

BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email editorial.bmjopen@bmj.com

BMJ Open

Identifying positively deviant elderly medical wards using routinely collected NHS Safety Thermometer data: an observational study.

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2017-020219
Article Type:	Research
Date Submitted by the Author:	24-Oct-2017
Complete List of Authors:	Baxter, Ruth; Bradford Institute for Health Research, Yorkshire Quality and Safety Research Group; University of Leeds, School of Psychology Taylor, Natalie; Cancer Council, Cancer Research Division Kellar, Ian; University of Leeds, School of Psychology Pye, Victoria; Macquarie University, Australian Institute of Health Innovation Mohammed, Mohammed ; faculty of Health care studies, University of Bradford, Reserach Lawton, Rebecca; Bradford Institute for Health Research, Yorkshire Quality and Safety Research Group; University of Leeds, School of Psychology
Primary Subject Heading:	Health services research
Secondary Subject Heading:	Research methods
Keywords:	Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, STATISTICS & RESEARCH METHODS, positive deviance, quality improvement, implementation science

SCHOLARONE™
Manuscripts

1
2
3
4 **Identifying positively deviant elderly medical wards using routinely collected**
5 **NHS Safety Thermometer data: an observational study.**
6
7
8
9
10
11
12

13 **Dr. Ruth Baxter^{1,2*}, Dr. Natalie Taylor³, Dr Ian Kellar², Ms. Victoria Pye⁴, Prof. Mohammed A**
14 **Mohammed^{1,5}, Prof Rebecca Lawton^{1,2}.**
15

16 1 - Yorkshire Quality and Safety Research group, Bradford Institute for Health Research, Bradford, UK
17

18 2 – School of Psychology, University of Leeds, Leeds, UK
19

20 3 – Cancer Research Division, Cancer Council, Sydney, Australia
21

22 4 – Australian Institute of Health Innovation, Macquarie University, NSW, Australia
23

24 5 – Faculty of Health Studies, University of Bradford, Bradford, UK
25

26 *corresponding author: ruth.baxter@bthft.nhs.uk; 01274 383413; Yorkshire Quality and Safety
27 Research group, Bradford Institute for Health Research, Bradford Royal infirmary, Duckworth Ln,
28 Bradford, BD9 6RJ.
29
30
31
32
33
34
35
36
37

38 **Word Count: 3766**
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

ABSTRACT (300/300 words)

Objective: The positive deviance approach seeks to identify and learn from exceptional performers. Although a framework exists to apply positive deviance within healthcare organisations, there is limited guidance to support its implementation. The approach has also rarely explored exceptional performance on broad outcomes, been implemented at ward level, or applied within the UK. This study develops and critically appraises a pragmatic method for identifying positively deviant wards using a routinely collected, broad measure of patient safety.

Design: A two-phased observational study was conducted. During phase 1, cross-sectional and temporal analyses of NHS Safety Thermometer data were conducted to identify a discrete group of positively deviant wards that consistently demonstrated exceptional levels of safety. A group of matched comparison wards with above average performances were also identified. During phase 2, multidisciplinary staff and patients on the positively deviant and comparison wards completed surveys to explore whether their perceptions of safety supported the identification of positively deviant wards.

Setting: 34 elderly medical wards within a northern region of England, UK.

Participants: Multidisciplinary staff (n=161) and patients (n=188) clustered within nine positively deviant and comparison wards.

Results: Phase 1: A combination of analyses identified five positively deviant wards that performed best in the region, outperformed their organisation, and performed consistently well over 12 months. Five above average, matched comparator wards were also identified. Phase 2: Staff and patient perceptions of safety generally supported the identification of positively deviant wards using Safety Thermometer data, although patient perceptions of safety were less concordant with the routinely collected data.

Conclusions: This study tentatively supports a pragmatic method of using routinely collected data to identify positively deviant elderly medical wards; however, it also highlights the various challenges that are faced when conducting the first stage of the positive deviance approach.

Registration: UK Clinical Research Network Portfolio (reference-18050).

Key words

Positive deviance, implementation science, quality improvement, NHS Safety Thermometer, elderly care, patient safety

Strengths and limitations of this study

- There is limited guidance to support the identification of positive deviants in healthcare settings. This study develops a method for identifying positive deviants using routinely collected data.
- A combination of four different analyses (including performance rankings, comparisons with organisational level performances, and Statistical Process Control methods) were conducted to provide a pragmatic yet robust method for identifying a discrete group of positively deviant wards that performed exceptionally well on a broad outcome of safety.
- Staff and patient perceptions of safety were measured using validated surveys to explore whether they supported the identification of positively deviant wards using routinely collected data.
- Due to the small sample size (n=9 wards) it was not possible to statistically assess whether staff and patient perceptions of safety supported the identification of positive deviants using routinely collected Safety Thermometer data.
- The study was conducted on elderly medical wards and so further research is required to explore whether the methods can be generalised to identify positive deviants other healthcare settings.

BACKGROUND

Despite extensive efforts to improve, patient safety continues to be a pervasive problem across the globe.(1, 2) Traditionally, these efforts have focused on past errors and harm, but there are increasing calls to also explore how 'safe' patient care is delivered.(3, 4) Positive deviance provides an asset based approach to improving the quality and safety of healthcare.(5, 6) The approach seeks to identify and learn from those who demonstrate exceptional performance on an outcome of interest.(6) It assumes that solutions to problems already exist within communities and that positive deviants (individuals, teams, or organisations) identify these solutions and succeed despite facing similar constraints as others.

Bradley et al.(7) have proposed a four stage framework to implement the positive deviance approach within healthcare organisations. Positive deviants, who display exceptionally high performance, are identified using routinely collected data (stage 1). Hypotheses about how they succeed are generated using qualitative methods (stage 2). These hypotheses are quantitatively tested in larger, more representative samples (stage 3), and then disseminated to others with the help of key stakeholders (stage 4). Despite the increasing popularity of the positive deviance approach,(8) there is little evidence or practical guidance to support its application within healthcare organisations.(9, 10)

Identifying positive deviants who demonstrate exceptional performance

Identifying positive deviants may be the most crucial stage of the Bradley et al.(7) framework as subsequent stages hinge on its perceived legitimacy. Misidentification could lead to the generation of hypotheses that do not capture the factors that facilitate exceptional performance. Bradley et al.(7) suggest that positive deviants should be identified by ranking routinely collected data, and previous healthcare applications have, for example, identified three of the top 10 clinics with the best anticoagulation control(11) or the top quintile of primary care medical homes with the most improved diabetes outcomes.(12) However, performance rankings can differ depending on the rating systems that are used creating confusion and contradiction about who demonstrates high and low performances.(13-15) Positive deviants are also supposed to demonstrate *exceptional* rather than just *good* performances on the outcome of interest,(7) but rankings or league tables simply appraise performances along a continuum without differentiating a distinct groups of 'outliers' or positive deviants from the rest of a population.

1
2
3 Previous healthcare applications of the approach have predominantly identified positively
4 deviant organisations (e.g. hospitals) or individuals.(9) This is despite greater amounts of variation
5 existing at the level of a hospital ward or unit(16-18) and the majority of frontline care being
6 delivered by the multidisciplinary teams that work within these clinical microsystems.(19) Previous
7 applications have also typically focused on quite specific processes or outcomes of care such as hand
8 hygiene compliance and the incidence of healthcare associated infections.(9) Although it is relevant
9 to explore positive deviance in this way, factors that contribute to safety often operate across
10 various levels of the system and affect multiple outcomes.(20, 21) If ward teams succeed on broad
11 outcomes of care there are likely to be some underlying, latent factors that facilitate their success.
12 Understanding these factors and spreading the associated strategies may generate more far-
13 reaching improvements in quality and safety. However, it remains unknown whether positively
14 deviant wards or units can be identified accurately using a routinely collected, broad and
15 multidimensional measure of patient safety.

16
17
18
19
20
21
22
23
24 Based on these gaps in the literature this study was conducted in two phases to develop and
25 critically appraise a robust yet pragmatic method for identifying positive deviants at ward level using
26 a routinely collected, broad outcome of safety. The shortcomings of using routinely collected data,
27 such as publication lags, coding differences, and data gaming, are well documented.(22-24)
28 However, if positive deviance is to become a useful improvement approach, its methods must be
29 pragmatic and accessible for healthcare organisations, networks, and frontline improvers to use.
30 Routinely collected and publicly available data are therefore required, especially when applying the
31 approach across several different wards, units, or organisations. In preparation for this study various
32 routinely collected measures of safety within the NHS were identified (e.g. Hospital Episode
33 Statistics,(25) and the NHS Staff Survey(26)), but the NHS Safety Thermometer (ST) provides the only
34 routinely collected, broad measure of safety that is publicly available at ward level.

35
36
37
38
39
40
41
42
43 Each month the NHS ST measures four commonly occurring patient harms – falls, pressure ulcers,
44 venous-thromboembolism, and urinary tract infections – and uses this data to report on the
45 proportion of ‘harm-free care’ that is delivered at ward, speciality and organisational levels.(27)
46 During phase 1 of this study, a rigorous and robust analysis (compared to simply ranking the data)
47 was applied to the ST’s harm-free care data to identify a distinct group of positively deviant wards
48 that demonstrated exceptional levels of safety over a 12 month period. To try and identify positively
49 deviant ward teams that were delivering exceptionally safe patient care under challenging
50 circumstances, elderly medical wards were sampled because older patients are particularly
51 vulnerable to safety incidents and harms such as falls and pressure ulcers.(28, 29)
52
53
54
55
56
57

1
2
3 A key challenge to improving the quality and safety of healthcare is convincing people to adopt a
4 chosen solution.(30) The positive deviance approach seeks to identify solutions from within that are
5 acceptable, feasible, and sustainable in order to disseminate them to others.(6) However, if staff
6 and/or patients do not perceive positive deviants to be performing exceptionally well, they may not
7 engage with the positively deviant strategies that are disseminated during stage 4 of the Bradley et
8 al.(7) framework. As this study was one of the first applications of the approach to identify positive
9 deviants at ward level using a broad measure of safety, phase 2 of this study explored whether staff
10 and patient perceptions of safety were similar to, or at odds with, the routinely collected data.
11
12
13
14
15

16 Although this particular study focuses on stage 1 of the Bradley et al.(7) framework, it also
17 contributes to a wider application of the positive deviance approach(31) which seeks to generate
18 hypotheses about how positively deviant ward teams deliver safe patient care. The qualitative
19 findings from this wider application (stage 2 of the framework) will be published elsewhere. To the
20 authors' knowledge, this study is one of the first applications of positive deviance within the English
21 National Health Service (NHS).
22
23
24
25
26
27
28
29

30 **PHASE 1 METHOD**

31 **Sample**

32
33
34
35
36 Elderly medical wards (n=36) were identified from within 13 acute NHS Trusts (healthcare
37 organisations) in a northern region of England, UK. This represented 10% of all acute NHS Trusts in
38 England. Elderly medical wards were defined as those providing 24 hour, acute, medical care for
39 elderly patients (>65 years); with dedicated multidisciplinary teams; and patient stays typically
40 exceeding 48 hours. Speciality wards (e.g. stroke, rehabilitation, and assessment units) were
41 excluded to maximise homogeneity within the sample.
42
43
44
45

46 **Data extraction**

47
48
49 For all wards the ST data were extracted from a publicly accessible website(32) for a 12 month
50 period – August 2013 to July 2014. Data were extracted for the ST's 'harm-free care' measure and all
51 of the individual ST harms at two different levels: ward level for all patients; and Trust level for acute
52 patients over 70 years.(31) Double blinding during both study phases ensured that researchers (RB,
53
54
55
56
57
58
59
60

1
2
3 RL, NT, and IK), staff and patients were not aware of how wards compared to each other on their ST
4 performances.
5

6 7 **Analysis**

8
9
10 Pragmatic cross-sectional and temporal analyses were conducted to identify a distinct group of
11 positively deviant elderly medical wards that displayed exceptionally high performance on the ST
12 harm-free care measure. Initially, in line with guidance(7) and previous applications,(9) wards were
13 ranked on their average harm-free care performances and then three further analyses were
14 conducted. First, to ensure that positive deviants' success was not simply a function of
15 organisational performance, a scatterplot compared wards with their respective NHS Trust level
16 data. Second, as small sample sizes increase the likelihood of variability being attributable to
17 chance,(33) a funnel plot compared ward level performances against their average sample sizes.
18 Third, performances were assessed over the 12 month period using run charts.(34) Monthly ward
19 level data were compared with the monthly regional average, and run charts were visually assessed
20 to identify consistent outperformers. Statistical Process Control methods such as run charts and
21 funnel plots are increasingly being promoted to assess variation within healthcare.(35, 36)
22
23
24
25
26
27
28
29

30 **Identification of wards**

31
32 Analyses were compared to identify a distinct group of positively deviant elderly medical wards.
33 Individual harms data were also assessed to ensure that positive deviants performed well across all
34 measures in the ST's harm-free care composite. In preparation for phase 2 of this study, comparison
35 wards with slightly above average ST performances were identified. As is the case for many sources
36 of publicly available data, it was not possible to conduct any case-mix adjustments on the ST data.
37 Consequently, comparator wards were matched to positive deviants on three key variables to
38 increase homogeneity within the sample: patient gender – mixed, female, or male; NHS Trust type -
39 teaching and/or foundation trusts; and a routinely collected measure of deprivation.(37) To ensure
40 that positive deviants did not simply care for younger, and thus, comparatively more healthy
41 patients,(28, 29) administrative average patient age data were analysed post-hoc.
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

PHASE 1 RESULTS

Data were analysed for 34 elderly medical wards clustered within 13 NHS Trusts. Two wards with over 50% missing data were excluded. Average harm-free care performances ranged from 70.56% to 92.68% (supplementary file 1 presents all ward rankings). Wards 7, 4, 17, 31, 36, and 29 were the only wards to outperform their respective Trusts on the ST harm-free care data (Figure 1). The first five of these also ranked the highest in the sample while ward 29 ranked eighth. Visual assessment of the run charts (Figure 2) indicated that wards 7, 17, 31, 36, 4, and 15 consistently outperformed the regional average over 12 months, with greater certainty held for those wards listed first. Although none of the wards exceeded the funnel plot's three standard error control limits, wards 7, 4, 17, 31, and 36 exceeded them at two standard errors (Figure 3).

Using a combination of these four analyses, five wards (7, 4, 17, 31, and 36) were identified to form a distinct group of positive deviants. These wards demonstrated the best performances (rankings); out-performed their respective NHS Trusts (scatterplot); consistently outperformed over 12 months (run charts); and their performance variation was attributable to more than chance alone (funnel plot). They also performed well – around/above average – for each individual ST harm (supplementary file 1). Wards 29 and 15, which were identified through the scatterplot and run charts respectively, did not exceed the funnel plot control limits and so were not deemed to be positively deviant.

In total, five matched comparison wards with slightly above average harm-free care performances were also identified. Independent samples t-tests indicated that positively deviant wards ($M=91.33, SD=0.92$) significantly differed from comparators ($M=87.46, SD=1.31$, $t(8)=5.42, p=0.001$) and all other wards in the region ($M=83.85, SD=4.57$, $t(32)=3.61, p=0.01$) for average ST harm-free care performance. Supplementary file 2 presents the key characteristics of the positively deviant and comparison wards. Although it was not sufficiently powered due to the small sample size ($n=9$), a post hoc analysis of administrative data for average patient age suggests that positively deviant wards ($M=85.1, SD=2.11$) did not care for younger, and thus more healthy patients, than the comparison wards ($M=84.92, SD=1.42$, $t(7)=0.15, p=0.88$).

