Considering Chance in Quality and Safety Performance Measures:

An analysis of performance reports by boards in English NHS trusts

Schmidtke, K. A.,* Poots, A. J., Carpio, J., Vlaev, I., Kandala, N., & Lilford, R. J.

Keywords: Chart, Data Interpretation, Statistical, Governing Board

Word count, excluding title page, abstract, references, figures and tables: 3,591

*Corresponding Author: Kelly A. Schmidtke

Behavioural Science Group
Warwick Business School
The University of Warwick
Coventry, CV4 7AL, UK

Kelly.Schmidtke@wbs.ac.uk

Phone number: 07758933026

Alan J. Poots, NIHR CLAHRC Northwest London, Department of Medicine, Imperial College London, London, United Kingdom

Juan Carpio, Warwick Business School, The University of Warwick, Coventry, United Kingdom

Ivo Vlaev, Warwick Business School, The University of Warwick, Coventry, United Kingdom

Ngianga-Bakwin Kandala, Warwick Medical School, The University of Warwick, Coventry, United Kingdom

Richard J. Lilford, Warwick Medical School, The University of Warwick, Coventry, United Kingdom

ABSTRACT

Objectives: Hospital board members are asked to consider large amounts of quality and safety data with a duty to act on signals of poor performance. However, in order to do so it is necessary to distinguish signals from noise (chance). This article investigates whether data in English NHS acute care hospital board papers are presented in a way that helps board members consider the role of chance in their decisions.

Methods: Thirty English NHS trusts were selected at random and their board papers retrieved. Charts depicting quality and safety were identified. Categorical discriminations were then performed to document the methods used to present quality and safety data in board papers, with particular attention given to whether and how the charts depicted the role of chance, i.e., by including control lines or error bars.

Results: Thirty board papers were sampled, containing a total of 1488 charts. Only 88 (6%) of these charts depicted the role of chance, and only 17 of the 30 board papers included any charts depicting the role of chance. Of the 88 charts that attempted to represent the role of chance; 16 included error bars and 72 included control lines. Only six (8%) of the 72 control charts indicated where the control lines had been set (e.g., 2 vs 3 SD's).

Conclusions: Hospital board members are expected to consider large amounts of information. Control charts can help board members distinguish signals from noise, but often boards are not using them. We discuss demand- and supply-side barriers that could be overcome to increase use of control charts in healthcare.

Considering Chance in Quality and Safety Performance Measures:
An analysis of performance reports by boards in English NHS trusts

Introduction

Hospitals collect large amounts of data related to quality and safety. This information is presented to hospital board members who have a duty to scrutinize the data to help identify problems with care. However deriving inferences from data is not straightforward. A key issue concerns the role of chance, i.e., random variation. There is a need to distinguish a signal (sometimes called special-cause) from noise (common-cause) variation. Therefore, it is sometime difficult to distinguish signals from noise purely by visual inspection.[1]

This article is concerned with the presentation of data in such a way as to help board members make this distinction by identifying the role of chance.[2] First we document how quantitative data is presented in NHS board papers and then discuss potential barriers to representing the play of chance in charts and how they may be overcome.

NHS Hospital Boards

Whilst accountability for hospital safety and quality lies with the whole board, many boards establish special committees dedicated to such purposes which may have access to more information than is provided to the whole board. The board is supported by an elected council of patient, staff, and local resident Governors, all reporting up through the NHS infrastructure (Clinical Commissioning Groups, Public Health England, and the Department of Health) to the Secretary of State for Health, with monitoring and regulation provided by other agencies.[3]

Why focus on charts?

Data relating to quality and safety can be presented in tables or charts. While tables are an excellent presentation method to help decision makers identify past, unique data (e.g., what was the infection rate in July?), charts better portray patterns in data (e.g., is the infection rate increasing?).[4] As quality improvement relies on recognizing patterns in data, we concentrate on charts. The following section provides a classification of chart presentations.

Classification of chart presentation methods

Line and bar charts.

Line and bar charts are the most commonly chosen presentation methods.[5] Line charts better highlight trends across time and bar charts differences between discrete groups (e.g., patients, staff, hospitals).[6] More complicated charts combine information across time and between-groups. The interpretation of information in line and bar charts may be facilitated by including reference indicators, as we now describe.

Reference indicators that do not depict the role of chance.

Reference indicators are any features of a chart, that helps the user interpret the data. Reference indicators may indicate a standard that is external to the data, e.g., a regulator may require that 95% of patients attending an Accident and Emergency department are seen within four hours of arrival. Reference indicators of this type facilitate identification of data that exceed pre-set thresholds.[7] Examples of such charts are in Figure 1A and 1B. Reference indicators that indicate trends (e.g., lines of best fit) reveal patterns internal to the structure of the data. Examples of such charts are in Figure 2A and 2B. Neither of these reference indicators depicts the role of chance.

