OPEN ACCESS

Cluster randomised evaluation of a training intervention to increase the use of statistical process control charts for hospitals in England: making data count

Kelly Ann Schmidtke ,[1,2] Laura Kudrna ,[3] Laura Quinn ,[3] Paul Bird,[4,5] Karla Hemming,[6] Zoe Venable,[1] Richard Lilford [7]

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/bmjqs-2024-017094).

For numbered affiliations see end of article.

Correspondence to
Dr Kelly Ann Schmidtke; Kelly.Schmidtke@uhsp.edu

Received 8 January 2024
Accepted 1 August 2024

To cite: Schmidtke KA, Kudrna L, Quinn L, *et al*. *BMJ Qual Saf* Epub ahead of print: [please include Day Month Year]. doi:10.1136/bmjqs-2024-017094

ABSTRACT

Background The way that data are presented can influence quality and safety initiatives. Time-series charts highlight changes but do not clarify whether data lie outside expected variation. Statistical process control (SPC) charts make this distinction and have been demonstrated to be effective in supporting hospital initiatives. To improve the uptake of the SPC methodology by hospitals in England, a training intervention was created. The current study evaluates the effectiveness of that training against the background of a wider national initiative to encourage the adoption of SPC charts.

Methods A parallel cluster randomised trial was conducted with 16 English NHS hospitals. Half were randomised to the training intervention and half to the control. The primary analysis compares the difference in use of SPC charts within hospital board papers in a postrandomisation period (adjusting for baseline use). Trainees completed feedback forms with Likert scale and open-ended items.

Results Fifteen hospitals participated across the study arms. SPC chart use increased in both intervention and control hospitals between the baseline and postrandomisation period (29 and 30 percentage points, respectively). There was no statistically significant difference between the intervention and control hospitals in use of SPC charts in the postrandomisation period (average absolute difference 9% (95% CI −34% to 52%). In the feedback forms, 93.9% (n=31/33) of trainees affirmed learning and 97.0% (n=32/33) had formed an intention to change their behaviour.

Conclusions Control chart use increased in both intervention and control hospitals. This is consistent with a rising tide and/or contamination effect, such that the culture of control chart use is spreading across hospitals in England. Further research is needed to support hospitals implementing SPC training initiatives and to link SPC implementation to quality and safety outcomes. Such research could support future quality and safety initiatives nationally and internationally.

Trial registration number NCT04977414.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Statistical process control (SPC) charts can guide improvement efforts better than data presentations that do not depict chance variation.
⇒ Prior to 2017, SPC charts were seldom included in quality and safety charts presented to hospital boards in the English National Health Service (NHS).
⇒ A training intervention was rolled out in the NHS and an observational study showed improved uptake of SPC methodology.

WHAT THIS STUDY ADDS

⇒ A prospective randomised controlled trial among slow adopters of the above training did not show a 'treatment effect' but uptake improved in both groups.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ SPC charts are becoming a new standard of reporting for NHS hospital trusts, thereby reducing the headroom for further improvements to show in a trial.

INTRODUCTION/BACKGROUND

Conversations between hospital board members, middle managers and front-line staff often revolve around performance metrics, for example, wait times, mortality rates, and hospital-acquired infection rates.[1] How these data are presented influences the selection of

quality improvement initiatives.[2] Where data presentations focus on two-point comparisons (before/after) or performance targets (often visualised with Red-Amber-Green codes), improvement initiatives could chase statistical noise rather than sound statistical signals of poor or superior performance. Statistical process control (SPC) charts make the distinction between the signals and the noise clearer to better guide quality improvement initiatives.[3 4]

While SPC charts were developed to improve manufacturing processes, their utility extends to healthcare.[5] A 2017 literature review[6] found that SPC charts have been used for many purposes in healthcare, for example, to monitor mortality rates[7] and to optimise staffing levels.[8] Two multicentre cluster randomised controlled trials evaluating SPC chart interventions to improve patient safety have been conducted. The first found that SPC chart use reduced hospital-acquired infections.[9] The second found that SPC chart use decreased adverse surgical events.[10] As these trials were randomised, they allow for a causal inference about the effects of SPC chart use on patient safety. In both trials, SPC charts were produced by an external organisation, rather than by an in-house data analytics team.

Despite the benefits of using control charts, a 2017 study showed that SPC charts seldom appeared in hospital board papers in England.[11] A 2021 study reviewed board papers for every hospital in England and found that one-third (75/217) lacked any SPC charts and that most data were presented as two-point comparisons or Red-Amber-Green codes.[12] This paper describes a quality improvement initiative developed by the National Health Service in England (NHS England—previously NHS Improvement) to improve data presentations, called 'Making Data Count'. Making Data Count includes SPC training sessions to enable and motivate NHS institutions to produce SPC charts in-house. In 2022, a retrospective evaluation of the SPC training sessions was conducted comparing the proportion of SPC charts appearing in board papers from hospitals that adopted the training relatively early with a matched-control group that had not adopted training.[13] The results showed that the use of SPC charts in the postintervention period was nine times higher in the intervention group than in the control group, but the research was not based on a randomised design and differences related to unobservable factors of the implementation initiative could not be fully accounted for.[14]

The current research further evaluates the effectiveness of the Making Data Count SPC training with a cluster randomised controlled trial to better assess causality. Our main objective is to evaluate the effectiveness of a training intervention aiming to increase the usage of SPC charts in hospital board papers.