PHASE 2 METHODS

Participants and recruitment

1
2
3 The positively deviant and comparison wards identified during phase 1 were invited to participate
4 in phase 2 of the study. One positively deviant ward was unable to take part. Up to 20 patients and a
5 minimum of 50% of the multidisciplinary ward team were recruited opportunistically to complete
6 surveys assessing their perceptions of safety. Eligible patients were 65 years or older, were deemed
7 to have capacity, and were physically well enough. Double blinding was retained.
8
9
10
11
12
13

14 **Data collection tools**

15 Patient survey

16
17 Patients completed the Patient Measure of Safety (PMOS) which gathers feedback from
18 hospitalised patients about factors that contribute to safety. A total of 44 items are scored on 5-
19 point Likert scales ranging from strongly agree to strongly disagree. The PMOS has been validated, is
20 reliable, and considered acceptable to patients.(38, 39) Patients also completed the NHS Friends and
21 Family Test (FFT) – a single item measure of patient experience used in the UK,(40) and three items
22 that had previously been a part of the NHS Commissioning for Quality and Innovation (CQUIN)
23 payment framework.(41)
24
25
26
27
28
29
30

31 Staff survey

32
33 Multidisciplinary staff completed the Patient Safety Grade (PSG) which asks them to grade their
34 ward on overall safety using a 5-point Likert scale (excellent to failing). The PSG is one of four
35 outcomes within the validated Hospital Survey on Patient Safety Culture (HSOPSC) where non-
36 required outcomes can be removed.(42-44) This single item was used to maximise response
37 rates.(45) The patient and staff surveys were published with the study protocol.(31)
38
39
40
41
42
43
44
45
46

47 **Procedure**

48
49 Eligible patients were identified by clinical members of the ward team. The patient survey was
50 completed electronically and, due to high levels of frailty, researchers supported most patients to
51 read questions and record their responses. Paper copies of the staff surveys were distributed to
52 multidisciplinary staff at convenient times (e.g. clinical handovers) for independent completion.
53 Participation was incentivised by a prize draw (£20 gift voucher per ward).
54
55
56
57

Analyses

Blinding was removed, data were aggregated to ward level, and average scores were calculated for all measures. To assess whether staff and patient perceptions of safety supported the identification of positive deviants, wards were ranked and z-scores were compared in a scatterplot. The small sample size (n=9 wards) and dichotomised performance groups prevented statistical analysis.

PHASE 2 RESULTS

Data were collected from 188 patients and 161 multidisciplinary staff, clustered within nine participating elderly medical wards (supplementary file 3 reports all recruitment data). On average patients were 84.53 years old (SD=5.45), and staff were predominantly nurses or support workers. Table 1 reports the ward level descriptives for all of the measures. Positively deviant wards performed better than comparators across all four measures, although differences between the groups were small.

Table 1 Ward level descriptive statistics for all staff and patient survey measures

		ST Harm-free care (Phase 1) Mean %	Patient Measure of Safety Mean (SD) ^a	Friends and Family test Mean (SD) ^a	CQUIN Mean (SD) ^b	Patient Safety Grade Mean (SD) ^a
Positively deviant wards	Ward 1	90.14	4.33 (.45)	4.71 (.56)	2.48 (.45)	4.29 (.56)
	Ward 3	92.68	4.21 (.34)	4.55 (.67)	2.58 (.47)	4.21 (.70)
	Ward 6	91.48	3.94 (.37)	4.14 (1.15)	2.45 (.32)	4.09 (.54)
	Ward 10	90.97	4.52 (.26)	4.65 (.49)	2.53 (.48)	4.13 (.78)
Comparison wards	Ward 2	88.48	4.11 (.53)	4.26 (.75)	2.25 (.50)	3.50 (1.15)
	Ward 4	87.72	4.09 (.39)	4.26 (1.00)	2.43 (.45)	4.07 (.48)
	Ward 5	85.17	3.96 (.39)	4.15 (1.23)	2.18 (.58)	4.05 (.52)
	Ward 8	87.90	4.51 (.27)	4.75 (.44)	2.48 (.33)	3.69 (.79)
	Ward 9	88.01	4.30 (.36)	4.46 (.88)	2.50 (.36)	3.29 (1.16)
Average	PD group	91.33	4.24 (.41)	4.51 (.78)	2.51 (.43)	4.18 (.67)
	Comp group	87.46	4.20 (.43)	4.38 (.92)	2.38 (.46)	3.71 (.91)

Abbreviations: ST = Safety Thermometer; CQUIN = Commissioning for Quality and Innovation; PD = Positive Deviant; Comp = Comparison

^a Measured on a 0-5 Likert scale; ^b Measured on a 0-3 Likert scale. Higher scores represent safer perceptions of patient care on all measures

Ranked performances (Table 2) highlight that staff on positively deviant wards perceived care to be safer than staff on comparison wards as measured by the Patient Safety Grade. This was also largely true for patients, although their perceptions were less concordant with the ST data. Positively deviant ward 6 displayed the lowest PMOS and FFT scores, and two comparators (wards 8 and 9) performed better than some positive deviants on certain patient measures. The scatterplot of z-scores (supplementary file 4) compared performances across different normal distributions. Positively deviant wards predominantly performed above the mean and, as a group, generally performed better on all measures than the comparison wards supporting their identification using ST data.

Table 2: Visual representation of ranked ward level performance across all patient safety measures

Rank	ST Harm-free care (Phase 1)	Patient Measure of Safety	Friends and Family Test	CQUIN	Patient Safety Grade
1 (high)	Ward 3	Ward 10	Ward 8	Ward 3	Ward 1
2	Ward 6	Ward 8	Ward 1	Ward 10	Ward 3
3	Ward 10	Ward 1	Ward 10	Ward 9	Ward 10
4	Ward 1	Ward 9	Ward 3	Ward 1 and	Ward 6
5	Ward 2	Ward 3	Ward 9	Ward 8	Ward 4
6	Ward 9	Ward 2	Ward 4 and	Ward 6	Ward 5
7	Ward 8	Ward 4	Ward 2	Ward 4	Ward 8
8	Ward 4	Ward 5	Ward 5	Ward 2	Ward 2
9 (low)	Ward 5	Ward 6	Ward 6	Ward 5	Ward 9

Positively deviant wards are shaded in colour. Comparison wards are represented in white. Higher ranks represent safer perceptions of patient care on all measures.

Abbreviations: ST = Safety Thermometer; CQUIN = Commissioning for Quality and Innovation;

DISCUSSION

This study developed and critically appraised a method for conducting stage 1 of the Bradley et al.(7) positive deviance framework – identifying positive deviants. Previous applications have typically identified positively deviant individuals or organisations and have focused on narrow

1
2
3 processes or outcomes of care.(9) However, performance variation also exists between
4 wards/units(18). Although there is merit to focusing on specific aspects of care, this can divert
5 attention away from other important aspects of safety reducing opportunities for wider
6 improvement, for example, implementing cultural changes that improve several different outcomes
7 (e.g. improving multi-disciplinary teamwork).
8
9

10
11 During phase 1 of this study, a robust yet pragmatic analysis successfully identified a discreet
12 group of five statistically different positively deviant elderly medical wards with exceptionally high ST
13 harm-free care performances. Although these wards did rank top of the region, rankings alone did
14 not differentiate between positively deviant wards and those that performed well. This study
15 therefore advanced the previous methods that have been used to identify positive deviants by
16 identifying a distinct and statistically different group of wards that not only ranked best within the
17 region but also outperformed their NHS Trust (organisation), performed consistently over time, and
18 demonstrated success beyond what would be expected through chance. Statistical Process Control
19 methods such as funnel plots and run charts are increasingly promoted for assessing performance
20 variation within healthcare.(35, 36) They combine statistical rigour with sensitive measurement to
21 differentiate between variation that is expected by chance and variation that has an assignable
22 cause.(46, 47) The methods are also considered to be relatively intuitive and pragmatic enough for
23 use by improvers on the frontline.(36)
24
25
26
27
28
29
30
31
32

33 However, the extent to which wards truly demonstrated exceptional performance can be
34 questioned. Wards did not exceed the funnel plot's three standard error control limits and so their
35 exceptional, outlier status was limited. Furthermore, although positively deviant wards differed
36 statistically from others, the minimal performance differences between them and the 'next best'
37 wards highlight the importance of considering clinical significance – that is, whether differences
38 meaningfully affect patient treatment.(48) Positive deviants are supposed to demonstrate
39 'exceptionally' high performance,(7) but there is little consensus in the literature about how to
40 differentiate between high performance and positively deviant performances. Control limits are
41 considered to be conservative and can be made more or less stringent depending on the
42 context.(34) In the absence of extreme outliers, it may still be possible to generate useful learning
43 from 'positive deviants' who simply perform well.
44
45
46
47
48
49
50

51 During phase 2 of the study, staff and patient perceptions of safety did, in the main, corroborate
52 the routinely collected ST data, providing tentative support for the methods used to identify
53 positively deviant wards. Patients on positively deviant wards, though, did not uniformly perceive
54
55
56
57

1
2
3 their care to be safer than those on comparator wards. This could be explained by wards adopting
4 different approaches to delivering safe patient care; for example, if staff emphasise guideline
5 compliance over patient centred care this may influence patient perceptions of safety. Furthermore,
6 patients may have unique perspectives of safety, which perhaps encompass the culture of a ward
7 rather than just the outcomes that are reflected in the routinely collected data.(49, 50) The lack of
8 agreement between the FFT, CQUIN, and ST measures may also have arisen by measuring two
9 associated but distinct quality domains – patient experience versus safety.(51) Fundamentally
10 though, inconsistent patient perceptions highlight that different positively deviant wards may have
11 been identified had a different broad measure of safety (other than the ST) been used.
12
13
14
15
16
17

18 In addition to these considerations, there are various overarching considerations and challenges
19 that are faced when applying the Bradley et al. framework.(7) First, there are few sources of
20 routinely collected data within the UK's NHS or further afield that broadly measure safety and are
21 publicly available at ward level.(52) This makes it difficult to adopt a pragmatic approach and identify
22 positive deviants across different organisations. Although this study provides tentative support for
23 using ST data within a UK healthcare setting, the harms measured within this tool are particularly
24 pertinent to older people. Consequently this measure may lack relevance to applications of the
25 approach which seek to identify positive deviants in different settings, for example, paediatric wards
26 or emergency departments.
27
28
29
30
31
32

33 Second, positive deviants are assumed to succeed *despite facing the same constraints as*
34 *others*(6) and so it is critical to identify them from within a homogenous population. This study
35 increased homogeneity by defining elderly medical wards as stringently as possible and by sampling
36 matched comparators. However, numerous factors are known to contribute to patient safety
37 incidents(53) and case-mix adjustments are notoriously difficult.(54) One can never fully control for
38 all confounding variables when identifying positive deviants, but this is likely to be especially difficult
39 when comparing performance on broad outcomes of safety, when using publicly available data, and
40 when adopting a pragmatic approach.
41
42
43
44
45
46

47 Third, although the problems associated with routinely collected data are well documented,(22-
48 24) there are also wider implications of using routine data to identify positive deviants. Performance
49 variation can arise because measurement is conducted in a social context – staff don't make
50 decisions about the same things nor decide things in the same way.(55) This is problematic when the
51 positive deviance approach seeks to compare performances across several different healthcare
52 providers. Furthermore, healthcare organisations retrospectively measure the absence rather than
53
54
55
56
57
58
59
60

1
2
3 presence of safety(56) and measurement and monitoring systems say nothing about how safe
4 patient care currently is or how safe it will be in the future.(57) This compounds the ability to
5 accurately identify positive deviants and thus the ability to reliably conduct subsequent stages of the
6 approach.
7
8
9

10 **Study limitations**

11
12
13 Various study limitations have already been highlighted including measuring statistical rather
14 than clinical differences between positively deviant and comparison wards. Due to resource
15 constraints, it was also not possible to assess staff and patient perceptions of safety across all 34
16 wards that were sampled during phase 1 of the study. The resulting small and dichotomised sample
17 during phase 2 meant that the associations between staff and patient perceptions of safety and the
18 ST data could not be assessed statistically. Furthermore, some of the differences in performance
19 between positively deviant and comparison wards were small. Had the comparison group comprised
20 negative deviants (the worst performers) rather than above average performers, the differences
21 between the two groups on each of the measures may have been more stark. Many previous
22 applications of the positive deviance approach have sampled positive and negative deviants.(31)
23 However, this specific comparison group was chosen with the wider application of positive deviance
24 in mind,(31) so that when we explore how positive deviants succeed (stage 2 of the framework(7))
25 we can distinguish between *exceptional* and good performances, not just explore how teams differ
26 from the worst in the population.
27
28
29
30
31
32
33
34
35

36 Finally, as with many routinely collected measures of quality and safety, the reliability and validity
37 of the ST has been questioned. ST data are collected opportunistically at a single monthly time
38 point,(27) harm definitions are subject to interpretation,(58, 59) and data collection was previously
39 incentivised. However, the ST is used to measure performance in most acute NHS Trusts and it is the
40 only routinely collected, broad measure of safety that is publicly available at ward level in the UK.
41 Additional ST tools, such as the Medications Safety Thermometer, have been developed(60, 61) and
42 so, if these can also be used to identify positively deviant wards, then the methods tested in this
43 study could have greater impact across the NHS.
44
45
46
47
48
49
50
51

52 **Conclusions**

53
54
55
56
57

1
2
3 This study has shown that a distinct group of positively deviant wards that perform exceptionally
4 well on a broad measure of safety can be identified using a robust yet pragmatic method, and that
5 staff and patient perceptions of safety do, in the main, support their identification. It has highlighted
6 the challenges faced when selecting a source of routinely collected data that provides a valid and
7 reliable measure at the appropriate level in order to facilitate performance comparisons across
8 wards or units in several organisations. Many of these challenges are applicable to a variety of
9 different settings and applications of the approach and so this study may provide generalisable
10 guidance on the methods that can be used to effectively identify positive deviants.
11
12
13
14
15
16
17
18
19
20
21

22 **Funding:** The Health Foundation - PhD in Improvement Science.
23

24 **Competing interests:** None declared.
25
26

27 **Acknowledgements:** The authors express their sincere gratitude to all of the ward staff and patients
28 who participated in and supported this study. Thank you to Carolyn Reynolds and Alex Howat who
29 helped collect some of the data. RB, RL and IK are members of the NIHR Collaboration for Leadership
30 in Applied Health Research and Care, Yorkshire and Humber – Evidence Based Transformation
31 Theme. The research was supported by the NIHR CLAHRC Yorkshire and Humber. [www.clahrc-](http://www.clahrc-yh.nihr.ac.uk)
32 [yh.nihr.ac.uk](http://www.clahrc-yh.nihr.ac.uk). The views expressed are those of the author(s), and not necessarily those of the NHS,
33 the NIHR or the Department of Health.
34
35
36
37
38

39 **Author contributions:** RB, RL, NT and IK developed the idea for the study. RB, MM and VP conducted
40 the analyses to identify positive deviants. RB collected and analysed the primary data to assess staff
41 and patient perceptions of safety. RB drafted the manuscript. All authors provided comments and
42 approved the final version.
43
44
45
46

47 **Patient consent:** Obtained.
48
49

50 **Ethics approval:** South East Scotland Research Ethics Committee (ref 14/SS/1085). NHS permissions
51 were gained from all five NHS trusts involved.
52
53

54 **Data sharing statement:** Anonymised data are available on request.
55
56
57
58
59
60