Reference indicators depicting the role of chance.

There are at least two commonly used types of reference indicators that depict the role of chance graphically: control charts and error bars. Control charts are a presentation method that includes reference indicators that make the role of chance explicit. They were originally developed for use in the manufacturing industry. Their use has since expanded to healthcare.[8] Control charts contain at least three reference indicators: a center line to signify the central tendency of data collected from a working process, and control lines surrounding the center line to signify variation due to chance. The amount of variation for which control lines account is at the creator's discretion; typically they are placed two or three standard deviations from the center line.[9]

The idea is that data falling between the control lines are likely to be the result of chance (common-cause variation). Data falling outside the control lines are more likely to be signals (result of special-cause variation). [1] Control lines act as thresholds based on statistical calculations to help target further investigations efficiently.[10, 11]

The horizontal axis for charts making comparisons between groups can be arranged such that data representing the group with the smallest sample-size appears first followed by data with increasingly larger sample-sizes. This rearrangement causes the control lines to take on a funnel appearance, termed a "funnel chart."[12]

Different methods of presenting the same information on a chart found in a NHS board paper are shown in Figure 3. Figure 3A, copied directly from the board paper, is a time-series line chart showing readmission rates by month. In this chart, the peak readmission rate, December, stands out, and so may trigger a board member to call for an investigation. Figure 3B shows the same information remade as a control chart. The peak is still shown, but the addition of control lines contextualizes the peak readmission rate as falling within the play of chance at 3 standard deviations (SD) and the lowest datum in May becomes more apparent. In so doing a

[1] When considering time-series data, special cause variations are also indicated when data series follow a statistically aberrant pattern, such as five data points all ascending or descending. Using multiple sets of control lines can facilitate the identification of some such patterns. For additional information see Champ, et al.[9]

board member's desire to investigate the high point may wane and their attention to the low point may wax. An example of a chart in board papers that could be remade into a funnel charts (i.e., that shown in Figure 3B) cannot be shown in the current paper as the necessary information is missing, i.e., the sample-size from which the data arose. One may note that error bars and control lines both represent dispersion of data, but in different ways. A more complete discussion of the distinction between hypothesis testing and control charts can be found in the literature.[13]

Our aim is to survey the quality and safety charts presented in NHS acute care trusts' board papers. In the following section we describe the methods by which we obtained and analyzed the charts in NHS publically available board papers according to the described classification system.

METHODS

Of the 163 English acute care trusts in the NHS Choices' service directory, 30 trusts were selected at random.[14] Each Trust was assigned a number and then Excel's random number generator was used to generate 30 numbers without replacement. The Trusts for which the assigned numbers were generated were selected. No geographical constraints were applied, but by chance these trusts include all nine historic regions of England, and remain anonymous. After selecting the trust to be included, temporal constraints were applied to ensure the analysis encompassed an entire year (May 2013 - April 2014); each selected trust was randomly assigned a month without replacement so that every month was selected at least twice, but no more than three times.[2] One board paper from each selected trust was obtained through each trust's website.

[2] As some trusts do not meet every month, a month could be randomly selected that was not available. When this occurred, trusts' months were exchanged. For example, Trust 1's randomly selected month may have been December, but during that month there was no board meeting. In addition, Trust 2's randomly selected month may have been February. If Trust 1 had a February meeting and Trust 2 had a December meeting, than their selected

Categorical discriminations were performed to understand the contents of the charts in these board papers. The first discrimination noted the charts' broad content: quality and safety, financial, patient surveys, staff, and activity. These categories were informed by past literature on hospital performance measures.[15] Further discriminations analyzed only the charts containing quality and safety data, using the classification of presentation methods discussed in the introduction (summarized in Table 1). All discriminations were performed independently by the first author (KS) and a co-author blind to the purpose of this article (JC).

RESULTS

Categorical discriminations

The initial, inter-rater reliabilities were high (average Cohen's kappa = 0.94) across the different discriminations and the majority of disagreements were settled after discussion between KS and JC (average Cohen's kappa = 0.99). The remaining 11 disagreements were arbitrated by a third author (RL). More detailed information appears in Table 1.

Table 1. Inter-rater reliabilities

Discrimination	Number of charts	Number of Categories	Kappa	95% confidence interval
1. Broad Content	1488	6*		
Initial			0.90 (p<0.01)	0.88 - 0.92
After discussion			1.00 (p<0.01)	0.99 -1.00
2. Quality and Safety Content	589	14**		
Initial			0.92 ($p < 0.01$)	0.89 - 0.94
After Discussion			1.00 ($p < 0.01$)	1.00 - 1.00

months were exchanged. Accordingly, Trust 1's month was now February and Trust 2's month was now December. If Trust 2 also did not have a December meeting, then Trust 3 would have been considered.