METHODS

A study protocol was published on the ClinicalTrials.gov (NCT04977414). The Standards for Quality Improvement Reporting Excellence guideline checklist was completed (online supplemental file 1).[15] No patients were involved in the conduct or write-up of this trial. The results of the trial have been reported back to the Making Data Count team and disseminated to public contributors involved with quality improvement efforts in NHS England.

Context

In England, hospital trusts may be composed of single or multiple hospitals. Within the current paper, we refer to a single trust as a 'hospital' and multiple trusts as 'hospitals' to align with international nomenclature. At the time of the present evaluation, England was recovering from the COVID-19 pandemic and hospitals were struggling with long waitlists.[16] As previously stated, the SPC training intervention is embedded within a larger quality improvement initiative called Making Data Count. The implementation efforts included presentations at national events, tweets about the training, and posting the training materials online.[14] Since the time of our trial, similar efforts to promote the uptake of SPC training have been rolled out in Australia.[17]

Intervention

The Template for Intervention Description and Replication(TIDieR) checklist describing Making Data Count's SPC training is provided in online supplemental file 2.[18] Training sessions are tailored for each hospital using their data and for two groups of trainees: board members (lasting 1.5 hours) and data analysts (lasting 3 hours). Board member sessions focus on the benefits of SPC charts compared with other charts. Analyst sessions focus on the structure and interpretation of the individual and moving range charts (X-mR charts), where control limits are set at three-sigma, that is, a 0.997 probability of the observation arising under a scenario where variation was entirely random around the mean value for the data considered. The analysts who construct SPC charts can share their thinking with decision-makers in a textbox. Here, they can justify deviations from the three-sigma standard, record possible explanations for variations, or recommendations for action. Analysts also use colour-codes to signal data likely to indicate common causes and potentially adverse or beneficial special causes. Figure 1 contains an example control chart documenting inpatient falls along with supplementary text describing that an increase in falls may be due to poor staffing. Preformatted Microsoft Excel templates are provided to hospitals, which are available from the Making Data Count section on the NHS Futures website.

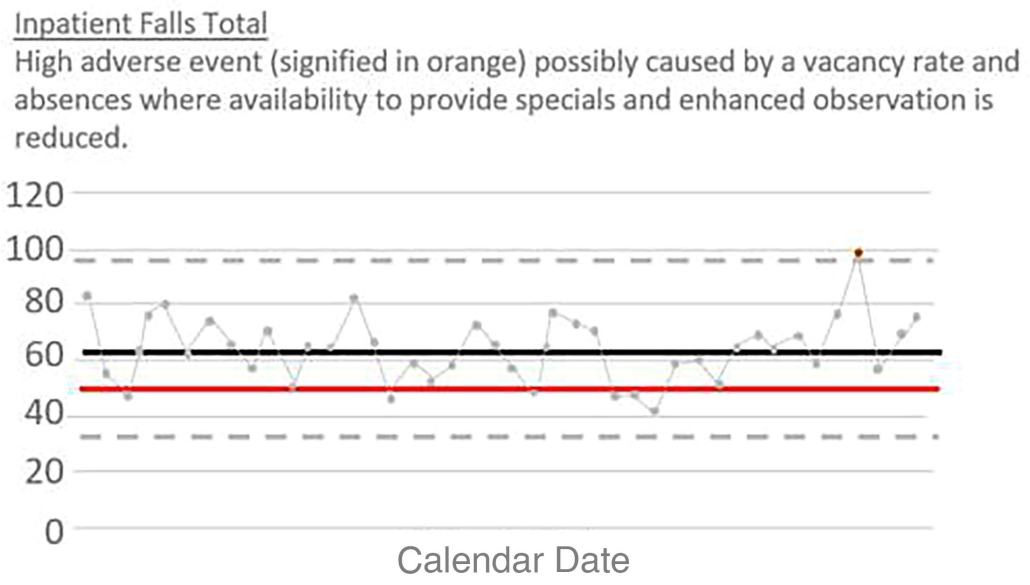

Figure 1 An example control chart with supporting text remade from a hospital board paper reviewed in this study. The black line represents the average fall rate and the grey dashed lines represent the three-sigma variation around that average. The red line represents a desirable target. The grey data represent variations falling within the dashed lines (expected variation) and the orange datum represents variations falling outside the dashed lines (special cause).

While analysts are encouraged to create a large number of charts, they are asked to be selective about which complete SPC charts are included in board papers. A novel method was used to summarise a number of complete SPC charts in a dashboard of 'icons' that then do not need to be included in their full form in the board papers. This reduces the heft of the board papers to make information easier to assimilate.

Figure 2 shows a dashboard of icons constituted of eight individual complete control charts.[19] Thus, board papers may contain a mix of complete SPC charts and SPC icons. In the present study, to capture the use of SPC charts we add together each complete SPC chart with each icon representing a unique complete SPC chart. As the icon methodology is a particular feature of our training programme, the use of the icons serves