References

1. Hogan H, Healey F, Neale G, et al. Preventable deaths due to problems in care in English acute hospitals: a retrospective case record review study. *BMJ Qual Saf*. 2012:Published Online First: 07 July 2012.
2. Landrigan CP, Parry GJ, Bones CB, et al. Temporal trends in rates of patient harm resulting from medical care. *New England Journal of Medicine*. 2010;363(22):2124-34.
3. Hollnagel E, Braithwaite J, Wears RL. Resilient health care. UK: Ashgate; 2013.
4. Vincent C, Amalberti R. Safer healthcare: Strategies for the real world.: Springer Open; 2016. Available from: <http://link.springer.com/book/10.1007%2F978-3-319-25559-0>.
5. Lawton R, Taylor N, Clay-Williams R, et al. Positive deviance: a different approach to achieving patient safety. *BMJ Quality & Safety*. 2014:Published online first 21 July 2014. PubMed PMID: 25049424. Epub 2014/07/23. Eng.
6. Marsh DR, Schroeder DG, Dearden KA, et al. The power of positive deviance. *BMJ*. 2004;329(7475):1177-9. PubMed PMID: WOS:000225169600033.
7. Bradley EH, Curry LA, Ramanadhan S, et al. Research in action: using positive deviance to improve quality of health care. *Implementation Science*. 2009;4:25.
8. Dixon-Woods M, Martin GP. Does quality improvement improve quality? *Future Hospital Journal*. 2016;3(3):191-4.
9. Baxter R, Taylor N, Kellar I, et al. What methods are used to apply positive deviance within healthcare organisations? A systematic review. *BMJ Quality and Safety*. 2015:Published online first 20th November.
10. Rose AJ, McCullough MB. A practical guide to using the positive deviance method in health services research. *Health Serv Res*. 2016;Published online 28th June PubMed PMID: 27349472.
11. Rose AJ, Petrakis BA, Callahan P, et al. Organizational characteristics of high- and low-performing anticoagulation clinics in the Veterans Health Administration. *Health Serv Res*. 2012;47(4):1541-60. PubMed PMID: 22299722. Pubmed Central PMCID: 3401398.
12. Gabbay RA, Friedberg MW, Miller-Day M, et al. A positive deviance approach to understanding key features to improving diabetes care in the medical home. *Ann Fam Med*. 2013;1(Suppl 1):S99-107. PubMed PMID: 23690393. Pubmed Central PMCID: PMC3707253 [Available on 11/01/13].
13. Austin JM, Jha AK, Romano PS, et al. National hospital ratings systems share few common scores and may generate confusion instead of clarity. *Health Affairs*. 2015;34(3):423-30. PubMed PMID: 25732492.
14. Healthcare Association of New York State. HANYS' report on report cards: understanding publicly reported hospital quality measures. US, New York: HANYS Quality Institute, 2013.
15. Rothberg MB, Morsi E, Benjamin EM, et al. Choosing the best hospital: the limitations of public quality reporting. *Health Aff (Millwood)*. 2008;27(6):1680-7. PubMed PMID: 18997226. Epub 2008/11/11. eng.
16. Pannick S, Wachter RM, Vincent C, et al. Rethinking medical ward quality. *BMJ*. 2016;355(i5417).
17. Rose JS, Thomas CS, Tersigni AR, et al. A leadership framework for culture change in health care. *The Joint Commission Journal on Quality and Patient Safety*. 2006;32(8):433-42.
18. Schwendimann R, Zimmermann N, Küng K, et al. Variation in safety culture dimensions within and between US and Swiss Hospital Units: an exploratory study. *BMJ Qual Saf*. 2013;22:32-41.
19. Nelson EC, Batalden PB, Huber TP, et al. Microsystems in health care: Part 1. Learning from high-performing front-line clinical units. *Jt Comm J Qual Patient Saf*. 2001;28(9):472-93.
20. Reason J. Understanding adverse events: human factors. *Quality in Health Care*. 1995;4(2):80-9.

21. Vincent C, Taylor-Adams S, Stanhope N. Framework for analysing risk and safety in clinical medicine. *BMJ*. 1998;316(7138):1154-7.
22. Lilford R, Mohammed MA, Spiegelhalter D, et al. Use and misuse of process and outcome data in managing performance of acute medical care: avoiding institutional stigma. *The Lancet*. 2004;363(9415):1147-54.
23. Shaw J, Taylor R, Dix K. Uses and abuses of performance data in healthcare. UK: Dr Forster, 2015.
24. Zhan C, Miller MR. Administrative data based patient safety research: a critical review. *Quality and Safety in Health Care*. 2003;12(Supp 2):58-63.
25. Digital N. Hospital Episode Statistics 2017 [cited 2017 13th October]. Available from: <http://content.digital.nhs.uk/hes>.
26. Institute TP. The NHS Staff Survey 2017 [cited 2017 13th October 2017]. Available from: <http://www.nhsstaffsurveys.com/Page/1056/Home/NHS-Staff-Survey-2017/>.
27. Power M, Stewart K, Brotherton A. What is the NHS Safety Thermometer? *Clinical Risk*. 2012;18(5):163-9. PubMed PMID: 23136533. Pubmed Central PMCID: 3484316. Epub 2012/11/09. Eng.
28. Sari AB, Cracknell A, Sheldon TA. Incidence, preventability and consequences of adverse events in older people: Results of a retrospective case-note review. *Age and Ageing*. 2008;37(3):265-9. PubMed PMID: 18332053.
29. Thomas EJ, Brennan TA. Incidence and types of preventable adverse events in elderly patients: population based review of medical records. *BMJ*. 2000;320(7237):741-4.
30. Dixon-Woods M, McNicol S, Martin G. Ten challenges in improving quality in healthcare: lessons from the Health Foundation's programme evaluations and relevant literature. *BMJ Qual Saf*. 2012:Published Online First: 28 April 2012.
31. Baxter R, Taylor N, Kellar I, et al. Learning from positively deviant wards to improve patient safety: an observational study protocol. *BMJ Open*. 2015;5:e009650.
32. Health and Social Care Information Centre. NHS Safety Thermometer. 2014 [cited 2014 29th April]. Available from: <http://www.hscic.gov.uk/thermometer>.
33. Mayer EK, Bottle A, Rao C, et al. Funnel plots and their emerging application in surgery. *Ann Surg*. 2009;249(3):376-83. PubMed PMID: 19247021.
34. Perla RJ, Provost LP, Murray SK. The run chart: a simple analytical tool for learning from variation in healthcare processes. *BMJ Qual Saf*. 2011;20(1):46-51. PubMed PMID: 21228075.
35. Mountford J, Wakefield D. From stoplight reports to time series: equipping boards and leadership teams to drive better decisions. *BMJ Quality & Safety*. 2016;Published online first: 7th March.
36. 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*. 2016:Published online first 31 March.
37. Office of the Deputy Prime Minister. The English Indices of Deprivation 2004: Summary (revised). UK: Office of the Deputy Prime Minister; 2004.
38. Giles SJ, Lawton RJ, Din I, et al. Developing a patient measure of safety (PMOS). *BMJ Quality and Safety*. 2013;22:554-62.
39. McEachan RRC, Lawton RJ, O'Hara JK, et al. Developing a reliable and valid patient measure of safety in hospitals (PMOS): a validation study. *BMJ Quality and Safety*. 2014:Published online first 24 December 2013.
40. NHS England. Friends and family test. 2014 [cited 2014 29th April]. Available from: <http://www.england.nhs.uk/statistics/statistical-work-areas/friends-and-family-test/>.
41. Department of Health. Using the Commissioning for Quality and Innovation (CQUIN) payment framework – Guidance on national goals for 2011/12. 2010 [cited 2016 3rd April]. Available from:

- https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/215049/dh_133859.pdf.
42. Sarac C, Flin R, Mearns K, et al. Hospital survey on patient safety culture: psychometric analysis on a Scottish sample. *BMJ Quality and Safety*. 2011;20(10):842-8. PubMed PMID: 21690247. Epub 2011/06/22. eng.
 43. Sorra J, Dyer N. Multilevel psychometric properties of the AHRQ hospital survey on patient safety culture. *BMC Health Service Research*. 2010;10:199. PubMed PMID: 20615247. Pubmed Central PMCID: 2912897. Epub 2010/07/10. eng.
 44. Sorra J, Nieva V. Hospital survey on patient safety culture. USA, Rockville: Westat, AHRQ Publication No. 04-0041, 2004.
 45. Edwards P, Roberts I, Clarke M, et al. Increasing response rates to postal questionnaires: systematic review. *BMJ*. 2002;324(7347):1183.
 46. Benneyan JC, Lloyd RC, Plsek PE. Statistical process control as a tool for research and healthcare improvement. *Qual Saf Health Care*. 2003;12:458-64.
 47. Mohammed MA. Using statistical process control to improve the quality of health care. *Quality and Safety in Health Care*. 2004;13(4):243-5.
 48. Oberst MT. Clinical versus statistical significance. *Cancer Nursing*. 1982;5(6):475-6.
 49. Ocloo J, Matthews R. From tokenism to empowerment: progressing patient and public involvement in healthcare improvement. *BMJ Quality & Safety*. 2016:Published online first 18th March.
 50. Weingart SN, Pagovich O, Sands DZ, et al. What can hospitalized patients tell us about adverse events? Learning from patient-reported incidents. *J Gen Intern Med*. 2005;20(9):830-6. PubMed PMID: 16117751. Pubmed Central PMCID: 1490203. Epub 2005/08/25. eng.
 51. Doyle C, Lennox L, Bell D. A systematic review of evidence on the links between patient experience and clinical safety and effectiveness. *BMJ open*. 2013;3(1):e001570.
 52. Pannick SAJ, Wachter RM, Vincent C, et al. Rethinking medical ward quality. *BMJ*. 2016;355:i5417.
 53. Lawton R, McEachan RR, Giles SJ, et al. Development of an evidence-based framework of factors contributing to patient safety incidents in hospital settings: a systematic review. *BMJ Quality & Safety*. 2012:Published Online First: 15 March 2012.
 54. Shahian DM, Normand SLT. What is a performance outlier? *BMJ Qual Saf*. 2015;24:95-9.
 55. Dixon-Woods M, Leslie M, Bion J, et al. What counts? An ethnographic study of infection data reported to a patient safety program. *The Milbank Quarterly*. 2012;90(3):548-91.
 56. Hollnagel E, Wears RL, Braithwaite J. From Safety-I to Safety-II: A white paper. Published simultaneously by the University of Southern Denmark, University of Florida, USA, and Macquarie University, Australia: The Resilient Health Care Net, 2015.
 57. Vincent C, Burnett S, Carthey J. The measurement and monitoring of safety. UK: The Health Foundation, 2013.
 58. Buckley C, Cooney K, Sills E, et al. Implementing the Safety Thermometer tool in one NHS trust. *British Journal of Nursing*. 2014;23(5):268-72.
 59. Power M, Fogarty M, Madsen J, et al. Learning from the design and development of the NHS Safety Thermometer. *International Journal for Quality in Health Care*. 2014;26(3):287-97.
 60. NHS Quality Observatory. NHS Safety Thermometer 2013 [cited 2017 12th May]. Available from: https://www.safetythermometer.nhs.uk/index.php?option=com_content&view=article&id=1&Itemid=101.
 61. Rostami P, Power M, Harrison A, et al. Learning from the design, development and implementation of the Medication Safety Thermometer. *International Journal for Quality in Health Care*. 2016;29(2):301-9.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

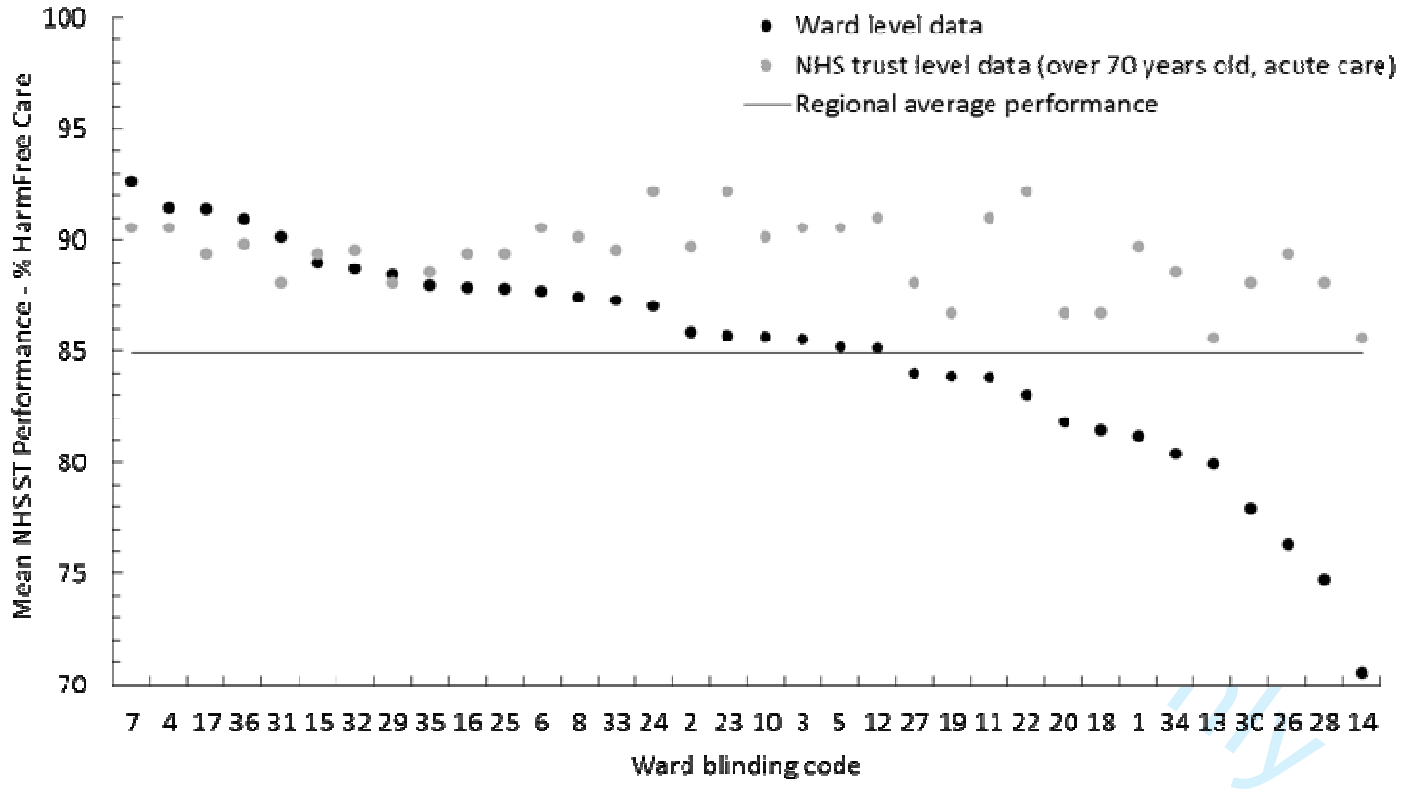


Figure 1: Scatterplot comparing average harm-free performances at ward and NHS Trust (organisation) levels.

