser, D. (2006). Walter A Shewhart, 1924, and the Hawthorne factory. Quality & safety in health care, 15(2), 142–143. https://doi.org/10.1136/qshc.2006.018093

[12] Thor, J., Lundberg, J., Ask, J., Olsson, J., Carli, C., Härenstam, K. P., & Brommels, M. (2007). Application of statistical process control in healthcare improvement: systematic review. Quality & safety in health care, 16(5), 387–399. https://doi.org/10.1136/qshc.2006.022194

[13] Suman, G., & Prajapati, E. (2018). Control chart applications in healthcare: a literature review. Int. J. Metrol. Qual. Eng. 9(5). https://doi.org/10.1051/ijmqe/2018003

[14] W.H. Woodall, S.L. Fogel, S.H. Steiner, The monitoring and improvement of surgical outcome quality, J. Qual. Technol. 47, 383–400 (2015)

[15] M.A. Mohammed, Using statistical process control to improve the quality of healthcare, Qual. Saf. Healthc. 13, 243–245 (2004)

[16] U. Gabbay, M. Bukchin, Does daily nurse staffing match ward workload variability? Three hospital experiences, Int. J. Healthc. Qual. Assu. 22, 625–641 (2009)

[17] Curran E, Harper P, Loveday H, Gilmour H, Jones S, Benneyan J, Hood J, Pratt R. Results of a multicentre randomised controlled trial of statistical process control charts and structured diagnostic tools to reduce ward-acquired meticillin-resistant Staphylococcus aureus: the CHART Project. J Hosp Infect. 2008 Oct;70(2):127-35. doi: 10.1016/j.jhin.2008.06.013. Epub 2008 Aug 23. PMID: 18723251.

[18] Duclos A, Chollet F, Pascal L, Ormando H, Carty M J, Polazzi S et al. Effect of monitoring surgical outcomes using control charts to reduce major adverse events in patients: cluster randomised trial BMJ 2020; 371 :m3840 doi:10.1136/bmj.m3840

[19] (Kudrna, et al in review at BMJ Quality and Safety).

[20] Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American Psychologist, 55, 68 –78. http://dx.doi.org/10.1037/0003- 066X.55.1.68

[21] Curado, C., Henriques, P. L., & Ribeiro, S. (2015). Voluntary or mandatory enrollment in training and the motivation to transfer training. International Journal of Training and Development, 19, 98 –109. http://dx.doi .org/10.1111/ijtd.12050

[22] Salas, E., Tannenbaum, S. I., Kraiger, K., & Smith-Jentsch, K. A. (2012). The science of training and development in organizations: What matters in practice. Psychological Science in the Public Interest, 13, 74 –101. http://dx.doi.org/10.1177/1529100612436661

[23] Stewart, Greg L. PhD; Manges, Kirstin A. RN; Ward, Marcia M. PhD Empowering Sustained Patient Safety, Journal of Nursing Care Quality: July/September 2015 - Volume 30 - Issue 3 - p 240-246 doi: 10.1097/NCQ.0000000000000103

[24] https://www.england.nhs.uk/2021/05/nhss-160-million-accelerator-sites-to-tackle-waiting-lists/

[25] https://www.england.nhs.uk/wp-content/uploads/2021/06/B0642-ics-design-framework-june-2021.pdf

[26] NHS Improvement. Making Data Count website. Retrieved 14 August 2019 from https://improvement.nhs.uk/resources/making-data-count/

[27] https://www.england.nhs.uk/wp-content/uploads/2019/12/making-data-count-strengthening-your-decisions.pdf Slide 27

[28] Robinson, C. D., Pons, G. A., Duckworth, A. L., & Rogers, T. (2018). Some Middle School Students Want Behavior Commitment Devices (but Take-Up Does Not Affect Their Behavior). Frontiers in psychology, 9, 206. https://doi.org/10.3389/fpsyg.2018.00206

[29] Dolan, P., Hallsworth, M., Halpern, D., King, D., Metcalfe, R., Vlaev, I. (2012). Influencing behaviour: The mindspace way. Journal of Economic Psychology Volume 33, Issue 1, February 2012, Pages 264-277

[30] Kirkpatrick DL, Kirkpatrick JD. 2006. Evaluating training programs: The four levels. 3rd ed. San Francisco: Berrett-Koehler Publication

[31] Frison L, Pocock SJ. Repeated measures in clinical trials: analysis using mean summary statistics and its implications for design. Statistics in Medicine. 1992;11(13):1685-704. https://doi.org/10.1002/sim.4780111304

[32] http://www.england.nhs.uk/statistics/statistical-work-areas/bed-availability-and-occupancy/bed-data-overnight/

[33] NHS Digital (2020). Peer Finder Too. Retrieved 18 May 2020 from https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/innovative-uses-of-data/multi-dataset-analysis/nhs-trust-peer-finder-tool

[34] Virginia Braun & Victoria Clarke (2006) Using thematic analysis in psychology, Qualitative Research in Psychology, 3:2, 77-101, DOI: 10.1191/1478088706qp063oa

[35] (Kudrna, et al in review in review at BMJ Quality and Safety).

[36] Lilford RJ, Chilton PJ, Hemming K, Girling AJ, Taylor CA, Barach P. Evaluating policy and service interventions: framework to guide selection and interpretation of study end points. BMJ. 2010 Aug 27;341:c4413. https://doi.org/10.1136/bmj.c4413