



OPEN ACCESS

Governing patient safety: lessons learned from a mixed methods evaluation of implementing a ward-level medication safety scorecard in two English NHS hospitals

Angus I G Ramsay,¹ Simon Turner,¹ Gillian Cavell,² C Alice Osborne,³ Rebecca E Thomas,³ Graham Cookson,⁴ Naomi J Fulop¹

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/bmjqs-2012-001730>).

¹Department of Applied Health Research, University College London, London, UK

²Department of Pharmacy, King's College Hospital NHS Foundation Trust, London, UK

³Pharmacy Department, Guy's and St Thomas' NHS Foundation Trust, London, UK

⁴Department of Health Care Management & Policy, University of Surrey, Guildford, UK

Correspondence to

Dr Angus I G Ramsay, Department of Applied Health Research, University College London, 1-19 Torrington Place, London WC1E 7HB, UK; angus.ramsay@ucl.ac.uk

Received 2 December 2012

Revised 9 August 2013

Accepted 13 August 2013

Published Online First

12 September 2013



To cite: Ramsay AIG, Turner S, Cavell G, et al. *BMJ Qual Saf* 2014;**23**:136–146.

ABSTRACT

Background Relatively little is known about how scorecards presenting performance indicators influence medication safety. We evaluated the effects of implementing a ward-level medication safety scorecard piloted in two English NHS hospitals and factors influencing these.

Methods We used a mixed methods, controlled before and after design. At baseline, wards were audited on medication safety indicators; during the 'feedback' phase scorecard results were presented to intervention wards on a weekly basis over 7 weeks. We interviewed 49 staff, including clinicians and managers, about scorecard implementation.

Results At baseline, 18.7% of patients (total n=630) had incomplete allergy documentation; 53.4% of patients (n=574) experienced a drug omission in the preceding 24 h; 22.5% of omitted doses were classified as 'critical'; 22.1% of patients (n=482) either had ID wristbands not reflecting their allergy status or no ID wristband; and 45.3% of patients (n=237) had drugs that were either unlabelled or labelled for another patient in their drug lockers. The quantitative analysis found no significant improvement in intervention wards following scorecard feedback. Interviews suggested staff were interested in scorecard feedback and described process and culture changes. Factors influencing scorecard implementation included 'normalisation' of errors, study duration, ward leadership, capacity to engage and learning preferences.

Discussion Presenting evidence-based performance indicators may potentially influence staff behaviour. Several practical and cultural factors may limit feedback effectiveness and should be considered when developing improvement interventions. Quality scorecards

should be designed with care, attending to evidence of indicators' effectiveness and how indicators and overall scorecard composition fit the intended audience.

INTRODUCTION

Medication safety is recognised as an important component of patient safety.¹ Medication errors occur frequently, increasing risk of harm to patients and costs to health systems.^{2–10} Interventions to improve medication safety take numerous forms,^{11–14} with significant efforts in recent years to encourage prioritisation of this safety issue across the English National Health Service (NHS).^{3 10 15}

Performance indicators are an established tool of governance applied at macro-level (national), meso-level (organisational) and micro-level (service, ward or individual).^{16–20} Indicators have two broad functions in healthcare settings: to support accountability systems and to support front-line quality and safety improvement.^{17 20 21} Within hospital organisations, senior leadership may use quality indicators to assess organisational progress against strategic objectives. Research on hospital board governance in the UK, US and Canada reports increasing use of quality indicators in organisational 'dashboards' to support quality assurance and improvement, particularly among 'high-performing' organisations.^{22–24} Presenting performance indicators to staff may align staff with organisational priorities and support quality and safety improvement.^{17 20 25 26}

Evidence on using performance indicators to support quality and safety improvement suggests that both indicator selection and approach to presentation are important. Evidence suggests that an indicator is more effective when there is strong evidence of association with improved outcomes, when it closely reflects the process it aims to measure and when the measured process is closely related to the desired outcome.¹⁸ Adaptation to local contexts is recognised as important.^{16 17} Evidence suggests that indicators are best used formatively to guide discussions of local improvement. This process benefits from involving stakeholders and in harnessing ‘soft’ data to strengthen the indicators’ message.¹⁷

Reviews of the evidence on audit and feedback in healthcare settings indicate that this technique can positively influence behaviour, but that such effects tend to be small and variable.^{27–30} Reviews indicate that factors underlying the varying effects of audit and feedback, and how the technique might be tailored to different settings, require further exploration. In support of this, rigorous mixed methods evaluation is recommended.^{27–29} As with audit and feedback, reviews identify a need for greater empirical evidence on the use of performance indicators,^{16 17} with a particular focus on issues of implementation.¹⁷ Reviews of quality and safety improvement and spreading innovations suggest it is important to understand the nature of an intervention, how it is implemented and the context into which it is introduced.^{31–35}