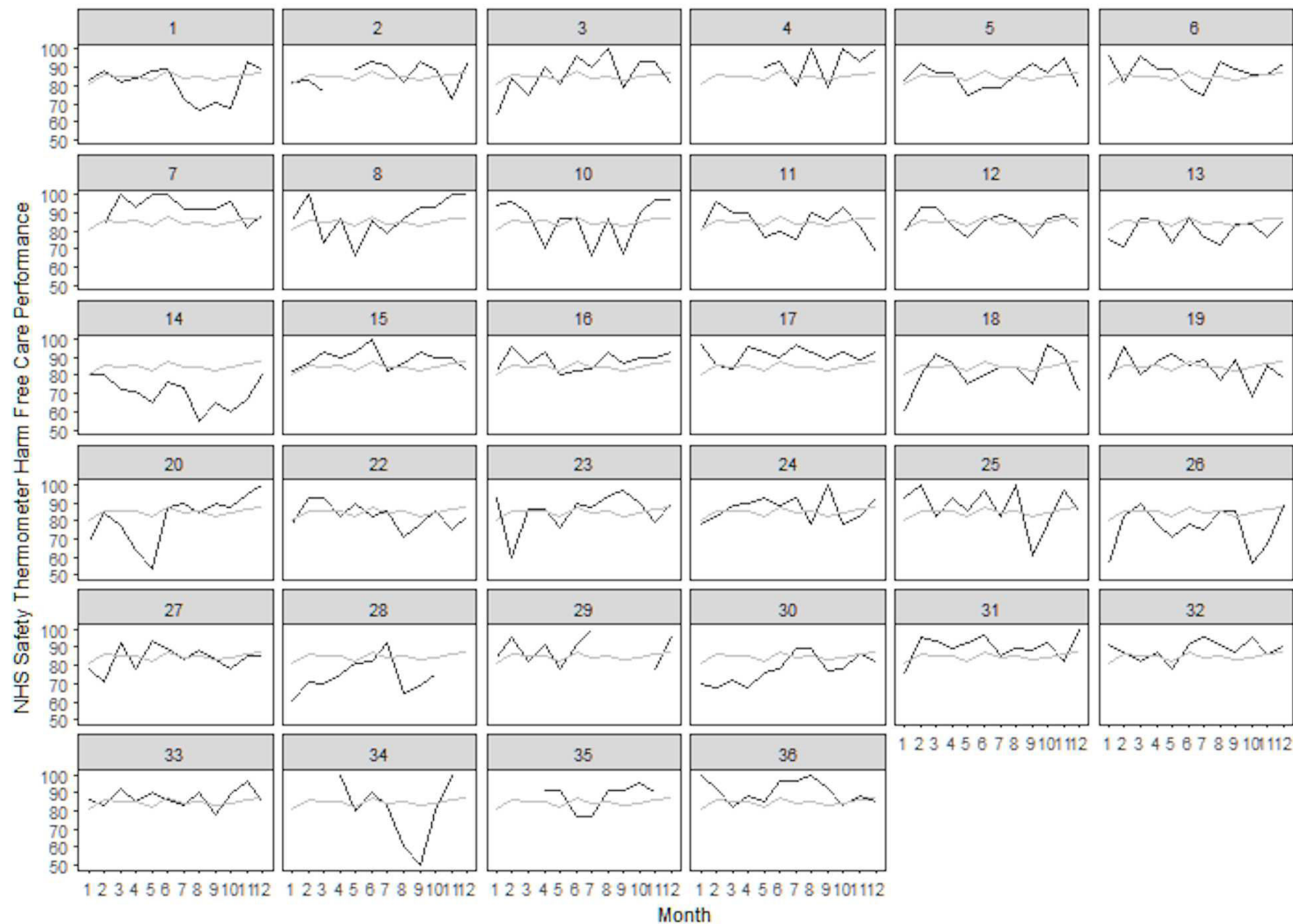


Figure 2: Run charts comparing ward and regional level monthly harm-free care performance across a 12 month period.

Each square represents an individual ward within the population. Wards are numbered consecutively according to their pseudonyms (from top left to bottom right across the rows).

Black lines represent a ward's monthly harm-free care performance. Red lines represent the population's average monthly performance.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

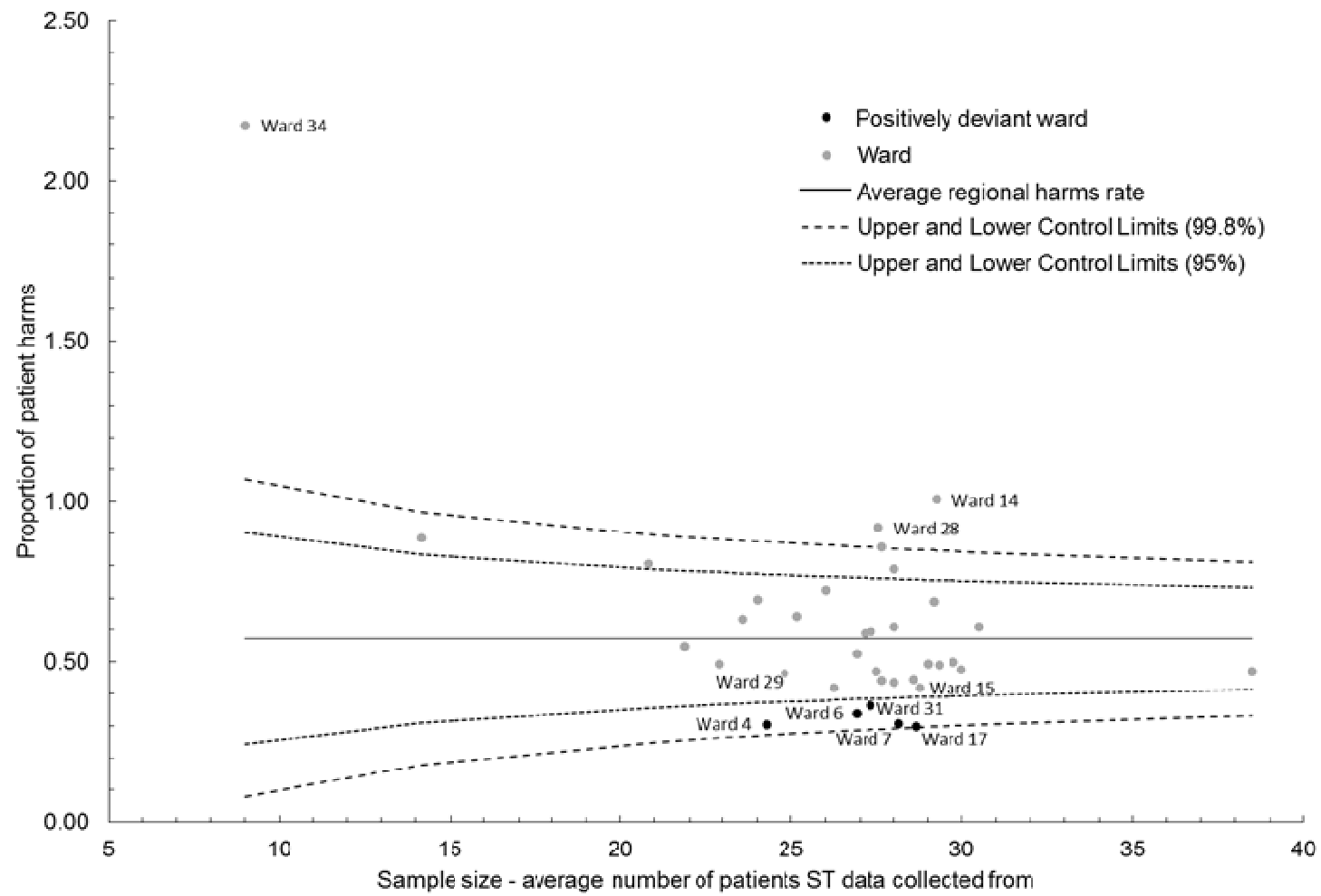


Figure 3: Funnel plot of average ST harm-free care performance and average sample size.

Supplementary file 1

Ward rankings for each of the NHS Safety Thermometer measures

Ranking	Ward blinding code	ST harmfree care performance ^a	New PUs ^b	Falls	New UTI	New VTE
1	4	92.68	1.01	0.36	0.00	1.36
2	7	91.48	0.00	0.74	0.40	0.74
3	17	91.40	0.58	0.60	0.29	0.60
4	36	90.97	2.55	0.32	0.30	0.00
5	31	90.14	2.09	1.53	0.31	0.30
6	15	88.97	1.16	0.00	0.00	0.00
7	32	88.70	2.21	1.09	0.36	2.93
8	29	88.48	2.09	0.79	1.20	1.20
9	35	88.19	1.65	0.82	0.27	1.65
10	16	88.01	1.14	1.16	0.57	2.27
11	25	87.90	1.15	1.75	1.15	1.99
12	6	87.72	1.51	0.60	0.30	1.80
13	8	87.46	1.11	0.00	1.67	0.56
14	33	87.28	1.99	1.19	0.00	1.14
15	24	86.52	0.00	1.31	1.33	1.32
16	2	85.87	0.66	0.69	0.65	1.97
17	23	85.71	0.00	1.13	0.58	0.84
18	10	85.68	3.32	0.83	0.83	0.28
19	3	85.57	2.22	1.42	0.00	0.27
20	5	85.17	1.44	2.11	0.36	0.00
21	12	85.14	3.05	2.25	1.70	0.00
22	27	84.06	0.60	4.57	1.21	0.30
23	19	83.90	1.58	1.52	0.95	1.91
24	11	83.81	3.99	2.20	0.29	0.58
25	22	83.04	1.79	0.60	0.60	0.60
26	20	81.83	4.42	5.98	0.00	0.43
27	18	81.45	2.15	0.79	0.32	1.13
28	1	81.20	2.74	1.62	0.93	1.63
29	34	80.42	2.50	1.25	2.50	3.75
30	13	79.91	3.52	1.14	2.08	0.84
31	30	77.95	0.58	2.09	5.65	0.88
32	26	76.28	0.89	6.32	2.41	3.92
33	28	74.69	4.70	1.03	2.67	3.27
34	14	70.56	8.02	1.13	5.14	2.01
Average		84.90	2.03	1.52	1.08	1.21

* PUs = pressure ulcers; UTI = urinary tract infections, VTE = venous thromboembolism

^a The wards highlighted blue are positive deviants and the wards highlighted green are matched comparison wards.

^b Performances for each individual harm are listed in columns 4-8 (not ranked in order). The top five performers across the region are highlighted yellow.

Supplementary file 2

Key characteristics of the positively deviant and comparison wards

	Ward	ST Harm-free care (%)	Trust number / type	Patient gender ^a	Index of Multiple Deprivation ^b
Positively deviant wards	1	90.14	Trust 1 / Teaching & Foundation	Mixed	More deprived 40-50%
	3	92.68	Trust 2 / Teaching	Female	More deprived 30-40%
	6	91.48	Trust 2 / Teaching	Female	More deprived 30-40%
	7	91.40	Trust 3 / Teaching & Foundation	Mixed	Less deprived 30-40%
	10	90.97	Trust 5 / Foundation	Mixed	More deprived 30-40%
Comparison wards	2	88.48	Trust 1 / Teaching & Foundation	Mixed	More deprived 40-50%
	4	87.72	Trust 2 / Teaching	Female	More deprived 30-40%
	5	85.17	Trust 2 / Teaching	Male	More deprived 30-40%
	8	87.90	Trust 3 / Teaching & Foundation	Mixed	Less deprived 30-40%
	9	88.01	Trust 4 / Foundation	Mixed	More deprived 30-40%

^a One positively deviant ward had to be matched to a comparison ward that cared for patients of the opposite gender.

^b IMD overall rank data (extracted from the 2012/13 Hospital Episode Statistics data) are categorised into deciles. Geographic areas are ranked and then described as falling within the most or least deprived % of England. Categories change in increments of 10% up to the more/least deprived 40-50% of England.

Supplementary file 3: Patient and staff recruitment information

Table 1: Ward level recruitment data for staff and patients

		Patient response rate n (% of those approached) ^a	Staff response rate n (approx. % of the MDT)
Positively deviant wards	T1W1	21 (51)	22 (44)
	T2W3	22 (61)	14 (40)
	T2W6	21 (47)	11 (31)
	T5W10	20 (51)	30 (67)
Comparison wards	T1W2	17 (53)	18 (45)
	T2W4	23 (48)	14 (40)
	T2W5	20 (49)	19 (54)
	T3W8	20 (69)	16 (46)
	T4W9	24 (83)	17 (35)
All wards		188 (55)	161 (45)

^a Response rates include patients who explicitly refused to participate. It does not include: those whom nurses reported to be eligible but were subsequently considered unsuitable; and those who could not be followed up after providing time to consider participation.

Table 2: Characteristics of patient participants

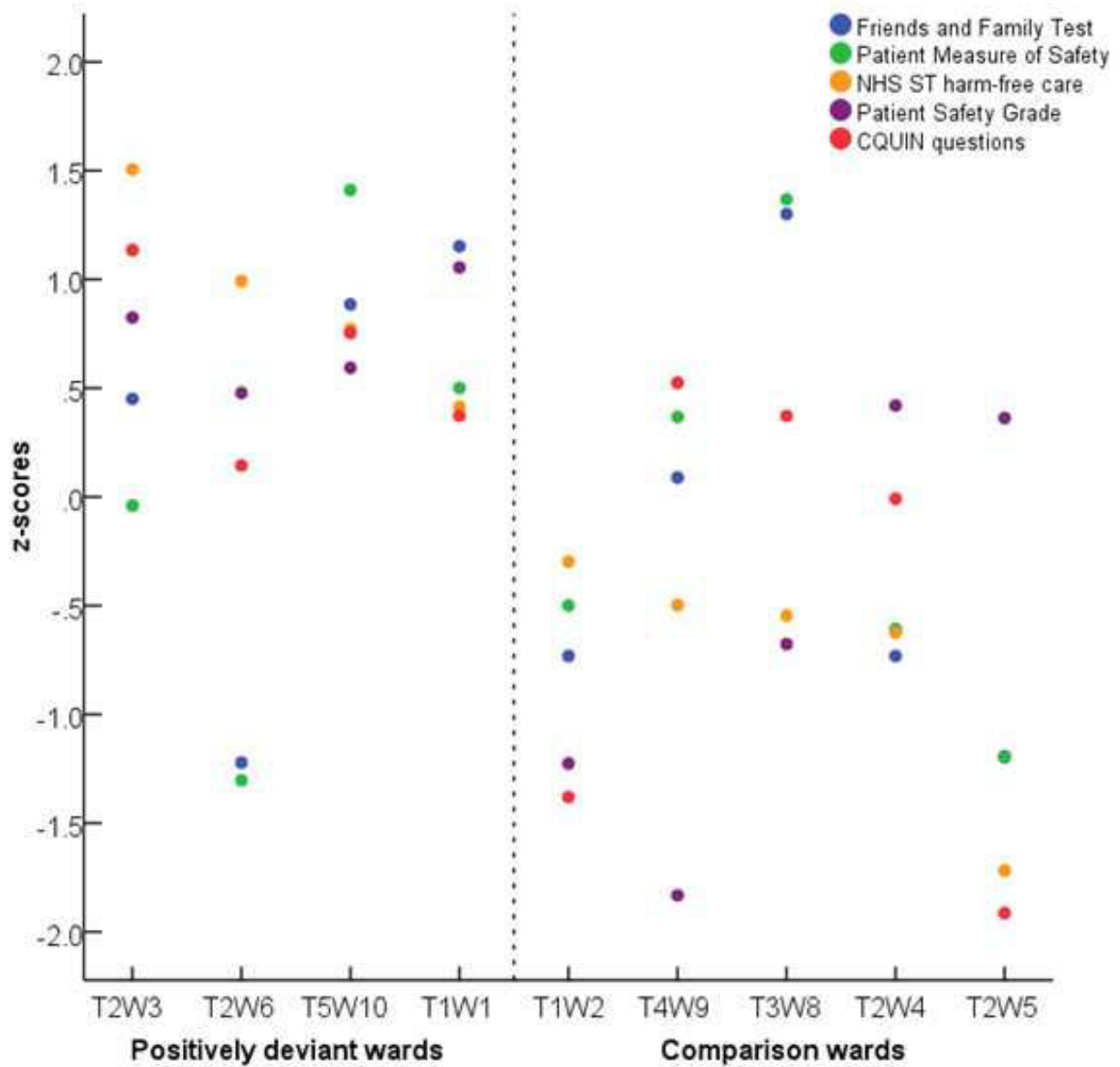
	Patient age mean years (SD)	Ongoing hospital treatment % yes	Time in hospital mean days (SD)	Inpatient frequency mean (SD)
Positively deviant wards	84.49 (5.60)	33%	13.6 (12.87)	2.27 (2.46)
Comparison wards	84.56 (5.36)	46% (3 missing)	15.71 (19.64)	2.71 (3.32)
All wards	84.53 (5.45)	41%	14.75 (16.91)	2.51 (2.97)

Table 3: Professional roles of staff participants

	Nursing %	Healthcare Assistants %	Allied Health Profes %	Doctors %	Other %
Positively deviant wards	46.8	19.5	15.6	6.5	11.7
Comparison wards	36.9	31.0	13.1	1.2	17.8
All wards	41.6	25.5	14.3	3.7	15.0

Supplementary file 4

Scatterplot of z-scores to assess whether other perceptions of safety support the use of ST data to identify positive deviants.