3. Methods used to Present Quality and Safety Data				
3a. Appearance	589	8***		
Initial			0.88 ($p < 0.01$)	0.85- 0.91
After Discussion			0.98 ($p < 0.01$)	0.96 - 0.99
3b. Comparison	589	3****		
Initial			0.91 ($p < 0.01$)	0.87 - 0.94
After Discussion			1.00 ($p < 0.01$)	1.00 - 1.00
3c. Reference indicators Depiction	285	2*****		
Initial			0.84 ($p < 0.01$)	0.76 - 0.91
After Discussion			1.00 ($p < 0.01$)	1.00 - 1.00

*quality and safety, financial, activity, patient surveys, staffing, and other
**waiting/delays, healthcare acquired infections, incidents reports, mortality, pressure ulcers, falls, length of stay, readmissions, venous thromboembolism, cleanliness, catheter, medication errors, others not consistently enough appearing to warrant a specific category and one for graphs that were placed in multiple categories
***line, bar, both line and bar, line with reference, bar with reference, both line and bar with reference, pie, other
****time-series, between-groups, both time-series and between-groups
*****reference indicators depicting either a standard or trend, or the role of chance

Broad content

In total, 1488 charts were located in the 30 board papers. The median board paper was 148 pages (range = 53-456) and contained 39.5 charts (range = 0-124). Quality and safety was the most frequent type (*Mdn* = 16, range = 0-54), followed by: financial information (*Mdn* = 7.5, 0-34), patient surveys (*Mdn* = 4.5, range = 0-38), staffing (*Mdn* = 4, range = 0-50), activity (*Mdn* = 2, range = 0-15) and others (*Mdn* = 0, range = 0-27). This article will now focus on those charts presenting quality and safety information.

Quality and safety contents

In total, 589 quality and safety charts were located across the 30 board papers. The median board paper contained 16 charts of this type, but with a wide range of 0 to 54. The types of quality and safety issues depicted, from most to least common were: waiting/delays (n = 112), incident reports[3] (*n* = 100), healthcare acquired infections (*n* = 99), and mortality (*n* = 85). Categories included less often, from most to least were: pressure ulcers (*n* = 30), falls (*n* = 27), length of stay (*n* = 19), venous thromboembolism prophylaxis (*n* = 15), readmissions (*n* = 14), cleanliness (*n* = 13), medication errors (*n* = 11) and information related to the management of catheters, urinary or vascular (*n* = 8), see Table 2. The results now presented relate to the 589 charts related to quality and safety.

Table 2. Quality and Safety Contents

Quality and Safety Content	Total Number*	Mean	Median	Range
Waiting-times	112	3.73	2	0- 16
Incident reports	100	3.33	2	0-16
Health care acquired infections	99	3.30	2	0-24
Mortality	85	2.83	2	0-18
Pressure Ulcers	30	1.00	0	0-5
Falls	27	0.90	0	0-5
Length of Stay	19	0.63	0	0-6
Venous thromboembolism	15	0.50	0	0-3
Readmissions	14	0.47	0	0-4
Cleanliness	13	0.43	0	0-5
Medication Error	11	0.37	0	0-4
Catheters (urinary, vascular)	8	0.27	0	0-4
Other[4]	89	2.97	1.5	0-11

*The numbers in this column will not add up to the total number of charts because eleven charts were placed in multiple categories

Classification of presentation methods for quality and safety charts

The 589 quality and safety charts will be classified in two different ways: first using the total number of charts as the denominator (e.g. 88 charts contained reference indicators that

[3] Incidents reports includes any graph which was an amalgamation of specific instances without specifically stating the type of incidents included, such as harm free days and serious incidents requiring investigation (SIRI's).
[4] Items that are displayed on less than eight of the charts, e.g., looked after children assessments.

depict the role of chance) and second using the median number of charts appearing in the 30 board papers as the denominator (e.g. the median board paper contained 1 chart depicting the role of chance), see Table 3. Figure 4 shows how the 589 total charts split into those including or not including a reference indicator, whether those indicators represented the role of chance and how they did so.

Table 3. Chart Presentation Methods

Chart Presentation Methods	Total Number	Median Number	Range Number
Line	347	8	0 – 48
Bar	158	4	0 – 21
Line and Bar	33	0	0 – 8
Other	51	0	0 – 13
Total	589		
Across Time	413	9	0 – 49
Between-Groups	112	1	0 – 18
Both Across and Between	64	1	0 – 12
Total	589		
Reference Indicators			
No	304	7	0 – 35
Yes	285	7	0 – 39
Total	589		
Reference Indicators depicting the Role of Chance			
No	197	4.5	0 – 37
Yes	88	1	0 – 16
Total	285		
Methods of Depicting Chance			
Error Bars	16	0	0 – 2
Control Lines	72	0	0 – 16
Total	88		

Line and bar.