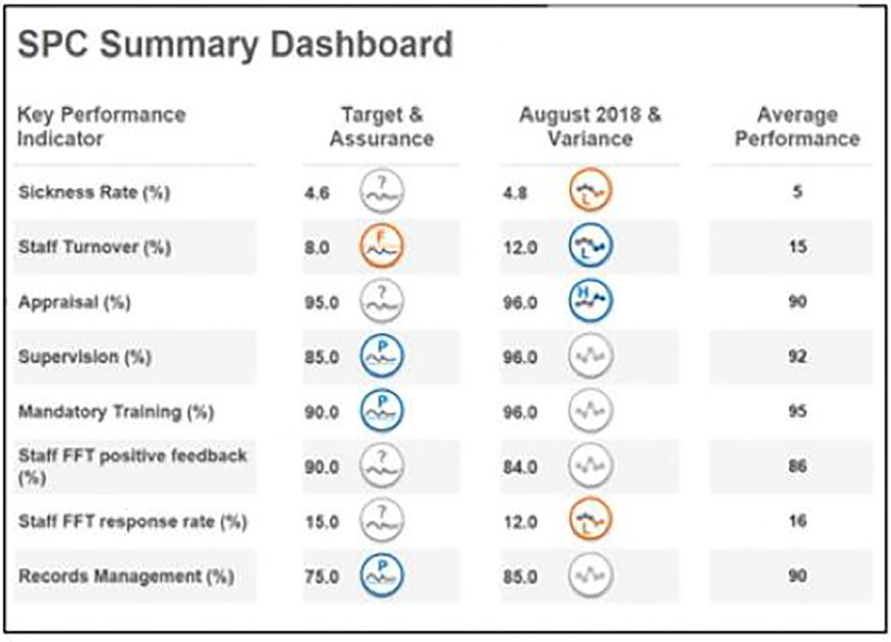

Figure 2 Example of statistical process control (SPC) icons presented on a dashboard. FFT stands for the "Friends and Family Test". The orange colour represents performance that is deteriorating (variation) or consistently below set targets (assurance). Blue colouring represents performance that is improving (variation) or consistently above targets (assurance). Grey colouring represents performance for which there is no significant change (variation) or for which performance is inconsistently hitting or falling short of targets.

as a sign that the principles taught in the SPC training intervention sessions have influenced local practice.

Study of the intervention

Design

A parallel cluster randomised trial was conducted. Hospitals were randomly allocated to either the intervention arm to receive the training or a waitlist control arm where they would be offered the same training at a later date. The use of SPC charts/icons was evaluated by monitoring the proportion of SPC charts/icons in the board papers during both a baseline period and a postrandomisation period. An analysis of the training feedback forms was conducted to improve future training.

Sample size calculation

Our sample size calculation suggested that a minimum sample of 16 hospitals in 2 equal groups with premeasures and postmeasures would provide 80% power to detect a 30 percentage point difference (ie, from 10% to 40%) between the intervention group and the control group in the use of SPC charts/icons (primary outcome) with an alpha of 0.05, and a correlation coefficient of 0.90 between baseline and postrandomisation measures.[20] This calculation assumed a t-test of the cluster-level proportions adjusting for baseline values and was implemented in the Shiny CRT Calculator to determine the required sample size.[21 22]

Selection and randomisation of hospitals

The Making Date Count team of NHS England invited 75 hospitals identified as not using the SPC methodology to participate. Hospitals were made aware that they would be randomised to receive training either immediately (intervention arm) within the context of the trial or to receive it at a later date (control arm). Of the 75 hospitals invited, 16 expressed willingness and availability to take part within a month of receiving the invitation. The research team then randomised hospitals to the intervention or control groups. Randomisation was stratified based on the number of overnight beds dichotomised at the median to create 'larger' and 'smaller' hospitals.[23]

Scheduling training

The Making Data Count team scheduled training for the intervention group over 6 months based on their availability to provide the training and the hospitals' availability to receive the training. These training dates were scheduled after randomisation and dependent on pragmatic and logistical constraints. The training dates for the control arm (under a waitlist design) were scheduled in a similar way—again dependent on pragmatic and logistical constraints, with the additional constraint that they would all run only after all the intervention hospitals had received their training.

Selection of board papers from hospitals

Our data collection parallels the previously conducted retrospective trial.[13] Two board papers were retrieved from each hospital. For the intervention group, we retrieved the papers published in the nearest month before the intervention was delivered (baseline observation) and then approximately 5 months after the intervention was delivered (postrandomisation observation). As boards do not publish papers every month, it is not always possible to sample precisely 1 month before or 5 months after the delivery of the intervention. Where a planned month was not available, the nearest month pre/post was selected. For the control group, we planned to match the months for which board papers were retrieved for the intervention group, based on the order in which the control group received their waitlist version of the training. A conceptual diagram describing how this training could be scheduled is provided in the preregistration protocol.

Outcome measures

In line with previous research,[13 24] our primary outcome measure was the use of 'SPC charts/icons' in board papers. This was defined as the number of unique SPC charts/icons about quality and safety measures (the numerator) out of the number of all quality and safety charts/icons in the board papers (the denominator). This was obtained by summing the complete control charts and unique control charts indicated by individual summary icons. The summary dashboard of icons represented in figure 2 includes icons from eight unique and complete SPC charts not otherwise represented in the board paper. For example, if the board paper contained just one complete SPC chart (figure 1) and these SPC icons (figure 2), then this hospital would have nine control charts.

A secondary outcome measure was the use of complete 'SPC charts (excluding icons)' in board papers.

Data extraction from board papers

The detailed data extraction instructions are provided in online supplemental file 3. Briefly, the board papers were retrieved by a single researcher. Then, two researchers independently identified charts and SPC icons describing quality and safety data. Disagreements were resolved through consensus discussions. One researcher also extracted any supporting textboxes (figure 1) for review by the next set of researchers.

Next, the first reviewer removed identifying features from charts and supporting text, for example, hospital names and calendar dates. The redacted charts were sent to two new researchers who independently coded chart type (line, bar, line and bar or other), and whether the chart was an SPC chart. SPC charts were further coded noting whether the control lines were recalculated, whether special causes were highlighted (if yes,

they coded whether special causes were coloured as recommended by training), and whether the process limits were labelled (if yes, they coded whether the labels differed from three-sigma). For the supporting text, these researchers coded whether it explained where control lines were set, whether reasons for variations were stated and whether suggestions for improvements were provided.