In this study, we aimed to evaluate the effects of implementing a ward-level medication safety scorecard in two NHS hospitals in a large English city and examine factors influencing these effects. We first present the scorecard’s development and implementation. In evaluating its effects, we consider quantitative and qualitative evidence of changes that occurred with the introduction of scorecard feedback. We also analyse the extent to which these effects were influenced by the scorecard contents, feedback approach and organisational setting in which the work took place. We are therefore able to examine factors influencing governance of medication safety. Further, based on the analysis, we identify lessons for using quality and safety indicators at ward level.

METHODS

Design

This was a mixed methods study. The quantitative component was a controlled before and after analysis of performance on a ward-level medication safety scorecard. The qualitative component was based on interviews with hospital staff about governance of medication safety and experiences of scorecard feedback. Drawing these components together permitted understanding of how and why the intervention influenced staff behaviour, whether there were any unintended consequences and what factors were influential.^{36 37}

A process of triangulation was applied, where the two research components were analysed separately, with the results drawn together where appropriate.³⁸ First, we examined qualitative and quantitative data to evaluate how scorecard feedback influenced staff behaviour. Second, the qualitative analysis explored implementation issues as a means of understanding these effects.

Participants

The scorecard was developed and piloted in two acute NHS hospital organisations situated in a large English city (hereafter referred to as Hospitals A and B).

In each hospital, we recruited three wards: two ‘intervention’ wards and one ‘control’. In Hospital A, all three wards were general medicine wards. In Hospital B, two general medicine wards and an elderly care ward were recruited, with one general medicine ward acting as the control.

We interviewed 25 staff in Hospital A and 24 staff in Hospital B. Interviewees included a range of professionals and managers working at ward level and in governance of quality and patient safety (see online supplementary appendix A).

Procedure

Developing the scorecard

The study was developed and implemented in collaboration with staff in both hospitals. Our approach sought to harness potential benefits of knowledge co-production, such as increased sensitivity to local context and greater relevance of findings to end users.^{39 40} A mapping exercise conducted in Hospital A identified a gap in knowledge of how medication safety is governed on wards.⁴¹ A group of senior clinicians and management with responsibilities for governing quality and safety in Hospital A proposed developing and piloting a ward-level medication safety scorecard. This intervention was proposed based on the view that a ‘scorecard culture’, where staff prioritise safety issues that appear on scorecards, existed in the hospital. A senior pharmacist led selection of indicators, which were chosen for their recognised importance to medication safety.^{3 10} Performance was coded red (non-compliant) or green (compliant) and presented relative to the previous week’s performance (figure 1A).

While the overall structure of the scorecard was retained for Hospital B, the content was amended. Following suggestions from both Hospital A and B staff, the revised scorecard featured fewer indicators overall; also, following Hospital B pharmacists’ recommendation, indicators related to prescribing were added (figure 1B). This process reflects the pragmatic nature of quality improvement, where learning and local intelligence guide adaptation of indicators.^{16 17} Performance was coded red (not compliant), amber (partly compliant) or green (compliant) and again presented relative to the previous week’s performance.