Of the 589 charts dealing with quality/safety, over half were line charts (*n* = 347 [58.9%] *Mdn* = 8) and approximately a quarter were bar charts (*n* = 158 [26.8%] *Mdn* = 4). Charts including both lines and bars or other formats, e.g. pie charts, were much less common (*n's* = 33 and 51 respectively, [5.6% and 8.7% respectively] *Mdn's* = 1).

Performance across time and between-groups.

Of the 589 charts, most displayed comparisons across time (*n* = 413 [70.1%] *Mdn* = 9), followed by charts presenting comparisons both across time and between-groups (*n* = 112 [19.0%] *Mdn* = 1) and those comparing groups, e.g., wards or hospitals, at a given time (*n* = 64 [10.9%] *Mdn* = 1).

Reference indicators not depicting the role of chance.

There were 285 charts that included reference indicators. Of these 285 charts, 197 (69.1%, *Mdn* = 4.5) did not depict the role of chance. Of these 197 charts, 137 (69.5%, *Mdn* = 2) depicted an externally imposed standard and 38 (19.3% *Mdn* < 1) depicted a trend. An even smaller number of charts (22) displayed both standards and trends (11.2%, *Mdn* < 1).

Reference indicators depicting the role of chance.

Of the 285 charts that included reference indicators, only 88 highlighted the role of chance (*n* = 88 [30.9%] *Mdn* = 1). Of the 88 charts depicting the role of chance, 16 included error bars (18.2%, *Mdn* < 1) and 72 included control lines (81.8%, *Mdn* < 1).

Of the 30 board papers, only 17 (56.7%) board papers displayed any charts depicting the role of chance. Nine board papers included at least 1 chart with error bars and 14 included at least 1 control chart. Thus over half of the board papers did not contain any control charts.

Of the 72 control charts, 40 (55.6%, *Mdn* < 1) featured time-series and 32 (44.4%, *Mdn* < 1) between-groups comparisons. Only six of the control charts specified the control limits (e.g., 2 vs 3 SD's). Certain types of quality/safety indicators were more likely to be featured as control charts. Of the 40 time-series control charts, the most frequently occurring contents in order, from most to least, include: safety incidents (*n* = 11), mortality (*n* = 11), infection (*n* = 7), waiting (*n* = 4), pressure ulcers (*n* = 2), length of stay (*n* = 2) medication errors (*n* = 1), falls (*n* = 1), and the number of times patients were moved (*n* = 1). Of the 32 between-groups control charts, 16 charts, all from one board paper, used straight lines to compare infection or infection rates between hospitals. The remaining 16 between-groups control charts were all funnel charts. The contents of these charts included, in order from most to least: mortality (*n* = 11), incidents (*n* = 2), infection (*n* = 1), doctor to patient ratios (*n* = 1), and knee replacement outcomes (*n* = 1).

DISCUSSION

This article surveyed the quality and safety charts presented in 30 NHS acute care trusts' board papers. To our knowledge this is the first article describing how such information is presented to boards. The quality and safety charts available in these papers differed quantitatively and qualitatively. Although not the intended focus of this research, the wide variation in the number of charts is surprising (range 0 – 124), suggesting that there is little consensus on the quantity and types of information that should be presented to the board in graphical form. It is plausible that the number of charts, specifically depicting summative incidents reflect an open culture conducive to safety.

The role of chance was rarely depicted and where it was depicted, the charts were silent as to where the control lines had been set. This is suboptimal because without this information the role of chance is easily overlooked and common-cause variations can be misdiagnosed.[16]

Our results pertain to England and whether these findings apply elsewhere is unanswered. We hope that our study will provoke investigation of how charts are presented to decision makers (and whether or not the role of chance is depicted). Our focus on English hospitals may seem solipsist at first glance, since it is focused on but one issue in but one country. However, the results speak to broader issues of public engagement in science and statistics. The hospital board is one of many places where citizens and managers need to be numerate in order to take a view on issues that affect them.

While both control lines and error bars convey the role of chance, there are reasons to prefer control lines. Error bars allow performance measures to be compared, but this often cannot be accomplished by visual inspection alone since inferences require an accompanying statistical test. In contrast, control charts allow the reader to use visual inspection to derive statistical inferences without separate statements of statistical significance. Further, error bars are poorly understood by lay people and academics alike.[17] By comparison control charts are a "powerful means of communicating results to lay audiences or clinical personnel who are unfamiliar with statistical tests, probability values, effect sizes, and confidence intervals."[18]

Some readers may note that analyses of time-series line charts, i.e., run charts, can be guided by four rules, wherein an unusual pattern is designated by a: (1) shift, (2) trend, (3) run, or (4) astronomical point in the data series.[19]. Precise definitions are available for the first three rules and no reference indicators are needed (e.g., a trend is five or more consecutive points all going in the same direction). These rules are based on a false positive rate of 0.05 for normally distributed data. Encouragingly, board members could be taught to identify these patterns. The last rule however depends on chance variation which can often be difficult to discern without control lines.[20]

Our data do not explain why control charts are seldom used. We postulate that this could be due to an issue in demand (board members not requesting the data on control charts) and supply (staff are not able to supply the data in this format).