BMJ Open

Identifying positively deviant elderly medical wards using routinely collected NHS Safety Thermometer data: an observational study.

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2017-020219.R1
Article Type:	Research
Date Submitted by the Author:	08-Dec-2017
Complete List of Authors:	Baxter, Ruth; Bradford Institute for Health Research, Yorkshire Quality and Safety Research Group; University of Leeds, School of Psychology Taylor, Natalie; Cancer Council, Cancer Research Division Kellar, Ian; University of Leeds, School of Psychology Pye, Victoria; Macquarie University, Australian Institute of Health Innovation Mohammed, Mohammed ; faculty of Health care studies, University of Bradford, Reserach Lawton, Rebecca; Bradford Institute for Health Research, Yorkshire Quality and Safety Research Group; University of Leeds, School of Psychology
Primary Subject Heading:	Health services research
Secondary Subject Heading:	Research methods
Keywords:	Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, STATISTICS & RESEARCH METHODS, positive deviance, quality improvement, implementation science

SCHOLARONE™
Manuscripts

1
2
3
4 **Identifying positively deviant elderly medical wards using routinely collected**
5 **NHS Safety Thermometer data: an observational study.**
6
7
8
9
10
11
12

13 **Dr. Ruth Baxter^{1,2*}, Dr. Natalie Taylor³, Dr Ian Kellar², Ms. Victoria Pye⁴, Prof. Mohammed A**
14 **Mohammed^{1,5}, Prof Rebecca Lawton^{1,2}.**
15

16 1 - Yorkshire Quality and Safety Research group, Bradford Institute for Health Research, Bradford, UK
17

18 2 – School of Psychology, University of Leeds, Leeds, UK
19

20 3 – Cancer Research Division, Cancer Council, Sydney, Australia
21

22 4 – Australian Institute of Health Innovation, Macquarie University, NSW, Australia
23

24 5 – Faculty of Health Studies, University of Bradford, Bradford, UK
25

26 *corresponding author: ruth.baxter@bthft.nhs.uk; 01274 383413; Yorkshire Quality and Safety
27 Research group, Bradford Institute for Health Research, Bradford Royal infirmary, Duckworth Ln,
28 Bradford, BD9 6RJ.
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Word Count: 4431

ABSTRACT (301/300 words)

Objective: The positive deviance approach seeks to identify and learn from exceptional performers. Although a framework exists to apply positive deviance within healthcare organisations, there is limited guidance to support its implementation. The approach has also rarely explored exceptional performance on broad outcomes, been implemented at ward level, or applied within the United Kingdom (UK). This study develops and critically appraises a pragmatic method for identifying positively deviant wards using a routinely collected, broad measure of patient safety.

Design: A two-phased observational study was conducted. During phase 1, cross-sectional and temporal analyses of Safety Thermometer data were conducted to identify a discrete group of positively deviant wards that consistently demonstrated exceptional levels of safety. A group of matched comparison wards with above average performances were also identified. During phase 2, multidisciplinary staff and patients on the positively deviant and comparison wards completed surveys to explore whether their perceptions of safety supported the identification of positively deviant wards.

Setting: 34 elderly medical wards within a northern region of England, UK.

Participants: Multidisciplinary staff (n=161) and patients (n=188) clustered within nine positively deviant and comparison wards.

Results: Phase 1: A combination of analyses identified five positively deviant wards that performed best in the region, outperformed their organisation, and performed consistently well over 12 months. Five above average, matched comparator wards were also identified. Phase 2: Staff and patient perceptions of safety generally supported the identification of positively deviant wards using Safety Thermometer data, although patient perceptions of safety were less concordant with the routinely collected data.

Conclusions: This study tentatively supports a pragmatic method of using routinely collected data to identify positively deviant elderly medical wards; however, it also highlights the various challenges that are faced when conducting the first stage of the positive deviance approach.

Registration: UK Clinical Research Network Portfolio (reference-18050).

Key words

Positive deviance, implementation science, quality improvement, NHS Safety Thermometer, elderly care, patient safety

Strengths and limitations of this study

- There is limited guidance to support the identification of positive deviants in healthcare settings. This study develops a method for identifying positive deviants using routinely collected data.
- A combination of four different analyses (including performance rankings, comparisons with organisational level performances, and Statistical Process Control methods) were conducted to provide a pragmatic yet robust method for identifying a discrete group of positively deviant wards that performed exceptionally well on a broad outcome of safety.
- Staff and patient perceptions of safety were measured using validated surveys to explore whether they supported the identification of positively deviant wards using routinely collected data.
- Due to the small sample size (n=9 wards) it was not possible to statistically assess whether staff and patient perceptions of safety supported the identification of positive deviants using routinely collected Safety Thermometer data.
- The study was conducted on elderly medical wards and so further research is required to explore whether the methods can be generalised to identify positive deviants in other healthcare settings.

BACKGROUND

Despite extensive efforts to improve, patient safety continues to be a pervasive problem across the globe.(1, 2) Traditionally, these efforts have focused on past errors and harm, but there are increasing calls to also explore how 'safe' patient care is delivered.(3, 4) Positive deviance provides an asset based approach to improving the quality and safety of healthcare.(5, 6) The approach seeks to identify and learn from those who demonstrate exceptional performance on an outcome of interest.(6) It assumes that solutions to problems already exist within communities and that positive deviants (individuals, teams, or organisations) identify these solutions and succeed despite facing similar constraints as others.

Bradley et al.(7) have proposed a four stage framework to implement the positive deviance approach within healthcare organisations. Positive deviants, who display exceptionally high performance, are identified using routinely collected data (stage 1). Hypotheses about how they succeed are generated using qualitative methods (stage 2). These hypotheses are quantitatively tested in larger, more representative samples (stage 3), and then disseminated to others with the help of key stakeholders (stage 4). Despite the increasing popularity of the positive deviance approach,(8) there is little evidence or practical guidance to support its application within healthcare organisations.(9, 10)

Identifying positive deviants who demonstrate exceptional performance

Identifying positive deviants may be the most crucial stage of the Bradley et al.(7) framework as subsequent stages hinge on its perceived legitimacy. Misidentification could lead to the generation of hypotheses that do not capture the factors that facilitate exceptional performance. Bradley et al.(7) suggest that positive deviants should be identified by ranking routinely collected data, and previous healthcare applications have, for example, identified three of the top 10 clinics with the best anticoagulation control(11) or the top quintile of primary care medical homes with the most improved diabetes outcomes.(12) However, performance rankings can differ depending on the rating systems that are used, creating confusion and contradiction about who demonstrates high and low performances.(13-15) Positive deviants are also supposed to demonstrate *exceptional* rather than just *good* performances on the outcome of interest,(7) but rankings or league tables simply appraise performances along a continuum without differentiating a distinct groups of 'outliers' or positive deviants from the rest of a population.

1
2
3 Previous healthcare applications of the approach have predominantly identified positively
4 deviant organisations (e.g. hospitals) or individuals.(9) This is despite greater amounts of variation
5 existing at the level of a hospital ward or unit(16-18) and the majority of frontline care being
6 delivered by the multidisciplinary teams that work within these clinical microsystems.(19) Previous
7 applications have also typically focused on specific processes or outcomes of care such as hand
8 hygiene compliance and the incidence of healthcare associated infections.(9) Although it is relevant
9 to explore positive deviance in this way, factors that contribute to safety often operate across
10 various levels of the system and affect multiple outcomes.(20, 21) If ward teams succeed on broad
11 outcomes of care there are likely to be some underlying, latent factors that facilitate their success.
12 Understanding these factors and spreading the associated strategies may generate more far-
13 reaching improvements in quality and safety. However, it remains unknown whether positively
14 deviant wards or units can be identified accurately using a routinely collected, broad and
15 multidimensional measure of patient safety.

16
17
18 The shortcomings of using routinely collected data, such as publication lags, coding differences,
19 and data gaming, are well documented.(22-24) Nonetheless, if positive deviance is to become a
20 useful improvement approach, its methods must be pragmatic and accessible for healthcare
21 organisations, networks, and frontline improvers to use. Routinely collected and publicly available
22 data are therefore required, especially when applying the approach across several different wards,
23 units, or organisations. In preparation for this study various routinely collected measures of safety
24 within the UK's National Health Service (NHS) were identified (e.g. Hospital Episode Statistics,(25)
25 and the NHS Staff Survey(26)), but the NHS Safety Thermometer (ST) provides the only routinely
26 collected, broad measure of safety that is publicly available at ward level. Each month the NHS ST is
27 used to measure four commonly occurring patient harms – falls, pressure ulcers, venous-
28 thromboembolism, and urinary tract infections. This data is used to report on the proportion of
29 'harm-free care' that is delivered at ward, speciality and organisational levels.(27) During this study,
30 the NHS ST data will be used to identify positively deviant ward teams that deliver exceptional levels
31 of safe patient care.

32
33
34 One of the key challenges to improving the quality and safety of healthcare is convincing people
35 to adopt a chosen solution.(28) The positive deviance approach seeks to identify solutions from
36 within with the assumption that these solutions will be acceptable to others, feasible to implement,
37 and sustainable over time.(5, 6) However, if staff and/or patients do not perceive positive deviants
38 to be performing exceptionally well, they may not engage with the positively deviant strategies that
39 are disseminated during stage 4 of the Bradley et al.(7) framework. As this study is one of the first
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 applications of the approach to identify positive deviants at ward level using a broad measure of
4 safety, this study also sought to assess the extent to which staff and patient perceptions of safety
5 supported the identification of positively deviant wards via the ST data. A number of validated
6 surveys exist to measure staff perceptions of 'safety culture' – the shared values, beliefs, norms, and
7 attitudes that guide how healthcare staff behave in order to maintain safety.(29) Furthermore, there
8 is increasing evidence to suggest that patients can be involved in maintaining their own safety –
9 patients are able to identify adverse events and can provide a unique perspective on the safety of
10 care.(30-32)
11
12
13
14
15
16
17
18

19 **Study design and aims**

20
21
22 Based on these gaps in the literature, the overarching aim of this observational study was to
23 develop and critically appraise a robust yet pragmatic method for identifying positive deviants at
24 ward level using a routinely collected, broad outcome of safety. The study was conducted in two
25 phases. During phase 1, we sought to apply a rigorous and robust analysis (compared to simply
26 ranking the data) to the ST's harm-free care data to identify a distinct group of positively deviant
27 wards that demonstrated exceptional levels of safety over a 12 month period. To try and identify
28 positively deviant ward teams that were delivering exceptionally safe patient care under challenging
29 circumstances, elderly medical wards were sampled because older patients are particularly
30 vulnerable to safety incidents and harms such as falls and pressure ulcers.(33, 34) Phase 2 of this
31 study aimed to explore whether staff and patient perceptions of safety (which were measured using
32 surveys) were similar to, or at odds with, the routinely collected NHS ST data that had been used to
33 identify the positively deviant wards.
34
35
36
37
38
39
40
41

42 Although this particular study focuses solely on stage 1 of the Bradley et al.(7) framework, it also
43 contributes to a wider application of the positive deviance approach(35) which seeks to generate
44 hypotheses about *how* the positively deviant ward teams deliver exceptionally safe patient care. The
45 qualitative findings from this wider application (stage 2 of the framework) will be published
46 separately. To the authors' knowledge, this study is one of the first applications of positive deviance
47 within the English NHS.
48
49
50
51
52
53
54
55
56
57
58
59
60

PHASE 1 METHOD

This paper adheres to the STROBE Statement (Stengthening the reporting of observational studies in epidemiology).(36) A complete checklist can be found in supplementary file 1.

Sample

Elderly medical wards (n=36) were identified from within 13 acute NHS Trusts (healthcare organisations) in a northern region of England, UK. This represented 10% of all acute NHS Trusts in England. Elderly medical wards were defined as those providing 24 hour, acute, medical care for elderly patients (>65 years); with dedicated multidisciplinary teams; and patient stays typically exceeding 48 hours. Speciality wards (e.g. stroke, rehabilitation, and assessment units) were excluded to maximise homogeneity within the sample.

Data extraction

For all wards the ST data were extracted from a publicly accessible website(37) for a 12 month period – August 2013 to July 2014. Data were extracted for the ST's 'harm-free care' measure and all of the individual ST harms at two different levels: ward level for all patients; and Trust level for acute patients over 70 years.(35) Double blinding during both study phases ensured that researchers (RB, RL, NT, and IK), staff, and patients were not aware of how wards compared to each other on their ST performances.

Analysis

Pragmatic cross-sectional and temporal analyses were conducted to identify a distinct group of positively deviant elderly medical wards that displayed exceptionally high performance on the ST harm-free care measure. Initially, in line with guidance(7) and previous applications,(9) wards were ranked on their average harm-free care performances and then three further analyses were conducted. First, to ensure that positive deviants' success was not simply a function of organisational performance, a scatterplot compared wards with their respective NHS Trust level data. Second, as small sample sizes increase the likelihood of variability being attributable to chance,(38) a funnel plot compared ward level performances against their average sample sizes. Third, performances were assessed over the 12 month period using run charts.(39) Monthly ward level data were compared with the monthly regional average, and run charts were visually assessed to identify consistent outperformers. Statistical Process Control (SPC) methods such as run charts

1
2
3 and funnel plots are increasingly being promoted to assess variation within healthcare.(40, 41)
4 Performance variations exist within any stable system and SPC methods can be used to distinguish
5 between variation that occurs by chance (i.e. noise in the system) and variation that has an
6 assignable cause (i.e. a signal of positive deviance).(42)
7
8
9

10 **Identification of wards**

11
12
13 Analyses were compared to identify a distinct group of positively deviant elderly medical wards.
14 Individual harms data were also assessed to ensure that positive deviants performed well across all
15 measures in the ST's harm-free care composite. In preparation for phase 2 of this study, comparison
16 wards with slightly above average ST performances were identified. As is the case for many sources
17 of publicly available data, it was not possible to conduct any case-mix adjustments on the ST data.
18 Consequently, comparator wards were matched to positive deviants on three key variables to
19 increase homogeneity within the sample: patient gender – mixed, female, or male; NHS Trust type -
20 teaching and/or foundation trusts; and a routinely collected measure of deprivation.(43) To ensure
21 that positive deviants did not simply care for younger, and thus, comparatively more healthy
22 patients,(33, 34) administrative average patient age data were analysed post-hoc.
23
24
25
26
27
28
29
30
31
32

33 **PHASE 1 RESULTS**

34
35
36 Data were analysed for 34 elderly medical wards clustered within 13 NHS Trusts. Two wards with
37 over 50% missing data were excluded. Average harm-free care performances ranged from 70.56% to
38 92.68% (supplementary file 2 presents all ward rankings). Wards 7, 4, 17, 31, 36, and 29 were the
39 only wards to outperform their respective Trusts on the ST harm-free care data (Figure 1). The first
40 five of these also ranked the highest in the sample while ward 29 ranked eighth. Visual assessment
41 of the run charts (Figure 2) indicated that wards 7, 17, 31, 36, 4, and 15 consistently outperformed
42 the regional average over 12 months, with greater certainty held for those wards listed first.
43 Although none of the wards exceeded the funnel plot's three standard error control limits, wards 7,
44 4, 17, 31, and 36 exceeded them at two standard errors (Figure 3).
45
46
47
48
49
50