Data extraction from feedback forms

The feedback forms were structured according to Kirkpatrick's four levels of evaluation: reaction, learning, behaviour change, and results (online supplemental file 4).[25] Trainees reported their overall reaction using a Likert scale (1=very poor to 5=very good). Then, trainees reported whether they learnt anything and whether they intended to change their behaviour with opportunities to describe what and how in open-text boxes. Next, trainees reported whether they believed the training would impact the organisation, that is, the results. Lastly, trainees could leave additional comments. All feedback is based on self-reports, and no further measures of the learning or results levels were collected. Behaviour change is indicated by our outcome variables, that is, SPC chart/icon use and SPC chart use (excluding icons).

Analyses

Data from board papers

Hospital characteristics were described including hospital size (based on the number of beds) and local deprivation status (based on NHS Digital Peer Finder Tool). Inter-rater reliability of the data extracted from board papers were calculated using percentage agreement and prevalence-adjusted bias-adjusted kappa (PABAK). Chart characteristics were described using counts and percentages.

For the primary analysis, the use of SPC charts/icons is presented as a proportion by extracting the number of SPC charts/icons about quality and safety measures (numerator) and dividing this by the number of all quality and safety charts/icons in the board paper (denominator), for baseline and postrandomisation periods. We first describe the change in use of SPC charts/icons between baseline and postrandomisation periods for each hospital, stratified by intervention and control groups. To determine the absolute effect of the intervention, using a cluster-level analysis, we used a t-test adjusting for baseline proportion of SPC charts/icons. Then, again, using a cluster-level analysis, to determine the relative effect (rate ratio (RR)) of the intervention, we fit a zero-inflated negative binomial regression model (outcome data likely to be overdispersed with a high number of zero counts), with outcome as the number of SPC charts/icons in the postrandomisation period, adjusting for the group (intervention or control) and proportion of SPC charts/icons in the baseline period (with offset number of charts). In a post hoc sensitivity analysis, to further investigate robustness of approach to overdispersed data, we fitted various alternative models. Results are not presented but were similar.

For the secondary analysis, we examined the effect of the intervention where SPC icons were not included in the number of SPC charts, that is, SPC charts (excluding icons). All estimates were reported with 95% CIs. All analyses were done on an intention-to-treat basis. There was no missing data. One site dropped-out before any data were collected. All analyses were completed in Stata V.18. The quantitative data are provided in online supplemental file 5.

Data from feedback forms

Responses to the feedback survey are described using a mean and SD for the reaction item and percentages of affirmative responses for the learning, behaviour, and results items. Text describing what participants learnt, how they intended to change, and additional comments are described using an inductive, semantic and (critical) realist approach.[26] Excel was used to view responses and apply codes. The initial framework contained each open-ended item and further codes were inductively derived to describe responses for each item by one researcher (KAS).[26] Codes were then further revised with two additional researchers who coded the charts for our primary analysis (PB and ZV). To promote transparency, the research team reflectively decided to report how often codes were applied; the quantitative findings should not be taken to indicate the generalisability of findings.

RESULTS

Hospital characteristics

Of the eight hospitals allocated to the intervention, one withdrew after the first month and was excluded from data collection thereafter. On average, intervention hospitals were similar in size; the mean was 657 (SD=628) in intervention hospitals and 588 (SD=516) in control hospitals. Deprivation (on the index of multiple deprivation) was descriptively higher in the intervention (median=26; range 13–32) than control hospitals (median=15; range 12–41).

Figure 3 shows the order in which the hospitals were trained and their board papers were retrieved. Training was delayed in three intervention hospitals, two of which took up training in the sixth month and one in the seventh month. However, all trainings in the control group still occurred after all trainings in the intervention group. Of the eight control hospitals, three took up training after the training intervention hospitals in the 7th month and one in the 12th month. The remaining four control hospitals did not complete training within the designated period. The timing of the data retrieval for these four control hospitals was matched with the final two intervention hospitals using a random number generator in Excel. The