Demand

Barrier 1: Board members may not be aware of control charts

As the use of statistical process control is relatively new in healthcare, it may not have been part of many board members' formal education.[21] This is a barrier because if board members are not aware of control charts they do not have the capability to request them.

Recommendation 1:

Active education may overcome this barrier. These efforts could take the form of a brief tutorial at a board meeting and/or instructional annotations on control charts as they are added to board reports. An introduction to control charts should be offered to new board members as part of their induction, at least until control charts become commonplace in healthcare. More generally, citizens need to understand simple statistical information to make more informed decisions about their care.

Barrier 2: Board members may not feel comfortable in their ability to interpret control charts

While we might expect board members to have more experience and knowledge of typical charting techniques, they might be unfamiliar or uncomfortable with the interpretation of such control charts, particularly regarding control lines.

Recommendation 2:

This barrier may be overcome initially by providing text or annotations to the charts to highlight when data are likely special-cause data. However, we worry that such text might overshadow the most important information, i.e., the data. In time it would be preferable that board members were empowered to identify special-cause data themselves. This is also a capability issue which again may be addressed through training. Below we now briefly propose information these educational efforts should include.

Control lines, which are typically dashed lines, are often set three standard deviations above and below the center line. This placement ensures that there is a small chance that an investigation of a signal will be unjustified. However, more cautious board members may consider this suitable for industrial uses but too stringent for health care, preferring a two standard deviation control line.[22] Such a practice increases the chance of a false positive signal more than many realize. Up to 25% of data can be located beyond two standard deviations (Chebyshev's inequality).[23] A common-sense approach may be to include both 2 and 3 SDs control limits. Determining where the control lines are set on charts for different measures should reflect the cost of investigation and the cost of not investigating, in terms of money, quality, and safety. This is a question of judgment and cannot be resolved statistically and it will vary from one type of measure to another.

Supply

Barrier 1: Staff may not know how to create control charts

Assuming staff know what control charts are, a reason staff do not provide control charts may be that they do not have the practical tools to do so at their fingertips – a question of opportunity.

Recommendation 1:

Staff should be encouraged to use to computer software to help them create control charts. However while numerous software tools exist, many are likely to cause more frustration than aid because they are unfamiliar, expensive, and create files that can only be shared if others have the same program (e.g., SPSS, Minitab, Sigma 6). Using a more common program such as Excel, might be easier for an organization first exploring control charts. Staff familiar with Excel's functions can set up templates for other staff to use. Another option is to install an Excel add-in, either at a cost or using peer-reviewed freeware.[24]

Barrier 2: Staff are not confident they have a sufficient number of data points to construct a control chart.

Another reason staff may not provide control charts is that they do not think they have a sufficient amount of data points to plot on the chart.

Recommendation 2:

While recommendations vary, a desirable number of data points required to set up the control lines ranges from 10 to 35.[25] The number of data points available to hospital staff often fall outside this range. The number of data points considered within a control charts can often be increased or reduced, for instance, by looking at shorter/larger time intervals and carefully aggregating data (e.g., plotting data by week rather than by day). As the number of available data can only be increased by collecting it more frequently, we urge hospitals to use automatic tools to record data as frequently as possible. For rare events, special control chart techniques have been described by Woodall and Driscoll.[26]

Barrier 3 - Staff are not confident they are selecting the correct type of control chart

Staff members may fail to use a control chart because they are uncertain which control chart is best.

Recommendation 3 –

Perfection should not become an enemy of the good. Fears about selecting the wrong chart may be mitigated by realizing that underlying principles are similar across different types. There are seven basic control charts types: Xbar, XmR, XmS, C, U, P and NP. As a default, we recommend using a XmR chart, which has proven robust for most time-series data and is a good place to start.[27] Many more sophisticated varieties are available, e.g., CUSUM and EWMA, for those who are comfortable with the basic types of control charts.[28] The table in the appendix 1 *(Adapted from Steven Wachs Integral Concepts, Inc.)* may be used to help select an appropriate control chart for different dimensions of quality and safety. Other decision tools are available in the literature.[29, 30]

A more fundamental issue relates to the importance of stating where the control lines have been set. Few of the control charts we located in board papers indicated where the control lines had been set. This is a concern because if the board members do not know where the control lines are set (e.g., 2 or 3 sigma) they cannot assess the chance of making type 1 and 2 errors in their decisions.

Of course board members' decisions should not be solely influenced by information provided in charts. Rather such data should be contextualized in the other information available to the board. For example if pressures sores increase within a ward but not enough to breach the upper control line, the board may want to consider the nurse to patient ratio in that ward before ultimately deciding whether further action is warranted. Further research is indicated not just for how decision makers can make the best use of statistical information within a single data set but also on how information across multiple data sets can be synthesized to inform decisions.