51 Using a combination of these four analyses, five wards (7, 4, 17, 31, and 36) were identified to
52 form a distinct group of positive deviants. These wards demonstrated the best performances
53 (rankings); out-performed their respective NHS Trusts (scatterplot); consistently outperformed over
54 12 months (run charts); and their performance variation was attributable to more than chance alone
55
56
57
58
59
60

1
2
3 (funnel plot). They also performed well – around/above average – for each individual ST harm
4 (supplementary file 2). Wards 29 and 15, which were identified through the scatterplot and run
5 charts respectively, did not exceed the funnel plot control limits and so were not deemed to be
6 positively deviant.
7
8

9
10 In total, five matched comparison wards with slightly above average ST harm-free care
11 performances were also identified. Independent samples t-tests indicated that positively deviant
12 wards (M=91.33,SD=0.92) significantly differed from comparators (M=87.46,SD=1.31,
13 t(8)=5.42,p=0.001) and all other wards in the region (M=83.85,SD=4.57, t(32)=3.61,p=0.01) for
14 average ST harm-free care performance. Supplementary file 3 presents the key characteristics of the
15 positively deviant and comparison wards. Although it was not sufficiently powered due to the small
16 sample size (n=9), a post hoc analysis of administrative data for average patient age suggests that
17 positively deviant wards (M=85.1,SD=2.11) did not care for younger, and thus more healthy patients,
18 than the comparison wards (M=84.92,SD=1.42, t(7)=0.15,p = 0.88).
19
20
21
22
23
24
25
26
27

28 **PHASE 2 METHODS**

29 **Participants and recruitment**

30
31
32
33
34 The positively deviant and comparison wards identified during phase 1 were invited to participate
35 in phase 2 of the study. One positively deviant ward was unable to take part. Up to 20 patients and a
36 minimum of 50% of the multidisciplinary ward team were recruited opportunistically to complete
37 surveys assessing their perceptions of safety on the ward. Eligible patients were 65 years or older,
38 were deemed to have capacity, and were considered to be physically well enough. Staff could hold
39 any job role and be of any professional grade. The patient sample size of 20 was determined by
40 previous research; recruiting beyond 20 participants only minimally narrows the confidence intervals
41 for the main measure in the patient survey.(44, 45) A 30-50% response rate for the staff survey has
42 also previously been reported.(46) All recruitment was conducted between February and August
43 2015 (due to publication lag of ST data and the time taken to gain ethical approvals and NHS
44 permissions). Double blinding was retained.
45
46
47
48
49
50
51
52
53
54

55 **Data collection tools**

Patient survey

Patients completed the Patient Measure of Safety (PMOS) which gathers feedback from hospitalised patients about the safety of their care and assesses perceptions about factors that contribute to safety. A total of 44 items are scored on 5-point Likert scales ranging from strongly agree to strongly disagree. These items measure nine safety domains: communication and team working; organisation and care planning; access to resources; ward type and layout; information flow; staff roles and responsibilities; staff training; equipment (design and functioning); and delays. A stand-alone item measures dignity and respect. The PMOS has been validated, is reliable, and considered acceptable to patients.(30, 47) Patients also completed the NHS Friends and Family Test (FFT) – a single item measure of patient experience used in the UK,(48) and three items that had previously been a part of the NHS Commissioning for Quality and Innovation (CQUIN) payment framework.(49)

Staff survey

Multidisciplinary staff completed the Patient Safety Grade (PSG) which asks them to grade their ward on overall safety using a 5-point Likert scale (excellent to failing). The PSG is one of four outcomes within the validated Hospital Survey on Patient Safety Culture (HSOPSC) where non-required outcomes can be removed.(46, 50, 51) This single item was used to maximise response rates.(52) The patient and staff surveys were published with the study protocol.(35)

Procedure

Eligible patients were identified by clinical members of the ward team. Researchers provided written and verbal information about the study and patients gave written informed consent. The patient survey was completed electronically and patients chose whether to do this independently or with the researcher's support. Where support was requested (e.g. due to frailty) the researcher simply read the questions and recorded the patient's answers for them.

Paper copies of the staff surveys were distributed by the researcher and ward managers to multidisciplinary staff at convenient times (e.g. clinical handovers). Staff placed their completed surveys into a 'drop box' which was stored securely on the ward. Staff participation was incentivised by a prize draw (£20 gift voucher per ward).

Analyses

Blinding was removed. Items within the PMOS and CQUIN measures were aggregated to create an overall PMOS and overall CQUIN score for each individual patient. Individual staff and patient level data were then aggregated to ward level by calculating an average ward level score for all measures – the PMOS, FFT, CQUIN and PSG. To assess whether staff and patient perceptions of safety supported the identification of positive deviants, wards were ranked and z-scores were compared in a scatterplot. The small sample size (n=nine wards) and dichotomised performance groups prevented statistical analysis.

PHASE 2 RESULTS

Data were collected from 188 patients and 161 multidisciplinary staff, clustered within nine participating elderly medical wards (supplementary file 4 reports all recruitment data). On average patients were 84.53 years old (SD=5.45), and staff were predominantly nurses or support workers. Table 1 reports the ward level descriptives for all of the measures. (Minimum and maximum values are presented in supplementary file 4.) Positively deviant wards performed better than comparators across all four measures, although differences between the groups were small.

Table 1 Ward level descriptive statistics for all staff and patient survey measures

		ST Harm-free care (Phase 1) Mean %	Patient Measure of Safety Mean (SD) ^a	Friends and Family Test Mean (SD) ^a	CQUIN Mean (SD) ^b	Patient Safety Grade Mean (SD) ^a
Positively deviant wards	Ward 1	90.14	4.33 (.45)	4.71 (.56)	2.48 (.45)	4.29 (.56)
	Ward 3	92.68	4.21 (.34)	4.55 (.67)	2.58 (.47)	4.21 (.70)
	Ward 6	91.48	3.94 (.37)	4.14 (1.15)	2.45 (.32)	4.09 (.54)
	Ward 10	90.97	4.52 (.26)	4.65 (.49)	2.53 (.48)	4.13 (.78)
Comparison wards	Ward 2	88.48	4.11 (.53)	4.26 (.75)	2.25 (.50)	3.50 (1.15)
	Ward 4	87.72	4.09 (.39)	4.26 (1.00)	2.43 (.45)	4.07 (.48)
	Ward 5	85.17	3.96 (.39)	4.15 (1.23)	2.18 (.58)	4.05 (.52)
	Ward 8	87.90	4.51 (.27)	4.75 (.44)	2.48 (.33)	3.69 (.79)
	Ward 9	88.01	4.30 (.36)	4.46 (.88)	2.50 (.36)	3.29 (1.16)
Average	PD group	91.33	4.24 (.41)	4.51 (.78)	2.51 (.43)	4.18 (.67)
	Comp group	87.46	4.20 (.43)	4.38 (.92)	2.38 (.46)	3.71 (.91)

Abbreviations: ST = Safety Thermometer; SD = Standard Deviation; CQUIN = Commissioning for Quality and Innovation; PD = Positive Deviant; Comp = Comparison

^a Measured on a 0-5 Likert scale; ^b Measured on a 0-3 Likert scale. Higher scores represent safer perceptions of patient care on all measures

Ranked performances (Table 2) highlight that staff on positively deviant wards perceived care to be safer than staff on comparison wards as measured by the PSG. This was also largely true for patients, although their perceptions were less concordant with the ST data. Positively deviant ward 6 displayed the lowest PMOS and FFT scores, and two comparators (wards 8 and 9) performed better than some positive deviants on certain patient measures. The scatterplot of z-scores (supplementary file 5) compared performances across different normal distributions. Positively deviant wards predominantly performed above the mean and, as a group, generally performed better on all measures than the comparison wards supporting their identification using ST data.

Table 2: A visual representation of how positively deviant and comparison wards, that were identified using routinely collected ST data, ranked on patient and staff perceptions of safety

Rank	Routinely collected ST Harm-free care (Phase 1)	Patient and staff perceptions of safety (Phase 2)			
		Patient Measure of Safety	Friends and Family Test	CQUIN	Patient Safety Grade
1 (high)	Ward 3	Ward 10	Ward 8	Ward 3	Ward 1
2	Ward 6	Ward 8	Ward 1	Ward 10	Ward 3
3	Ward 10	Ward 1	Ward 10	Ward 9	Ward 10
4	Ward 1	Ward 9	Ward 3	Ward 1 and Ward 8	Ward 6
5	Ward 2	Ward 3	Ward 9	Ward 8	Ward 4
6	Ward 9	Ward 2	Ward 4 and Ward 2	Ward 6	Ward 5
7	Ward 8	Ward 4	Ward 2	Ward 4	Ward 8
8	Ward 4	Ward 5	Ward 5	Ward 2	Ward 2
9 (low)	Ward 5	Ward 6	Ward 6	Ward 5	Ward 9

Positively deviant wards (as identified by the routinely collected ST data) are shaded in colour.

Comparison wards are represented in white. Higher ranks represent safer perceptions of patient care on all measures.

Abbreviations: ST = Safety Thermometer; CQUIN = Commissioning for Quality and Innovation;

DISCUSSION

This study developed and critically appraised a method for conducting stage 1 of the Bradley et al.(7) positive deviance framework – identifying positive deviants. More specifically, we applied a rigorous and robust analysis to the ST harm-free care data to identify positively deviant wards that demonstrated sustained exceptional levels of safety, and established the extent to which survey based staff and patient perceptions of safety aligned with this ST harm-free data. Previous applications of the approach have typically identified positively deviant individuals or organisations who demonstrate exceptional performance on narrow processes or outcomes of care.(9) However, performance variation also exists between wards/units.(18) Although there is merit to focusing on specific aspects of care, this can divert attention away from other important aspects of safety reducing opportunities for wider improvement, for example, by implementing cultural changes that improve several different outcomes (e.g. improving multi-disciplinary teamwork).

During phase 1 of this study, a robust yet pragmatic analysis successfully identified a discreet group of five statistically different positively deviant elderly medical wards with exceptionally high ST harm-free care performances. Although these wards did rank top of the region, rankings alone did not differentiate between positively deviant wards and those that performed well. This study therefore advanced the previous methods that have been used to identify positive deviants by identifying a distinct and statistically different group of wards that not only ranked best within the region but also outperformed their NHS Trust (organisation), performed consistently over 12 months, and demonstrated success beyond what would be expected through chance. SPC methods such as funnel plots and run charts are increasingly promoted for assessing performance variation within healthcare.(40, 41) They combine statistical rigour with sensitive measurement to differentiate between variation that is expected by chance and variation that has an assignable cause.(42, 53) The methods are also considered to be relatively intuitive and pragmatic enough for use by improvers on the frontline.(41)

However, the extent to which wards truly demonstrated exceptional performance can be questioned. Wards did not exceed the funnel plot's three standard error control limits and so their exceptional, outlier status was limited. Furthermore, although positively deviant wards differed statistically from others, the minimal performance differences between them and the 'next best' wards highlight the importance of considering clinical significance – that is, whether differences meaningfully affect patient treatment.(54) Positive deviants are supposed to demonstrate 'exceptionally' high performance,(7) but there is little consensus in the literature about how to

1
2
3 differentiate between high performance and positively deviant performances. Control limits are
4 considered to be conservative and can be made more or less stringent depending on the
5 context.(39) In the absence of extreme outliers, it may still be possible to conduct stage 2 of the
6 Bradley et al framework(7) in order to generate useful learning from 'positive deviants' who simply
7 perform *well*. However, regardless of whether one learns from outlying positive deviants or not, it is
8 important to note that there is currently a lack of evidence on the methods that should be used to a)
9 effectively uncover positively deviant success strategies, and b) disseminate them to others (stages
10 2 and 4 of the Bradley framework).(9, 10)
11
12
13
14
15

16 During phase 2 of this study, staff and patient perceptions of safety did, in the main, corroborate
17 the routinely collected ST data, providing tentative support for the methods used to identify
18 positively deviant wards. Patients on positively deviant wards, though, did not uniformly perceive
19 their care to be safer than those on comparator wards. This could be explained by wards adopting
20 different approaches to delivering safe patient care; for example, if staff emphasise guideline
21 compliance over patient centred care this may influence patient perceptions of safety. Furthermore,
22 patients may have unique perspectives of safety, which perhaps encompass the culture of a ward
23 rather than just the outcomes that are reflected in the routinely collected data.(55, 56) The lack of
24 agreement between the FFT, CQUIN, and ST measures may also have arisen by measuring two
25 associated but distinct quality domains – patient experience versus safety.(57) Fundamentally
26 though, inconsistent patient perceptions highlight that different positively deviant wards may have
27 been identified had a different broad measure of safety (other than the ST) been used.
28
29
30
31
32
33
34
35

36 In addition to these considerations, there are various overarching considerations and challenges
37 that are faced when applying stage 1 of the Bradley et al. framework(7) to identify positive deviants.
38 First, there are few sources of routinely collected data within the UK's NHS or further afield that
39 broadly measure safety and are publicly available at ward level.(58) This makes it difficult to adopt a
40 pragmatic approach and identify positive deviants across different organisations. Although this study
41 provides tentative support for using ST data within a UK healthcare setting, the harms measured
42 within this tool are particularly pertinent to older people. Consequently this measure may lack
43 relevance to applications of the approach which seek to identify positive deviants in different
44 healthcare settings, for example, paediatric wards or emergency departments.
45
46
47
48
49
50

51 Second, positive deviants are assumed to succeed *despite facing the same constraints as*
52 *others*(6) and so it is critical to identify them from within a homogenous population to ensure that,
53 as far as possible, one is comparing like for like. This study increased homogeneity by defining
54
55
56
57

1
2
3 elderly medical wards as stringently as possible and by sampling matched comparators. However,
4 numerous factors are known to contribute to patient safety incidents,(59) case-mix adjustments are
5 notoriously difficult, (60) and one can never fully control for all confounding variables when
6 identifying positive deviants. The complexity of healthcare means that all ward teams deliver patient
7 care within their own unique organisational contexts and so it will never be possible to sample a fully
8 homogenous population or identify positive deviants that face *exactly* the same constraints as
9 others. This challenge is likely to be especially pertinent when comparing performance on broad
10 outcomes of safety, when using publicly available data, and when adopting a pragmatic approach.
11
12
13
14
15

16 Third, although the problems associated with routinely collected data are well documented,(22-
17 24) there are also wider implications of using routine data to identify positive deviants. Performance
18 variation can arise because measurement is conducted in a social context – staff do not make
19 decisions about the same things nor do they decide things in the same way.(61) This is problematic
20 when the positive deviance approach seeks to compare performances across several different
21 healthcare providers. Furthermore, healthcare organisations retrospectively measure the absence
22 rather than presence of safety(62) and measurement and monitoring systems say nothing about
23 how safe patient care currently is or how safe it will be in the future.(63) This compounds the ability
24 to accurately identify positive deviants and thus the ability to reliably conduct subsequent stages of
25 the approach.
26
27
28
29
30
31
32

33 **Study limitations**

34
35
36 Various study limitations have already been highlighted including measuring statistical rather
37 than clinical differences between positively deviant and comparison wards. Due to resource
38 constraints, it was also not possible to assess staff and patient perceptions of safety across all 34
39 wards that were sampled during phase 1 of the study. The resulting small and dichotomised sample
40 during phase 2 meant that the associations between staff and patient perceptions of safety and the
41 ST data could not be assessed statistically. Furthermore, some of the differences in performance
42 between positively deviant and comparison wards were small. Had the comparison group comprised
43 negative deviants (the worst performers) rather than above average performers, the differences
44 between the two groups on each of the quantitative measures may have been more stark. Many
45 previous applications of stage 1 of the positive deviance approach have identified positive and
46 negative deviants.(35) However, this specific comparison group was chosen with the wider
47 application of positive deviance in mind,(35) so that when we qualitatively explore *how* positive
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 deviants succeed (stage 2 of the framework(7)) we can strive to distinguish between *exceptional* and
4 good performances, not just explore how teams differ from the worst in the population.
5
6