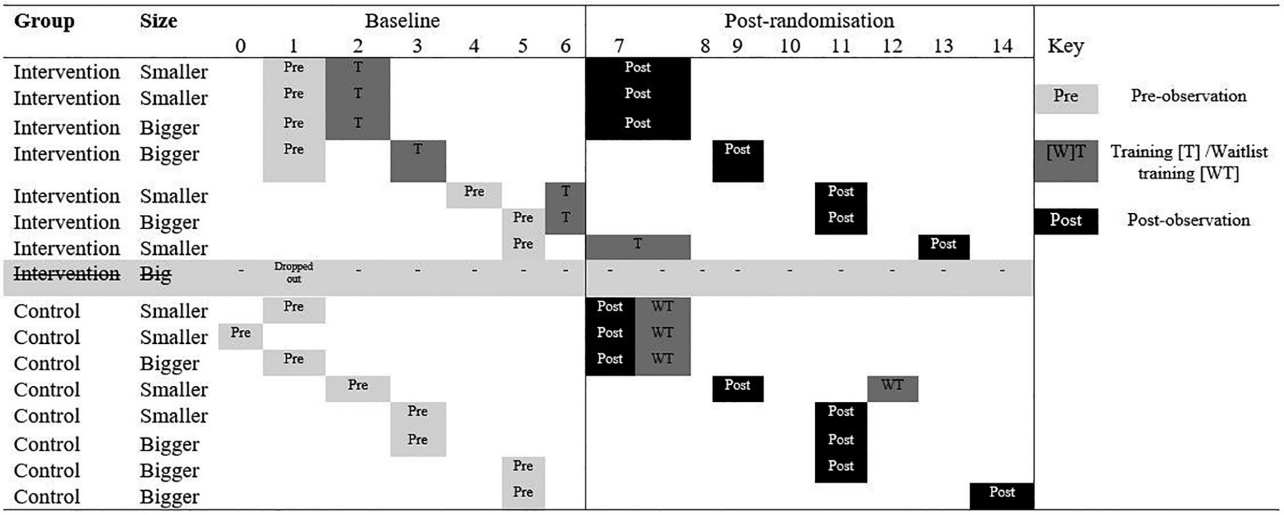

Figure 3 Board papers retrieved from each hospital during the baseline and postrandomisation observation periods. The top half depicts the hospitals that experienced the intervention (the intervention group) and the bottom half depicts hospitals that were placed on a waiting list to experience the intervention later (the control group). The darker grey cells with a 'T' show the month of the training intervention for intervention hospitals and the 'WT' shows the month of the waitlist control groups training. The lighter grey cells with 'pre' represent the month from which the baseline papers were retrieved, and the black cells with 'post' represent the month from which the postrandomisation papers were retrieved.

control hospitals' postrandomisation board papers were always published before any training took place. The intervention hospitals' postrandomisation board papers were always published after their trainings took place.

Inter-rater reliability and blinding

Agreement between reviewers for deciding if charts were quality and safety charts was 89.6% (95% CI 88.6 to 90.7), PABAK was 0.79 (95% CI 0.77 to 0.81). There was high agreement among reviewers when deciding whether an SPC icon indicated a unique chart (83.6%, 95% CI 80.3 to 86.9). No reviewers reported being unblinded to the hospital or the time period (online supplemental file 6).

Chart characteristics

In total 3182 unique charts/icons were identified, of which 1409 (44.3%) contained quality and safety data. Of these 1409 unique charts, the majority were line charts (914, 65%), followed by bar charts (209, 15%), SPC icons (156, 11%), other charts (83, 6%) and bar and line charts (47, 3%).

Three hundred and thirty-eight charts were complete SPC charts, of which 332 were line charts and/or bar charts (98%) and 6 (2%) were funnel charts; and 50 (16%) had recalculated control limits. Out of these 338 complete control charts, 213 (213/338= 63%) had special cause data highlighted of which 163 (163/213=76%) used the recommended colours (blue and orange). SPC icons were featured in six of the seven intervention hospitals and three of the eight control hospitals; these hospitals are designated with '*' in tables 1 and 2. See online supplemental file 7 for further descriptions.

Effects of training intervention on SPC charts/icons use

The number and proportion of SPC charts/icons out of all quality and safety charts are presented for each hospital by group (intervention or control) and period (baseline vs postrandomisation period) in table 1. Looking at the number of SPC charts/icons (the numerator of our primary outcome), all seven intervention hospitals experienced increased use, while only four control group hospitals did.

On average in the control group, hospitals used 29 percentage points more SPC charts/icons in the postrandomisation period compared with the baseline period (95% CI −5 to 63). In the intervention group, there was a similar increase in the (absolute) percentage increase in use of SPC charts/icons from baseline to postrandomisation (average difference 30%, 95% CI 2 to 59). Adjusting for baseline differences, during the postrandomisation period the use of SPC charts/icons was 9% (−34% to 52%) higher for the intervention group compared with the control group, although CIs were wide and compatible with an effect in either direction. Inferences estimating the relative effect of the intervention were similarly inconclusive (RR 0.93 (95% CI 0.35 to 2.50)).

Effects of training intervention on complete SPC charts use (excluding icons)

The number and proportion of complete SPC charts (excluding icons) out of all quality and safety charts for both groups and periods are shown in table 2. On average, the control groups used 24% more SPC charts (excluding icons) in the postrandomisation period compared with the baseline period (95% CI −6 to 54). In the intervention group, there was only a 6% average increase (95% CI −18 to 30). Adjusting

Table 1 SPC chart/icon usage by group, hospital and period out of all quality and safety charts

Control group				Intervention group			
	Baseline period	Postrandomisation	Postbaseline		Baseline period	Postrandomisation	Postbaseline
Hospital	SPC chart/icon (%)	SPC chart/icon (%)	% difference	Hospital	SPC chart/icon (%)	SPC chart/icon (%)	% difference
1	0/0 (0)	24/41 (59)	59	9*	0/172 (0)	8/177 (5)	5
2	4/8 (50)	2/6 (33)	−17	10*	21/34 (62)	36/54 (67)	5
3*	5/38 (13)	101/110 (92)	79	11*	0/0 (0)	14/26 (54)	54
4	0/22 (0)	0/30 (0)	0	12*	49/69 (71)	86/103 (83)	12
5	0/40 (0)	0/41 (0)	0	13*	11/42 (26)	42/64 (56)	30
6*	0/0 (0)	8/35 (23)	23	14	0/21 (0)	3/29 (10)	10
7*	4/47 (9)	4/47 (9)	0	15*	11/90 (12)	54/55 (86)	74
8	0/0 (0)	7/8 (88)	88				
Average change (baseline vs post) in control group (95% CI)			29 (−5 to 63)	Average change (baseline vs post) in intervention group (95% CI)			30 (2 to 59)
				Absolute difference between intervention and control group† (95% CI)			9 (−34 to 52)
				Relative difference between intervention and control group‡ (95% CI)			0.93 (0.4 to 2.5)