Conclusion

In summary, NHS board papers in England contain many quality and safety charts. Unfortunately, few of these charts allow board members to appreciate the role of chance in the data. To our knowledge, this is the first report documenting the types of charts used to present data to hospital boards. While the control charts are increasingly being used to monitor health related variables around the world, we suspect that they are still underused in many countries (and look forward to seeing such comparisons).[31] The introduction of control charts into NHS board papers is a simple process that would greatly improve board members' ability to consider the role of chance in their decisions, and ultimately provide better management for patient care.

ACKNOWLEDGEMENTS, COMPETING INTEREST, FUNDING AND ALL OTHER REQUIRED STATEMENTS

-NAME REMOVED FOR PEER REVIEW- is supported by the Academic Health and Science Network (AHSN). The paper presents independent research and the views expressed are those of the author(s) and not necessarily those of the AHSN

This article presents independent research commissioned by the National Institute for Health Research (NIHR) under the Collaborations for Leadership in Applied Health Research and Care (CLAHRC) programme, North West London and West Midlands. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

REFERENCES

1 Peymane A, Rouse A, Muhammed, AM. Performance league tables: The NHS deserves better, BMJ 2002;12:95-8.

2 Marshall M, Romano P. Impact of reporting hospital performance. Qual Saf Health Care 2005;14:77–8.

3 NHS 2013. The NHS Infrastructure explained. Available online [accessed 2015-11-15] http://www.nhs.uk/NHSEngland/thenhs/about/Pages/nhsstructure.aspx

4 Vessey I. Cognitive fit: A theory-based analysis of the graphs vs tables literature. Decision Sciences 1991;22:219-40. doi: 10.1111/j.1540-5915.1991.tb00344.x

5 Speier C. The influence of information presentation formats on complex task decision-making performance. Int J Hum Comput Stud 2006;64:1115-31. doi: 10.1016/j.ijhcs.2006.06.007

6 Lipkus IM. Numeric, verbal, and visual formats of conveying health risks: Suggested best practices and future recommendations. Med Decis Making 2007;27:696–713. doi: 10.1177/0272989X07307271

7 Cleveland WS. The Elements of Graphing Data (revised edition). Summit, NJ: Hobart Press 1994.

8 Mohammed MA, Cheng KK, Rouse A, Marshall T. Bristol, Shipman, and clinical governance: Shewhart's forgotten lessons.Lancet. 2001 Feb 10;357(9254):463–467.

9 Shewhart WA. Statistical Method from the viewpoint of Quality Control. Graduate School of the Department of Agriculture, Washington, D.C. 1939

10 Deming WE. On probability as a basis for action, Am Stat 1975;29:146–52. doi: 10.1080/00031305.1975.10477402

11 Orme JG, Cox ME. Analyzing single-subject design data using statistical process control charts. Soc Work Res, 2001 25:115–27. doi: 10.1093/swr/25.2.115

12 Spiegelhalter DJ. Funnel plots for comparing institutional performance. Stat Med 2005;24:1185-202. doi: 10.1002/sim.1970

13 Woodall WH. Controversies and contradictions in statistical process control. Journal of Quality Technology 2000;32: 341-50.

14 NHS Choices Authorities and trusts: All acute trusts. [cited April 20, 2014]. Avalible from: www.nhs.uk/servicedirectories/pages/acutetrustlisting.aspx.

15 Nerenz DR, Neil N. Performance Measures for Health Care Systems, Commissioned Paper for the Center for Health Management Research, 2001 [cited August 27, 2012]. Available at: http://depts.washington.edu/chmr/docs/commissioned_papers/performancemeasures_nerenz_2001.doc.

16 Wheeler DJ. Avoiding Man-Made Chaos, SPC Press, Inc 1998.

17 Hildon Z, Allwood D, Black N. Making data more meaningful: Patients' views of the format. Int J Qual Health Care 2012;24:55-64. doi: 10.1016/j.pec.2012.02.006

18 Polit DF, Chaboyer W. Statistical process control in nursing research. Res Nurs Health 2012;35*:*82-93, p. 89. doi: 10.1002/nur.20467

19 Perla R, Provost L, Murray, S. The run chart: A simple analytical tool for learning from variation in healthcare processes. BMJ Quality & Safety 2011;20:46-51. doi: 10.1136/bmjqs.2009.037895.