7 Finally, as with many routinely collected measures of quality and safety, the reliability and validity
8 of the ST has been questioned. ST data are collected opportunistically at a single monthly time
9 point,(27) harm definitions are subject to interpretation,(64, 65) and data collection was previously
10 incentivised. However, the ST is used to measure performance in most acute NHS Trusts and it is the
11 only routinely collected, broad measure of safety that is publicly available at ward level in the UK.
12 Additional ST tools, such as the Medications Safety Thermometer, have been developed(66, 67) and
13 so, if these can also be used to identify positively deviant wards, then the methods tested in this
14 study could have greater impact across the NHS.
15
16
17
18
19
20

21 **Conclusions**

22
23 This study has shown that a distinct group of positively deviant wards that perform exceptionally
24 well on a routinely collected, broad measure of safety can be identified using a robust yet pragmatic
25 method, and that staff and patient perceptions of safety do, in the main, support their identification.
26 It has highlighted the challenges faced when selecting a source of routinely collected data that
27 provides a valid and reliable measure at the appropriate level in order to facilitate performance
28 comparisons across wards or units in several organisations. Many of these challenges are applicable
29 to a variety of different settings and applications of the approach and so this study may provide
30 generalisable guidance on the methods that can be used to effectively apply stage 1 of the Bradley
31 et al framework(7) and identify positive deviants.
32
33
34
35
36
37
38
39
40
41

42 **Figures:**

43
44 Figure 1: Scatterplot comparing average ST harm-free care performances at ward and NHS Trust
45 (organisation) levels.
46
47

48 Figure 2: Run charts comparing ward and regional level monthly harm-free care performance across
49 a 12 month period.
50
51

52
53 Each square represents an individual ward within the population. Wards are numbered
54 consecutively according to their pseudonym (from top left to bottom right across the rows).
55
56
57

1
2
3 Black lines represent a ward's monthly harm-free care performance. Grey lines represent the
4 region's average monthly performance.
5
6

7 Figure 3: Funnel plot of average ST harm-free care performance and average sample size.
8
9
10

11
12 **Funding:** The Health Foundation - PhD in Improvement Science.
13
14

15 **Competing interests:** None declared.
16
17

18 **Acknowledgements:** The authors express their sincere gratitude to all of the ward staff and patients
19 who participated in and supported this study. Thank you to Carolyn Reynolds and Alex Howat who
20 helped collect some of the data. RB, RL and IK are members of the NIHR Collaboration for Leadership
21 in Applied Health Research and Care, Yorkshire and Humber – Evidence Based Transformation
22 Theme. The research was supported by the NIHR CLAHRC Yorkshire and Humber. [www.clahrc-
24 yh.nihr.ac.uk](http://www.clahrc-
23 yh.nihr.ac.uk). The views expressed are those of the author(s), and not necessarily those of the NHS,
25 the NIHR or the Department of Health.
26
27
28
29

30 **Author contributions:** RB, RL, NT and IK developed the idea for the study. RB, MM and VP conducted
31 the analyses to identify positive deviants. RB collected and analysed the primary data to assess staff
32 and patient perceptions of safety. RB drafted the manuscript. All authors provided comments and
33 approved the final version.
34
35
36

37 **Patient consent:** Obtained.
38
39

40 **Ethics approval:** South East Scotland Research Ethics Committee (ref 14/SS/1085). NHS permissions
41 were gained from all five NHS trusts involved.
42
43

44 **Data sharing statement:** Anonymised data are available on request.
45
46

47 References

48
49

- 50 1. Hogan H, Healey F, Neale G, et al. Preventable deaths due to problems in care in English
51 acute hospitals: a retrospective case record review study. *BMJ Qual Saf*. 2012:Published Online First:
52 07 July 2012.
- 53 2. Landrigan CP, Parry GJ, Bones CB, et al. Temporal trends in rates of patient harm resulting
54 from medical care. *New England Journal of Medicine*. 2010;363(22):2124-34.
- 55 3. Hollnagel E, Braithwaite J, Wears RL. Resilient health care. UK: Ashgate; 2013.
56
57

4. Vincent C, Amalberti R. Safer healthcare: Strategies for the real world.: Springer Open; 2016. Available from: <http://link.springer.com/book/10.1007%2F978-3-319-25559-0>.
5. Lawton R, Taylor N, Clay-Williams R, et al. Positive deviance: a different approach to achieving patient safety. *BMJ Quality & Safety*. 2014:Published online first 21 July 2014. PubMed PMID: 25049424. Epub 2014/07/23. Eng.
6. Marsh DR, Schroeder DG, Dearden KA, et al. The power of positive deviance. *BMJ*. 2004;329(7475):1177-9. PubMed PMID: WOS:000225169600033.
7. Bradley EH, Curry LA, Ramanadhan S, et al. Research in action: using positive deviance to improve quality of health care. *Implementation Science*. 2009;4:25.
8. Dixon-Woods M, Martin GP. Does quality improvement improve quality? *Future Hospital Journal*. 2016;3(3):191-4.
9. Baxter R, Taylor N, Kellar I, et al. What methods are used to apply positive deviance within healthcare organisations? A systematic review. *BMJ Quality and Safety*. 2015:Published online first 20th November.
10. Rose AJ, McCullough MB. A practical guide to using the positive deviance method in health services research. *Health Serv Res*. 2016;Published online 28th June PubMed PMID: 27349472.
11. Rose AJ, Petrakis BA, Callahan P, et al. Organizational characteristics of high- and low-performing anticoagulation clinics in the Veterans Health Administration. *Health Serv Res*. 2012;47(4):1541-60. PubMed PMID: 22299722. Pubmed Central PMCID: 3401398.
12. Gabbay RA, Friedberg MW, Miller-Day M, et al. A positive deviance approach to understanding key features to improving diabetes care in the medical home. *Ann Fam Med*. 2013;1(Suppl 1):S99-107. PubMed PMID: 23690393. Pubmed Central PMCID: PMC3707253 [Available on 11/01/13].
13. Austin JM, Jha AK, Romano PS, et al. National hospital ratings systems share few common scores and may generate confusion instead of clarity. *Health Affairs*. 2015;34(3):423-30. PubMed PMID: 25732492.
14. Healthcare Association of New York State. HANYS' report on report cards: understanding publicly reported hospital quality measures. US, New York: HANYS Quality Institute, 2013.
15. Rothberg MB, Morsi E, Benjamin EM, et al. Choosing the best hospital: the limitations of public quality reporting. *Health Aff (Millwood)*. 2008;27(6):1680-7. PubMed PMID: 18997226. Epub 2008/11/11. eng.
16. Pannick S, Wachter RM, Vincent C, et al. Rethinking medical ward quality. *BMJ*. 2016;355(i5417).
17. Rose JS, Thomas CS, Tersigni AR, et al. A leadership framework for culture change in health care. *The Joint Commission Journal on Quality and Patient Safety*. 2006;32(8):433-42.
18. Schwendimann R, Zimmermann N, Küng K, et al. Variation in safety culture dimensions within and between US and Swiss Hospital Units: an exploratory study. *BMJ Qual Saf*. 2013;22:32-41.
19. Nelson EC, Batalden PB, Huber TP, et al. Microsystems in health care: Part 1. Learning from high-performing front-line clinical units. *Jt Comm J Qual Patient Saf*. 2001;28(9):472-93.
20. Reason J. Understanding adverse events: human factors. *Quality in Health Care*. 1995;4(2):80-9.
21. Vincent C, Taylor-Adams S, Stanhope N. Framework for analysing risk and safety in clinical medicine. *BMJ*. 1998;316(7138):1154-7.
22. Lilford R, Mohammed MA, Spiegelhalter D, et al. Use and misuse of process and outcome data in managing performance of acute medical care: avoiding institutional stigma. *The Lancet*. 2004;363(9415):1147-54.
23. Shaw J, Taylor R, Dix K. Uses and abuses of performance data in healthcare. UK: Dr Forster, 2015.
24. Zhan C, Miller MR. Administrative data based patient safety research: a critical review. *Quality and Safety in Health Care*. 2003;12(Suppl 2):58-63.

- 1
2
3 25. Digital N. Hospital Episode Statistics 2017 [cited 2017 13th October]. Available from:
4 <http://content.digital.nhs.uk/hes>.
- 5 26. Institute TP. The NHS Staff Survey 2017 [cited 2017 13th October 2017]. Available from:
6 <http://www.nhsstaffsurveys.com/Page/1056/Home/NHS-Staff-Survey-2017/>.
- 7 27. Power M, Stewart K, Brotherton A. What is the NHS Safety Thermometer? *Clinical Risk*.
8 2012;18(5):163-9. PubMed PMID: 23136533. Pubmed Central PMCID: 3484316. Epub 2012/11/09.
9 Eng.
- 10 28. Dixon-Woods M, McNicol S, Martin G. Ten challenges in improving quality in healthcare:
11 lessons from the Health Foundation's programme evaluations and relevant literature. *BMJ Qual Saf*.
12 2012:Published Online First: 28 April 2012.
- 13 29. Scott T, Mannion R, Davies H, et al. The quantitative measurement of organisational culture
14 in healthcare: a review of the available instruments. *Health Services Research* 2003;38(3):923-44.
- 15 30. Giles SJ, Lawton RJ, Din I, et al. Developing a patient measure of safety (PMOS). *BMJ Quality
16 and Safety*. 2013;22:554-62.
- 17 31. Weingart SN, Pagovich O, Sands DZ, et al. What can hospitalized patients tell us about
18 adverse events? Learning from patient-reported incidents. *Journal of General Internal Medicine*.
19 2005;20(9):830-6.
- 20 32. Weissman JS, Schneider EC, Weingart PN, et al. Comparing patient-reported hospital adverse
21 events with medical record review: do patients know something that hospitals do not? *Annals of
22 Internal Medicine*. 2008;149(100-108).
- 23 33. Sari AB, Cracknell A, Sheldon TA. Incidence, preventability and consequences of adverse
24 events in older people: Results of a retrospective case-note review. *Age and Ageing*. 2008;37(3):265-
25 9. PubMed PMID: 18332053.
- 26 34. Thomas EJ, Brennan TA. Incidence and types of preventable adverse events in elderly
27 patients: population based review of medical records. *BMJ*. 2000;320(7237):741-4.
- 28 35. Baxter R, Taylor N, Kellar I, et al. Learning from positively deviant wards to improve patient
29 safety: an observational study protocol. *BMJ Open*. 2015;5:e009650.
- 30 36. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational
31 Studies in Epidemiology (STROBE)statement: guidelines for reporting observational studies. *The
32 Lancet*. 2007;370(9596):1453-7.
- 33 37. Health and Social Care Information Centre. NHS Safety Thermometer. 2014 [cited 2014 29th
34 April]. Available from: <http://www.hscic.gov.uk/thermometer>.
- 35 38. Mayer EK, Bottle A, Rao C, et al. Funnel plots and their emerging application in surgery. *Ann
36 Surg*. 2009;249(3):376-83. PubMed PMID: 19247021.
- 37 39. Perla RJ, Provost LP, Murray SK. The run chart: a simple analytical tool for learning from
38 variation in healthcare processes. *BMJ Qual Saf*. 2011;20(1):46-51. PubMed PMID: 21228075.
- 39 40. Mountford J, Wakefield D. From stoplight reports to time series: equipping boards and
40 leadership teams to drive better decisions. *BMJ Quality & Safety*. 2016;Published online first: 7th
41 March.
- 42 41. Schmidtke KA, Poots AJ, Carpio J, et al. Considering chance in quality and safety performance
43 measures: an analysis of performance reports by boards in English NHS trusts. *BMJ Quality & Safety*.
44 2016:Published online first 31 March.
- 45 42. Mohammed MA. Using statistical process control to improve the quality of health care.
46 *Quality and Safety in Health Care*. 2004;13(4):243-5.
- 47 43. Office of the Deputy Prime Minister. The English Indices of Deprivation 2004: Summary
48 (revised). UK: Office of the Deputy Prime Minister; 2004.
- 49 44. Lawton R, O'Hara JK, Sheard L, et al. Can staff and patient perspectives on hospital safety
50 predict harm-free care? An analysis of staff and patient survey data and routinely collected
51 outcomes. . *BMJ Quality and Safety*. 2015;24(6):369-76.
52
53
54
55
56
57
58
59
60

- 1
2
3 45. Sheard L, O'Hara J, Armitage G, et al. Evaluating the PRASE patient safety intervention - a
4 multi-centre, cluster trial with a qualitative process evaluation: study protocol for a randomised
5 controlled trial. *Trials*. 2014;15(1):2282.
- 6 46. Sorra J, Nieva V. Hospital survey on patient safety culture. USA, Rockville: Westat, AHRQ
7 Publication No. 04-0041, 2004.
- 8 47. McEachan RRC, Lawton RJ, O'Hara JK, et al. Developing a reliable and valid patient measure
9 of safety in hospitals (PMOS): a validation study. *BMJ Quality and Safety*. 2014:Published online first
10 24 December 2013.
- 11 48. NHS England. Friends and family test. 2014 [cited 2014 29th April]. Available from:
12 <http://www.england.nhs.uk/statistics/statistical-work-areas/friends-and-family-test/>.
- 13 49. Department of Health. Using the Commissioning for Quality and Innovation (CQUIN)
14 payment framework – Guidance on national goals for 2011/12. 2010 [cited 2016 3rd April].
15 Available from:
16 [https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/215049/dh_13385](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/215049/dh_13385_9.pdf)
17 [9.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/215049/dh_13385_9.pdf).
- 18 50. Sarac C, Flin R, Mearns K, et al. Hospital survey on patient safety culture: psychometric
19 analysis on a Scottish sample. *BMJ Quality and Safety*. 2011;20(10):842-8. PubMed PMID: 21690247.
20 Epub 2011/06/22. eng.
- 21 51. Sorra J, Dyer N. Multilevel psychometric properties of the AHRQ hospital survey on patient
22 safety culture. *BMC Health Service Research*. 2010;10:199. PubMed PMID: 20615247. Pubmed
23 Central PMCID: 2912897. Epub 2010/07/10. eng.
- 24 52. Edwards P, Roberts I, Clarke M, et al. Increasing response rates to postal questionnaires:
25 systematic review. *BMJ*. 2002;324(7347):1183.
- 26 53. Benneyan JC, Lloyd RC, Plsek PE. Statistical process control as a tool for research and
27 healthcare improvement. *Qual Saf Health Care*. 2003;12:458-64.
- 28 54. Oberst MT. Clinical versus statistical significance. *Cancer Nursing*. 1982;5(6):475-6.
- 29 55. Ocloo J, Matthews R. From tokenism to empowerment: progressing patient and public
30 involvement in healthcare improvement. *BMJ Quality & Safety*. 2016:Published online first 18th
31 March.
- 32 56. Weingart SN, Pagovich O, Sands DZ, et al. What can hospitalized patients tell us about
33 adverse events? Learning from patient-reported incidents. *J Gen Intern Med*. 2005;20(9):830-6.
34 PubMed PMID: 16117751. Pubmed Central PMCID: 1490203. Epub 2005/08/25. eng.
- 35 57. Doyle C, Lennox L, Bell D. A systematic review of evidence on the links between patient
36 experience and clinical safety and effectiveness. *BMJ open*. 2013;3(1):e001570.
- 37 58. Pannick SAJ, Wachter RM, Vincent C, et al. Rethinking medical ward quality. *BMJ*.
38 2016;355:i5417.
- 39 59. Lawton R, McEachan RR, Giles SJ, et al. Development of an evidence-based framework of
40 factors contributing to patient safety incidents in hospital settings: a systematic review. *BMJ Quality*
41 *& Safety*. 2012:Published Online First: 15 March 2012.
- 42 60. Shahian DM, Normand SLT. What is a performance outlier? *BMJ Qual Saf*. 2015;24:95-9.
- 43 61. Dixon-Woods M, Leslie M, Bion J, et al. What counts? An ethnographic study of infection
44 data reported to a patient safety program. *The Milbank Quarterly*. 2012;90(3):548-91.
- 45 62. Hollnagel E, Wears RL, Braithwaite J. From Safety-I to Safety-II: A white paper. Published
46 simultaneously by the University of Southern Denmark, University of Florida, USA, and Macquarie
47 University, Australia: The Resilient Health Care Net, 2015.
- 48 63. Vincent C, Burnett S, Carthey J. The measurement and monitoring of safety. UK: The Health
49 Foundation, 2013.
- 50 64. Buckley C, Cooney K, Sills E, et al. Implementing the Safety Thermometer tool in one NHS
51 trust. *British Journal of Nursing*. 2014;23(5):268-72.
- 52 65. Power M, Fogarty M, Madsen J, et al. Learning from the design and development of the NHS
53 Safety Thermometer. *International Journal for Quality in Health Care*. 2014;26(3):287-97.
- 54
55
56
57
58
59
60