A

XXX ward — medication safety	Target	Prev	Performance	Change
		DD/MM	DD/MM/YYYY	↑=↓
Allergy documentation				
Patients whose allergy documentation was completed	100%	Z%	X/ Y (Z%)	↑=↓
Patients whose allergy symptoms were documented	100%	Z%	X/ Y (Z%)	↑=↓
Patients whose allergy documentation was signed	100%	Z%	X/ Y (Z%)	↑=↓
Patients whose allergy documentation was dated	100%	Z%	X/ Y (Z%)	↑=↓
Patient identification band				
Patients who had an ID band	100%	Z%	X/ Y (Z%)	↑=↓
Patients who had an appropriately coloured ID band	100%	Z%	X/ Y (Z%)	↑=↓
POD lockers				
Total no. drugs in POD lockers	-	X	Y	
No. drugs labelled for incorrect patient in POD lockers	0%	Z%	X/ Y (Z%)	↑=↓
No. unlabelled drugs in POD lockers	0%	Z%	X/ Y (Z%)	↑=↓
No. patients whose POD locker contained drugs labelled for incorrect patient	0%	Z%	X/ Y (Z%)	↑=↓
No. patients whose POD locker contained unlabelled drugs	0%	Z%	X/ Y (Z%)	↑=↓
Drug omissions				
Total drugs prescribed	-	X	Y	
Drug omissions (percentage of drugs prescribed)	0%	Z%	X/ Y (Z%)	↑=↓
Unexplained omissions (percentage of omissions)	0%	Z%	X/ Y (Z%)	↑=↓
Patients who experienced any drug omission	0%	Z%	X/ Y (Z%)	↑=↓
Trigger drugs				
Patients administered Vitamin K	0%	X	X	
Patients administered Octaplex	0%	X	X	
Patients administered FFP	0%	X	X	
Patients administered Naloxone	0%	X	X	
Patients administered Flumenazil	0%	X	X	
Patients administered Glucagon	0%	X	X	
Patients administered any trigger drugs	0%	X	X	
Secure drug storage				
Badge entry to ward required?	Y	Y/N	Y/N	↑=↓
Clinical room locked?	Y	Y/N	Y/N	↑=↓
Drug cupboards locked?	Y/NA	Y/N/NA	Y/N/NA	↑=↓
ID challenged?	Y/NA	Y/N/NA	Y/N/NA	↑=↓
Access to secure drugs				
Morphine ampoules	N/NA	Y/N/NA	Y/N/NA	↑=↓
Midazolam ampoules	N/NA	Y/N/NA	Y/N/NA	↑=↓
Oral diazepam	N/NA	Y/N/NA	Y/N/NA	↑=↓
Oral codeine	N/NA	Y/N/NA	Y/N/NA	↑=↓
Insulin (from fridge)	N/NA	Y/N/NA	Y/N/NA	↑=↓
Safety notices				
Penicillin traffic light poster present	Y	Y	Y/N	↑=↓
Medication safety bulletin present	Y	N	Y/N	↑=↓
SureMed information present	Y	Y	Y/N	↑=↓

SUMMARY	Indicators that improved	↑	X
	Indicators that stayed the same	=	X
	Indicators that got worse	↓	X

Figure 1 (A) Scorecard used in Hospital A. (B) Scorecard used in Hospital B. Note: FFP stands for 'fresh frozen plasma'.

Scorecard data collection

In both hospitals, scorecard data were collected on unannounced ward visits. In Hospital A, in the baseline phase (July 2009–February 2010), data were collected by

AR and pharmacists. During the feedback phase, data were collected weekly by AR and a member of audit staff. In Hospital B, ST and pharmacists collected scorecard data over the period September–November 2011.

B

XXX ward	Target	Previous	Performance	Change
		DD/MM	DD/MM	↑=↓
Allergy and ID bands				
Patients whose allergy status was completed, symptoms documented, signed and dated	100%	X/ Y (Z%)	X/ Y (Z%)	↑=↓
Patients with ID band	100%	X/ Y (Z%)	X/ Y (Z%)	↑=↓
Appropriate ID band (if allergy)	100%	X/ Y (Z%)	X/ Y (Z%)	↑=↓
ID bands on same wrist (if allergy)	100%	X/ Y (Z%)	X/ Y (Z%)	↑=↓
POD lockers				
Drugs for correct patient in POD locker	100%	X/ Y (Z%)	X/ Y (Z%)	↑=↓
Drug omissions				
Doses omitted over the last 24 hours (except code 6)	0%	X/ Y (Z%)	X/ Y (Z%)	↑=↓
Unexplained omissions (% of omissions)	0%	X/ Y (Z%)	X/ Y (Z%)	↑=↓
'Drug not available' (code 7) for two consecutive omissions (% of omissions)	0%	X/ Y (Z%)	X/ Y (Z%)	↑=↓
Prescribing				
Prescriptions without prescriber's signature	0%	X/ Y (Z%)	X/ Y (Z%)	↑=↓
Doses administered without prescriber's signature	0%	X/ Y (Z%)	X/ Y (Z%)	↑=↓
Once-only prescriptions without drug start time	0%	X/ Y (Z%)	X/ Y (Z%)	↑=↓
Unacceptable abbreviation of units	0%	X/ Y (Z%)	X/ Y (Z%)	↑=↓
Brand prescribing (excluding exceptions)	0%	X/ Y (Z%)	X/ Y (Z%)	↑=↓
SUMMARY	Indicators that improved		↑	X
	Indicators that stayed the same		=	X
	Indicators that got worse		↓	X

Figure 1 Continued.