20 Anhøj J, Olesen AV. Run charts revisited: A simulation study of run chart rules for detection of non-random variation in health care processes. PLoS ONE 2014;9(11): e113825. doi:10.1371/journal.pone.0113825

21 NHS Institute for Innovation and Improvement. Quality and service improvement tools: Statistical process control, 2013. [Accessed March 19, 2015].Available from: NHS Institute for Innovation and Improvement 2006-2013. http://www.institute.nhs.uk/quality_and_service_improvement_tools/quality_and_service_improvement_tools/statistical_process_control.html#sthash.0L0qtk9J.dpuf

22 Noyez L. Control charts, cusum techniques and funnel plots. A review of methods for monitoring performance in healthcare. Interact Cardiovasc Thorac Surg 2009;9:494–9. doi: 10.1510/icvts.2009.204768

23 Kvanli AH, Pavur RJ, Keeling KB. Concise Managerial Statistics cEngage Learning 2006:81–2. ISBN 9780324223880

24 Buttery SE. An excel add-in for statistical process control charts, Journal of Statistical Software 2009;30(13).

25 Koetsier A. Control charts in healthcare quality improvement. A systematic review on adherence to methodological criteria. Methods Inf Med 2012;51:189-98. doi: 10.3414/ME11-01-0055

26 Woodall, W. H., Driscoll, A. G. Some Recent Results on Monitoring the Rate of a Rare Event, Chapter in: Frontiers in Statistical Quality Control 11. Knoth, S. and Schmid, W. (Eds.), Springer, 2015; 15-27.

27 Poots AJ, Woodcock T. Statistical process control for data without inherent order, BMC Med Inform Decis Mak 2012;12(86). doi: 10.1186/1472-6947-12-86

28 Provost LP, Murray S. The Health Care Data Guide: Learning from Data for Improvement. San Francisco: Jossey-Bass 2011.

29 Amin SG. (2001). Control charts 101: A guide to health care applications. Qual Manag Health Care 2001;9:1-27.

30 Mohammed M, Worthington P, Woodall WH. Plotting basic control charts: tutorial notes for healthcare practitioners. Qual Saf Health Care 2008;17:137-45. doi: 10.1136/qshc.2004.012047

31 Thor J, Lundberg J, Ask J, Olsson J, Carli C, Härenstam KP, Brommels J. Application of statistical process control in healthcare improvement: systematic review. Qual Saf Health Care 2007;16:387-399 doi:10.1136/qshc.2006.022194

Appendix 1. Decision tool to select appropriate control-chart.

DATA TYPE	CHARTS	MONITORS	APPLICATIONS
Variable	X-Bar and S	Process average and standard deviation	High volume, single characteristic Sample size 2 or larger
Variable	X-Bar and R	Process average and range	High volume, single characteristic Sample size between 2 and 6
Variable	X and MR	Process average and moving range	Sensitivity not required Sampling is costly Long cycle time (Note: Normality of data must be considered.)
Variable	Deviation from Nominal	Process average and range (or standard deviation)	Short production runs (multiple parts) All parts have similar standard deviation
Variable	Standardized X-Bar and R Standardized X-Bar and S	Process average and range Process average and standard deviation	Short production runs (multiple parts) Part standard deviations differ
Variable	X-Bar, Rb, d	Process average, range between and difference between extreme locations	Multiple locations within subgroup Location averages are statistically different
Variable	X-Bar, Rb, Rw X-Bar, Rb, S	Process average, range (or standard deviation) within and range between subgroup	Multiple locations within subgroup Variation within and between subgroups different Location averages are not statistically different
Variable	CUSUM	Cumulative deviations from mean	Charts for individuals when X and MR are not sensitive enough
Variable	EWMA	Weighted moving average	Charts for individuals when X and MR are not sensitive enough
--	--	--	--
Attribute	Np	Number of Defectives	Pass/Fail Data Constant Sample Size $n > 3/p$

Attribute	P	Proportion Defective	Pass/Fail Data Constant or Variable Sample Size n > 3/p
Attribute	Standardized p	Standardized Proportion Defective	Pass/Fail Data Variable Sample Size n > 3/p Can be used for short production runs
Attribute	C	Number of Defects	Multiple types of defects on unit Constant sample size n such that c > 7
Attribute	U	Number of Defects per unit	Multiple types of defects on unit Constant or variable sample size n such that c > 7
Attribute	Standardized u	Standardized Number of Defects per unit	Multiple types of defects on unit Variable sample size n such that c > 7 Can be used for short production runs

Figure 1. Charts taken from actual board papers with reference indicators depicting a standard that is external to the data (altered to be greyscale). Cell A contains a line chart, showing that trusts 4-hour performance target in the Accident and Emergency department, the current year, the previous year and the national target of 95%. Cell B contains a bar chart, showing different Trusts' cumulative C-Diff performances and their respective targets.