- 1
2
3 66. NHS Quality Observatory. NHS Safety Thermometer 2013 [cited 2017 12th May]. Available
4 from:
5 https://www.safetythermometer.nhs.uk/index.php?option=com_content&view=article&id=1&Itemid=101.
6
7 67. Rostami P, Power M, Harrison A, et al. Learning from the design, development and
8 implementation of the Medication Safety Thermometer. *International Journal for Quality in Health*
9 *Care*. 2016;29(2):301-9.
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

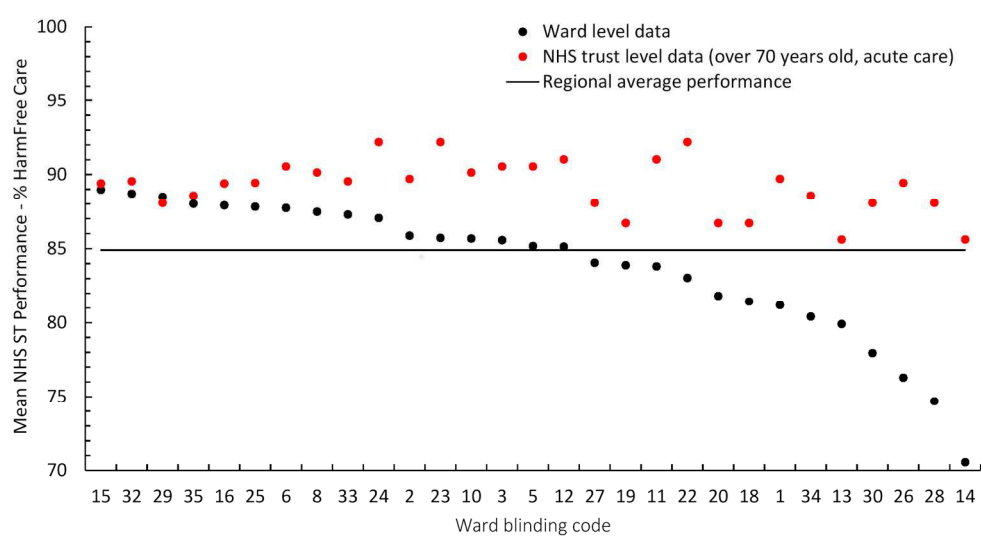


Figure 1: Scatterplot comparing average ST harm-free care performances at ward and NHS Trust (organisation) levels.

173x105mm (300 x 300 DPI)

Review only

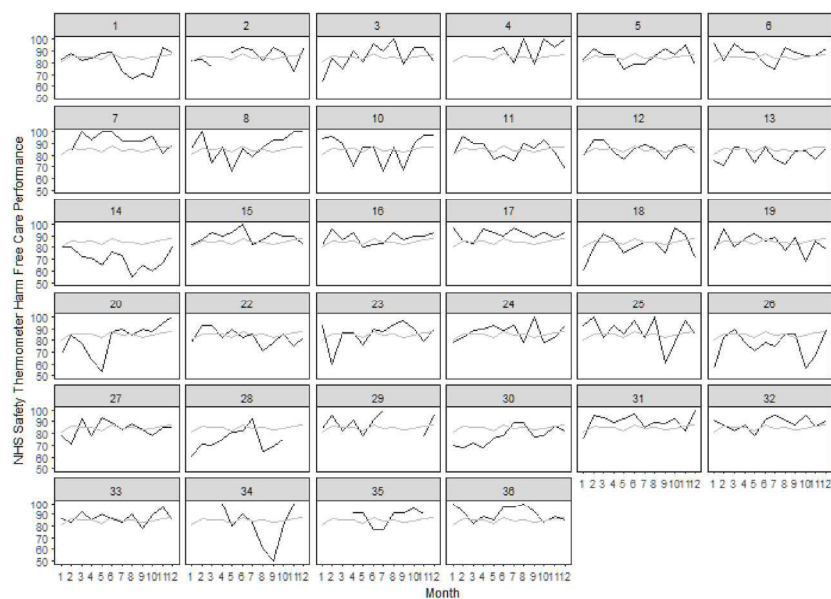


Figure 2: Run charts comparing ward and regional level monthly harm-free care performance across a 12 month period.
 Each square represents an individual ward within the population. Wards are numbered consecutively according to their pseudonyms (from top left to bottom right across the rows).
 Black lines represent a ward's monthly harm-free care performance. Grey lines represent the region's average monthly performance.

173x123mm (300 x 300 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

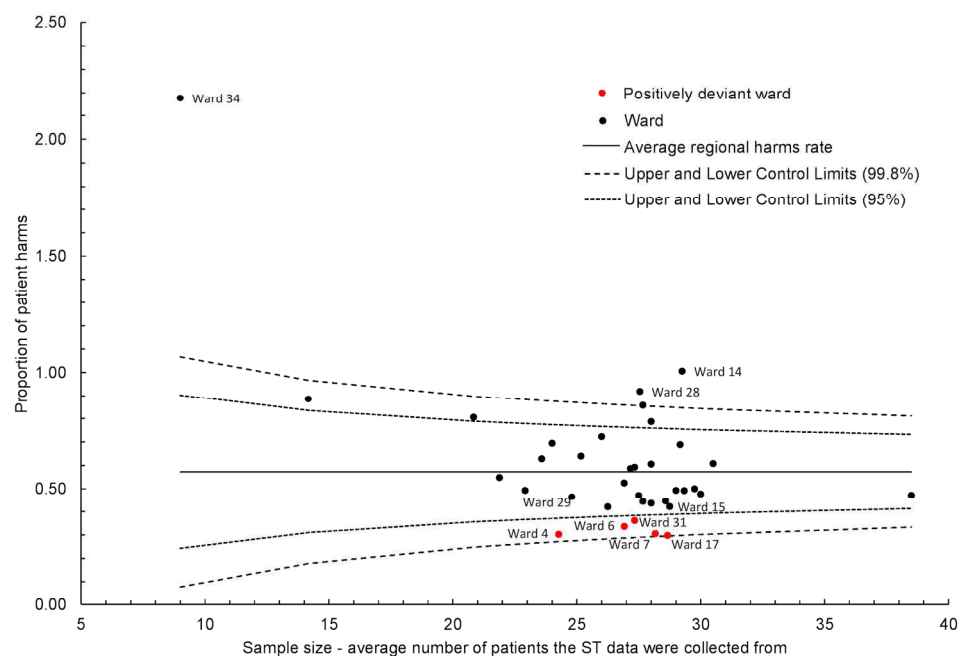


Figure 3: Funnel plot of average ST harm-free care performance and average sample size.

173x121mm (300 x 300 DPI)

new only

Supplementary File 1

STROBE Statement—Checklist of items that should be included in reports of *cross-sectional studies*

	Item No	Recommendation	Revised Manuscript Page number
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	1
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4-6
Objectives	3	State specific objectives, including any prespecified hypotheses	6
Methods			
Study design	4	Present key elements of study design early in the paper	6
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	7 and 9/10
Participants	6	(a) Give the eligibility criteria, and the sources and methods of selection of participants	7/8 and 9/10
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	7, 10 and published protocol
Data sources/measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	7 and 10
Bias	9	Describe any efforts to address potential sources of bias	7
Study size	10	Explain how the study size was arrived at	9
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	8 and 11
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	8 and 11
		(b) Describe any methods used to examine subgroups and interactions	n/a
		(c) Explain how missing data were addressed	8
		(d) If applicable, describe analytical methods taking account of sampling strategy	n/a
		(e) Describe any sensitivity analyses	n/a
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	10 and Supp file 4
		(b) Give reasons for non-participation at each stage	n/a
		(c) Consider use of a flow diagram	n/a
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	11 and Supp File 4

		(b) Indicate number of participants with missing data for each variable of interest	n/a
Outcome data	15*	Report numbers of outcome events or summary measures	12
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	12-13
		(b) Report category boundaries when continuous variables were categorized	n/a
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	n/a
Discussion			
Key results	18	Summarise key results with reference to study objectives	14-16
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	16
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	14-16
Generalisability	21	Discuss the generalisability (external validity) of the study results	14-17
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	18

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at www.strobe-statement.org.

Supplementary file 2

Ward rankings for each of the NHS Safety Thermometer measures

Ranking	Ward blinding code	ST harmfree care performance ^a	New PUs ^b	Falls	New UTI	New VTE
1	4	92.68	1.01	0.36	0.00	1.36
2	7	91.48	0.00	0.74	0.40	0.74
3	17	91.40	0.58	0.60	0.29	0.60
4	36	90.97	2.55	0.32	0.30	0.00
5	31	90.14	2.09	1.53	0.31	0.30
6	15	88.97	1.16	0.00	0.00	0.00
7	32	88.70	2.21	1.09	0.36	2.93
8	29	88.48	2.09	0.79	1.20	1.20
9	35	88.19	1.65	0.82	0.27	1.65
10	16	88.01	1.14	1.16	0.57	2.27
11	25	87.90	1.15	1.75	1.15	1.99
12	6	87.72	1.51	0.60	0.30	1.80
13	8	87.46	1.11	0.00	1.67	0.56
14	33	87.28	1.99	1.19	0.00	1.14
15	24	86.52	0.00	1.31	1.33	1.32
16	2	85.87	0.66	0.69	0.65	1.97
17	23	85.71	0.00	1.13	0.58	0.84
18	10	85.68	3.32	0.83	0.83	0.28
19	3	85.57	2.22	1.42	0.00	0.27
20	5	85.17	1.44	2.11	0.36	0.00
21	12	85.14	3.05	2.25	1.70	0.00
22	27	84.06	0.60	4.57	1.21	0.30
23	19	83.90	1.58	1.52	0.95	1.91
24	11	83.81	3.99	2.20	0.29	0.58
25	22	83.04	1.79	0.60	0.60	0.60
26	20	81.83	4.42	5.98	0.00	0.43
27	18	81.45	2.15	0.79	0.32	1.13
28	1	81.20	2.74	1.62	0.93	1.63
29	34	80.42	2.50	1.25	2.50	3.75
30	13	79.91	3.52	1.14	2.08	0.84
31	30	77.95	0.58	2.09	5.65	0.88
32	26	76.28	0.89	6.32	2.41	3.92
33	28	74.69	4.70	1.03	2.67	3.27
34	14	70.56	8.02	1.13	5.14	2.01
Average		84.90	2.03	1.52	1.08	1.21

* PUs = pressure ulcers; UTI = urinary tract infections, VTE = venous thromboembolism

^a The wards highlighted blue are positive deviants and the wards highlighted green are matched comparison wards.

^b Performances for each individual harm are listed in columns 4-8 (not ranked in order). The top five performers across the region are highlighted yellow.

Supplementary file 3

Key characteristics of the positively deviant and comparison wards

	Ward	ST Harm-free care (%)	Trust number / type	Patient gender ^a	Index of Multiple Deprivation ^b
Positively deviant wards	1	90.14	Trust 1 / Teaching & Foundation	Mixed	More deprived 40-50%
	3	92.68	Trust 2 / Teaching	Female	More deprived 30-40%
	6	91.48	Trust 2 / Teaching	Female	More deprived 30-40%
	7	91.40	Trust 3 / Teaching & Foundation	Mixed	Less deprived 30-40%
	10	90.97	Trust 5 / Foundation	Mixed	More deprived 30-40%
Comparison wards	2	88.48	Trust 1 / Teaching & Foundation	Mixed	More deprived 40-50%
	4	87.72	Trust 2 / Teaching	Female	More deprived 30-40%
	5	85.17	Trust 2 / Teaching	Male	More deprived 30-40%
	8	87.90	Trust 3 / Teaching & Foundation	Mixed	Less deprived 30-40%
	9	88.01	Trust 4 / Foundation	Mixed	More deprived 30-40%

^a One positively deviant ward had to be matched to a comparison ward that cared for patients of the opposite gender.

^b IMD overall rank data (extracted from the 2012/13 Hospital Episode Statistics data) are categorised into deciles. Geographic areas are ranked and then described as falling within the most or least deprived % of England. Categories change in increments of 10% up to the more/least deprived 40-50% of England.

Supplementary file 4: Patient and staff recruitment information

Table 1: Ward level recruitment data for staff and patients

		Patient response rate n (% of those approached) ^a	Staff response rate n (approx. % of the MDT)
Positively deviant wards	T1W1	21 (51)	22 (44)
	T2W3	22 (61)	14 (40)
	T2W6	21 (47)	11 (31)
	T5W10	20 (51)	30 (67)
Comparison wards	T1W2	17 (53)	18 (45)
	T2W4	23 (48)	14 (40)
	T2W5	20 (49)	19 (54)
	T3W8	20 (69)	16 (46)
	T4W9	24 (83)	17 (35)
All wards		188 (55)	161 (45)

^a Response rates include patients who explicitly refused to participate. It does not include: those whom nurses reported to be eligible but were subsequently considered unsuitable; and those who could not be followed up after providing time to consider participation.

Table 2: Characteristics of patient participants

	Patient age mean years (SD)	Ongoing hospital treatment % yes	Time in hospital mean days (SD)	Inpatient frequency mean (SD)
Positively deviant wards	84.49 (5.60)	33%	13.6 (12.87)	2.27 (2.46)
Comparison wards	84.56 (5.36)	46% (3 missing)	15.71 (19.64)	2.71 (3.32)
All wards	84.53 (5.45)	41%	14.75 (16.91)	2.51 (2.97)

Table 3: Professional roles of staff participants

	Nursing %	Healthcare Assistants %	Allied Health Profes %	Doctors %	Other %
Positively deviant wards	46.8	19.5	15.6	6.5	11.7
Comparison wards	36.9	31.0	13.1	1.2	17.8
All wards	41.6	25.5	14.3	3.7	15.0

Table 4: Ward level minimum and maximum values for all patient and staff measures of safety.

		Patient Measure of Safety		Friends and Family Test		CQUIN		Patient Safety Grade	
		Min	Max	Min	Max	Min	Max	Min	Max
Positively deviant wards	Ward 1	2.94	4.91	3	5	1.67	3.00	3	5
	Ward 3	3.59	4.74	3	5	1.67	3.00	3	5
	Ward 6	3.17	4.57	1	5	2.00	3.00	3	5
	Ward 10	3.93	4.91	4	5	1.00	3.00	2	5
Comparison wards	Ward 2	2.95	4.77	2	5	1.33	3.00	1	5
	Ward 4	3.31	4.81	1	5	1.67	3.00	3	5
	Ward 5	2.93	4.71	1	5	1.00	3.00	3	5
	Ward 8	4.00	4.98	4	5	2.00	3.00	2	5
	Ward 9	3.41	4.81	2	5	1.67	3.00	1	5

Supplementary file 5

Scatterplot of z-scores to assess whether other perceptions of safety support the use of ST data to identify positive deviants.