For each hospital in baseline and postrandomisation period, the number of SPC charts/icons, the number of quality and safety charts and the percentage of SPC charts/icons out of all reported.
*At least one board paper included SPC icons.
†T-test comparing the average difference in proportions between the intervention and control group; adjusting for baseline proportions. Percentage difference and 95% CI are reported.
‡Zero-inflated negative binomial regression models. The outcome is the number of SPC charts and icons in the postrandomisation period, adjusting for the baseline proportion of SPC charts/icons. Exposure is all quality and safety charts. Rate ratios and 95% CIs are reported.
SPC, statistical process control.

for baseline differences, during the postrandomisation period, the use of SPC charts (excluding icons) was 18% (95% CI −53 to 17) lower for the intervention group compared with the control group (equivalent to an RR of 0.60 (95% CI 0.20 to 1.60).

Regarding supporting textboxes, 131 were identified. Only one textbox (1.0%) explained where the control lines had been set, for example, a wider interval was used acknowledging a willingness to accept more variation in wait times. Approximately a quarter offered reasons for the variation (n=34, 26.0%), for example, understaffing; or suggested possible interventions (n=35, 26.7%), for example, hiring or onboarding initiatives.

Analysis of feedback data from trainees

Thirty-three trainees provided feedback; 21 from 5 of the 7 intervention hospitals and 12 from 4 of the 8 control hospitals. Overall, trainees expressed satisfaction with their sessions (M=4.7 out of 5; SD=0.5).

Table 2 SPC chart (excluding icons) usage by group, hospital and period out of all quality and safety charts

Control group				Intervention group			
	Baseline period	Postrandomisation	Postbaseline		Baseline period	Postrandomisation	Postbaseline
Hospital	SPC chart (%)	SPC chart (%)	% difference	Hospital	SPC chart (%)	SPC chart (%)	% difference
1	0/0 (0)	24/41 (59)	59	9*	0/172 (0)	8/177 (5)	5
2	4/8 (50)	2/6 (33)	−17	10*	21/34 (62)	36/54 (67)	5
3*	5/38 (13)	60/110 (55)	41	11*	0/0 (0)	14/26 (54)	54
4	0/22 (0)	0/30 (0)	0	12*	49/69 (71)	50/103 (49)	−22
5	0/40 (0)	0/41 (0)	0	13*	11/42 (26)	3/64 (5)	−22
6*	0/0 (0)	8/35 (23)	23	14	0/21 (0)	3/29 (10)	10
7*	4/47 (9)	4/47 (9)	0	15*	11/90 (12)	14/55 (25)	13
8	0/0 (0)	7/8 (88)	88				
Average difference in control group (95% CI)			24 (−6 to 54)	Average difference in intervention group (95% CI)			6 (−18 to 30)
				Absolute difference between intervention and control group† (95% CI)			−18 (−53 to 17)
				Relative difference between intervention and control group‡ (95% CI)			0.6 (0.2 to 1.6)

For each hospital in baseline and postrandomisation period, the number of SPC charts (excluding icons), the number of quality and safety charts and the percentage of SPC charts (excluding icons) out of all reported.
Supporting textboxes completed by analysts.
*At least one board paper included SPC icons.
†T-test comparing the average difference in proportions between the intervention and control group. Percentage difference and 95% CI are reported.
‡Zero-inflated negative binomial regression models. The outcome is the number of SPC charts (excluding icons) in the postrandomisation period, adjusting for the baseline proportion of SPC charts (excluding icons). Exposure is all quality and safety charts. Rate ratios and 95% CIs are reported.
SPC, statistical process control.

Most affirmed having learnt something (n=31/33; 93.9%) and intended to change their behaviour in the next 3 months (32/33; 97.0%). All believed that the training would positively impact their hospital's future performance (33/33, 100%).

Ninety-nine free-text responses were provided (mean=12.6 words, SD=16.7). Of the 33 quotes about learning, the two trainees who said they did not learn anything reported having attended a previous session or that the training only covered basic information. The remaining trainees reported that they learnt reasons to adopt the SPC methodology (7/33; 21%), how to present data (17/33; 52%) or how to interpret data (12/33; 36%). Of the 33 quotes about behaviour change, the one trainee who said 'no' did not think the training applied to non-executive directors. The remaining quotes were about 'who' or 'what' would change. Regarding 'who' would change, 26% (8/33) described how they or their team would change and 22% (7/33) described how others would change. Regarding 'what' would change, some trainees believed that organisational factors needed to change before they could personally change (4/33; 12.9%) or that they needed more training (4/33; 12.9%). Others discussed their intentions to simply use the SPC methodology (8/33; 25.8%), improve reporting (7/33; 22.6%) or to look at data differently (10/33; 32.3%).

Thirty-three participants provided additional comments. No negative comments appeared. Many took the opportunity to simply thank presenters (21/33; 64%). Two trainees (6%) wanted more content in the session, and 11 (33%) appreciated the presenters' passionate and knowledgeable approach to training that integrated data from their hospital.

DISCUSSION

Summary

The present prospective randomised trial evaluated the effectiveness of Making Data Count's SPC training intervention on the usage of control charts in hospital board papers. There was great variation in these changes. In the control group, changes ranged from −17% to 88%, and in the intervention group, changes ranged from 5% to 74%. While both groups increased on average, the difference in SPC chart/icon use between groups was not significant.