A. B.

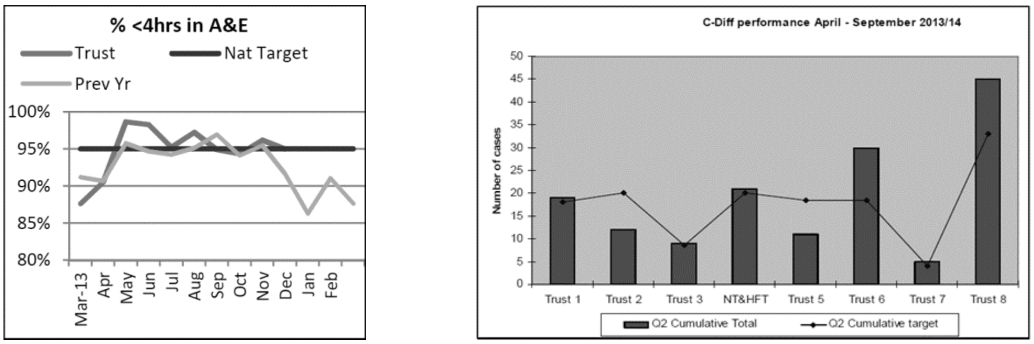
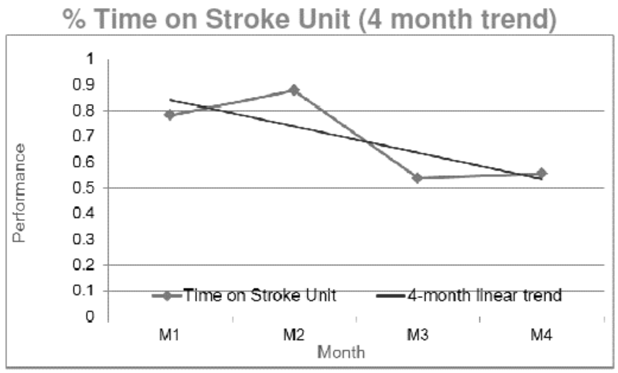


Figure 2. Charts taken from actual board papers with reference indicators depicting data trends (altered to be greyscale). Cell A contains a line chart; the unmarked line shows a 4-month linear trend. Cell B contains a bar chart; the line shows the average performance for the previous 12 months, including the 4 months displayed

A.



B.

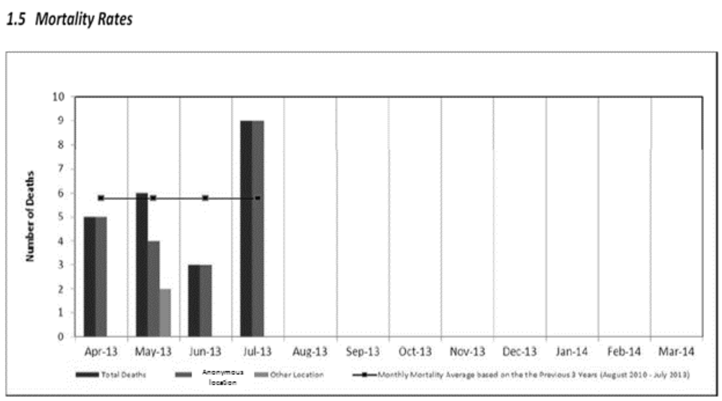


Figure 3. Chart in Cell A taken from actual board papers (altered to be greyscale). Cell A contains a time-series line chart showing patient readmissions across months. Cell B contains the same data remade into a control chart at three SDs

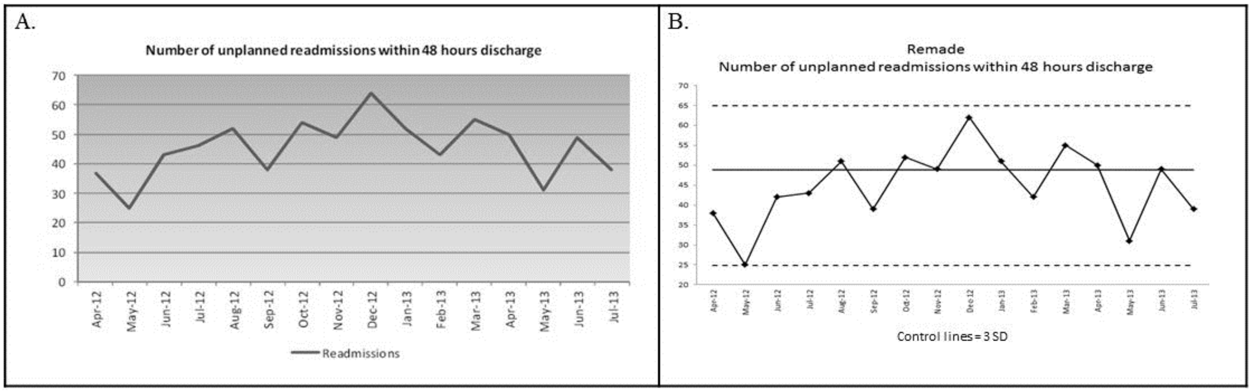


Figure 4. How the 589 total charts split according to whether they included reference indicators, whether the reference indicators highlight the role of chance, and how the role of chance was displayed