Possible interpretation

The randomised nature of the presented study allows us to conclude that the SPC training intervention was not a single casual factor producing increased control chart use among NHS hospitals in the context of our evaluation. Our finding that there was no intervention effect was coupled with the finding that there was increased uptake of SPC charts in both groups, intervention and control. A plausible explanation for our finding, therefore, is that the headroom for improvement was unexpectedly constrained. Such an effect could be due to a classical 'contamination' effect, a more general 'rising tide' effect spurred by the national implementation initiative, or a combination of both. Classical contamination effects occur when the intervention 'leaks' from the intervention sites to the control sites, thereby diluting differences between comparative groups. In contrast, a rising tide effect impacts sites whether they are in the study or not, thus causing changes across all hospitals in the direction anticipated by the intervention.[27]

A number of observations supports the rising tide interpretation. The use of control charts has been advocated since the 1980s, but a review of NHS hospital board papers in 2017 found little evidence of SPC use.[3 28] By the end of 2023, the training has been implemented in over 169 of England's NHS hospital trusts, and the Making Data Count team inform us that many early trainers have requested additional training to further improve their board papers. These observations suggest a pervasive temporal effect across the English health service. The rising tide interpretation is further supported by the appearance of icon methodology in the control hospitals' board papers. The control hospitals were not situated near the intervention hospitals, and since this method is bespoke to the NHS England's Making Data Count SPC training programme one must surmise that training spilled over from hospitals that had received the training to those that had not. Hospitals may be accessing this information from the board papers of peer hospitals as these are publicly available (which was part of the national initiative). Additionally, information may be spread by word-of-mouth from previous trainees, especially as our qualitative analysis of feedback forms suggests that the training was well received. A reasonable explanation for our findings is that headroom for improvement was reduced by control hospitals adopting similar interventions to the intervention hospitals, just by a different route. Other examples of rising tides have been cited in the literature.[28]

Strengths/Limitations

A strength of the present study involves its use of a randomised design to understand the causal relationship between training and control chart use. Several limitations of the current evaluation should be noted. Where possible, researchers attempt to control extraneous factors, but often cannot suppress factors that stakeholders believe could lead to overall improvements. The national initiative of implementation efforts that encouraged people to take up training was such an outside factor. The fact that the hospitals took up elements of the training before the training took place could not have been anticipated when this trial was preregistered. A combination of factors may have contributed, such as those described by the Consolidated Framework for Implementation Science and the MINDSPACE frameworks.[14 29] For instance,

promotional activities initiated by the national organisation, NHS Improvement, may have acted as an 'outer setting' factors encouraging implementation of the control chart approach irrespective of the training itself (ie, by eliciting a 'messenger' effect as described in MINDSPACE) and a gradual increase in SPC icons by peer hospitals may have acted as an individual or 'inner-setting' factor encouraging change (ie, through a 'social norm' effect as described in MINDSPACE).

Additional study limitations include that we were unable to blind the first set of reviewers to whether a hospital belongs to the intervention or control group as those reviewers extract initial information directly from the board papers. However, these reviewers did not code data for the primary outcome. Regarding generalisation, one limitation is the amount of data we consider. Including additional hospitals would increase statistical precision but would increase the time and resources needed to deliver the study which already practically informs future implementation efforts.

The present study provides little information about how control charts are used. Some information about how control charts could influence practices are provided in the supporting textboxes analysed in the present project, that is, explaining where control lines were set, reasons identified for variation, and suggested interventions. Previous qualitative research with board members suggests that control chart use redirects the focus of quality improvement practices.[12] Further information about how those charts influence practice is beyond the scope of the present study.

Lastly, the present study does not include information about patient outcomes. Training interventions are mediating variables. While a training session may directly influence the attitudes of trainees towards the intervention and their intended behaviour change, its downstream effects on patient safety will depend on further contextual factors.[23 30] While the existence of control charts is a necessary precondition 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.

CONCLUSIONS

The usage of control charts in hospital board papers increased for hospitals in both the intervention and control groups but experiencing SPC training was not a single causal factor. While this was not the anticipated effect, it is encouraging. The findings of this evaluation are important because they direct inquiry into other reasons why control chart use has increased beyond the training. Future work is needed to support hospitals in effectively implementing SPC methodologies to achieve improvements and to capture how upstream SPC training impacts downstream quality and safety outcomes. Such information can inform the implementation of other initiatives to increase quality and safety within hospitals and other health and social care settings.

Author affiliations
[1]Liberal Arts and Basic Sciences, University of Health Sciences and Pharmacy in St Louis, Saint Louis, Missouri, USA
[2]Warwick Medical School, University of Warwick, Coventry, UK
[3]University of Birmingham, Birmingham, UK
[4]Institute of Applied Health, University of Birmingham, Institute of Translational Medicine, University Hospitals Birmingham NHS Foundation Trust, Birmingham Health Partners, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK
[5]Health Innovation West Midlands, Birmingham, UK
[6]School of Health and Population Sciences, University of Birmingham, Birmingham, UK
[7]Institute of Applied Health Research, University of Birmingham, Birmingham, UK

X Richard Lilford @rjlilford

Contributors KAS wrote the initial draft of protocol and manuscript and revised subsequent drafts of protocol and manuscript. LK and PB attended NHSE SPC training and commented on drafts of the protocol and manuscript. LQ and KH contributed to the design of the study, developed the statistical analysis plan, commented on drafts of the protocol, wrote the results section of the manuscript and revised manuscript drafts. PB and ZV coded quality and safety charts. RL initiated idea, commented on drafts of protocol and manuscript and provided conceptual guidance. KAS acted as the guarantor. All authors have read and approved the manuscript.

Funding This study wad funded by Applied Research Centre (ARC) West Midlands (grant number: NIHR200165).

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by University of Warwick BSREC 100/20-21. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

ORCID iDs
Kelly Ann Schmidtke http://orcid.org/0000-0001-5993-0358
Laura Kudrna http://orcid.org/0000-0002-8163-7112
Laura Quinn http://orcid.org/0000-0001-9660-4631

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

REFERENCES

1 Soong C, Bell CM, Blackstien-Hirsch P. 'Show me the data!' Using time series to display performance data for hospital boards. *BMJ Qual Saf* 2023;32:69–72.
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;16:309–15.
3 Schmidtke KA, Watson DG, Vlaev I. The use of control charts by laypeople and hospital decision-makers for guiding decision making. *Q J Exp Psychol (Hove)* 2017;70:1114–28.
4 Mohammed MA, Cheng KK, Rouse A, *et al*. Bristol, Shipman, and clinical governance: Shewhart's forgotten lessons. *Lancet* 2001;357:463–7.
5 Thor J, Lundberg J, Ask J, *et al*. Application of statistical process control in healthcare improvement: systematic review. *Qual Saf Health Care* 2007;16:387–99.
6 Suman G, Prajapati D. Control chart applications in healthcare: a literature review. *Int J Metrol Qual Eng* 2018;9:5.
7 Mohammed MA. Using statistical process control to improve the quality of health care. *Qual Saf Health Care* 2004;13:243–5.
8 Gabbay U, Bukchin M. Does daily nurse staffing match ward workload variability? Three hospitals' experiences. *Int J Health Care Qual Assur* 2009;22:625–41.
9 Curran E, Harper P, Loveday H, *et al*. 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;70:127–35.
10 Duclos A, Chollet F, Pascal L, *et al*. Effect of monitoring surgical outcomes using control charts to reduce major adverse events in patients: cluster randomised trial. *BMJ* 2020;371:m3840.
11 Anhøj J, Hellesøe A-MB. The problem with red, amber, green: the need to avoid distraction by random variation in organisational performance measures. *BMJ Qual Saf* 2017;26:81–4.
12 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. *Leader* 2021;5:252–7.
13 Kudrna L, Bird P, Hemming K, *et al*. Retrospective evaluation of an intervention based on training sessions to increase the use of control charts in hospitals. *BMJ Qual Saf* 2023;32:100–8.
14 Damschroder LJ, Reardon CM, Widerquist MAO, *et al*. The updated Consolidated Framework for Implementation Research based on user feedback. *Impl Sci* 2022;17:75.
15 Ogrinc G, Davies L, Goodman D, *et al*. SQUIRE 2.0 (Standards for QUality Improvement Reporting Excellence): revised publication guidelines from a detailed consensus process. *BMJ Qual Saf* 2016;25:986–92.
16 NHS England. NHS's £160 million 'accelerator sites' to tackle waiting lists. Available: https://www.england.nhs.uk/2021/05/nhss-160-million-accelerator-sites-to-tackle-waiting-lists/ [Accessed 1 Oct 2023].
17 Royal Australasian College of Surgeons. Healthcare quality intelligence unite and health roundtable present: making data count - RACS: making data count (how to present data, UK NHS director sam riley). n.d. Available: https://events.teams.microsoft.com/event/5a86d55c-e951-420c-807a-c9c5f8776e98@35254406-146a-48d9-86e2-903d07f85ccb
18 Hoffmann TC, Glasziou PP, Boutron I, *et al*. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 2014;348:g1687.
19 NHS England and NHS Improvement. Making data count: strengthening your decisions. Available: https://www.england.nhs.uk/wp-content/uploads/2019/12/making-data-count-strengthening-your-decisions.pdf [Accessed 1 Oct 2023].
20 Frison L, Pocock SJ. Repeated measures in clinical trials: analysis using mean summary statistics and its implications for design. *Stat Med* 1992;11:1685–704.
21 Hayes RJ, Moulton LH. *Cluster randomised trials*. 1st edn. Chapman and Hall/CRC, 2009.
22 Hemming K, Kasza J, Hooper R, *et al*. A tutorial on sample size calculation for multiple-period cluster randomized parallel, cross-over and stepped-wedge trials using the Shiny CRT Calculator. *Int J Epidemiol* 2020;49:979–95.
23 NHS England. Bed availability and occupancy data – overnight. Available: http://www.england.nhs.uk/statistics/statistical-work-areas/bed-availability-and-occupancy/bed-data-overnight/ [Accessed 1 Oct 2023].
24 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 Qual Saf* 2017;26:61–9.
25 Kirkpatrick DL, Kirkpatrick JD. *Evaluating training programs: the four levels*. 3rd edn. San Francisco: Berrett-Koehler Publication, 2006.
26 Mercier J, Sanders J, Munford R. Fine companions: critical realism and framework analysis. *Int J Qual Methods* 2023;22.
27 Chen YF, Hemming K, Stevens AJ, *et al*. Secular trends and evaluation of complex interventions: the rising tide phenomenon. *BMJ Qual Saf* 2016;25:303–10.
28 Tennant R, Mohammed MA, Coleman JJ, *et al*. Monitoring patients using control charts: a systematic review. *Int J Qual Health Care* 2007;19:187–94.
29 Dolan P, Hallsworth M, Halpern D, *et al*. Influencing behaviour: the mindspace way. *J Econ Psychol* 2012;33:264–77.
30 Lilford RJ, Chilton PJ, Hemming K, *et al*. Evaluating policy and service interventions: framework to guide selection and interpretation of study end points. *BMJ* 2010;341:c4413.