Feedback to ward staff

In Hospital A, scorecard data were sent to ward leaders in intervention wards on a weekly basis over 7 weeks (February–April 2010). It was agreed that ward managers would present the scorecard to nurses at ward handover meetings and that consultants would present the data at junior doctors' educational meetings.

In Hospital B, data were fed back on a weekly basis to intervention wards over 7 weeks (October–November 2011). Feedback took place at established staff meetings (some uniprofessional, some multiprofessional) attended by ST and a Hospital B pharmacist. In these meetings, staff discussed the scorecard data, identifying possible causes of and potential solutions to the issues identified. Using established meetings in both hospitals permitted examination of existing approaches to governing medication safety in the participating organisations.

Interviews

In each hospital, interviews were conducted once the intervention phases were complete. Interviews lasted

between 15 and 60 min in locations agreed with interviewees. Interviewees discussed aspects of medication safety and their experiences of providing and receiving scorecard feedback. Interviews were digitally recorded and professionally transcribed. Transcripts were stored for analysis on QSR NVivo software.

Analysis

To analyse the effects of scorecard feedback, we applied a standard 'difference in difference' technique.⁴² Difference in difference is a quasi-experimental technique that measures the effect induced by a treatment, policy or intervention. It combines within-subject estimates of the effect (pretreatment and post-treatment) with between-subject estimates of the effect (intervention vs control groups). Using only between-subject estimates produces biased results if there are omitted variables that confound the difference between the two groups. Similarly, using only the within-group estimate produces biased results if there is a time trend present (eg, wards are always improving). The difference in

Table 1 Summary of medication safety risks analysed

Indicator	Medication safety risks	Collection method
Allergy documentation	The proportion of patients whose allergy documentation was incomplete (including symptoms documented, signed and dated)	Inspecting patient drug chart
Drug omissions	The proportion of patients who experienced a drug omission in the 24 h prior to data collection (excluding appropriate omissions, eg, patients refusing the dose)	Inspecting patient drug chart
ID wristbands	The proportion of patients who either had an ID wristband that did not reflect their allergy status or who had no ID wristband at all	Comparing allergy documentation on drug chart with colour of patient wristband
Patient's own drugs (POD) locker contents	The proportion of patients whose POD lockers contained either unlabelled drugs or drugs labelled for another patient (in Hospital B drugs that were not <i>currently</i> prescribed for patient were categorised as inappropriate)	Comparing patient name on drug chart with name appearing on drugs in POD locker; in Hospital B, also comparing with current patient drug chart

difference estimate combines the estimates and produces a more robust measure of the effect of an intervention.

The independent variables were 'ward type', 'study phase' and 'effect':

- ▶ 'Ward type' measured the difference between intervention wards, where scorecard data were presented to staff, and control wards, where no data were presented.
- ▶ 'Study phase' measured the difference between the baseline phase, during which data were collected but not fed back, and the feedback phase, during which data continued to be collected and were fed back to intervention wards.
- ▶ 'Effect' was the interaction (multiplication) of the 'ward type' and 'study phase' variables and isolated 'effect' of the intervention on the intervention wards. This was the so-called 'difference in difference' estimator.

Dependent variables were wards' performance on four indicators featured on the medication safety scorecard: we measured the proportion of patients experiencing a risk on these indicators (table 1). These were selected because they were identified by both the participating hospitals and in guidance as important factors in medication safety.³ Standard ordinary least squares regression was used to estimate the four difference in difference models.

Qualitative analysis

Interviews were analysed using data-driven thematic analysis.⁴³ In part, the quantitative analysis guided our focus. Analysing staff response to scorecard feedback permitted examination of effects not captured by the scorecard, and factors influencing the process and effects of feedback. Based on this, lessons for future implementation of scorecards of this kind were identified.

RESULTS

Quantitative analysis

Table 2 presents performance on medication safety indicators at baseline in terms of the percentage of instances in which patients experienced a medication safety risk (see column '% error'). It indicates that while performance varied across indicators, a

significant proportion of patients were exposed to medication safety risks in all cases (see online supplementary appendix B for further breakdown of patients).

Each drug omission (ie, each dose) was categorised using a list of 'critical drugs not to omit'¹⁵ developed by Hospital B. We found that, at baseline, 22.5% of omitted doses (N=865) could be classified as 'critical omissions', for example, heparin, oxycodone and levotiracetam, where omission or delay in doses increases the likelihood of outcomes such as mortality, haemorrhage and infection (see online supplementary appendix C).

Figure 2 presents performance across the four selected medication safety indicators disaggregated by Hospital (A or B), ward type (control or treatment) and phase (baseline or feedback). Where the gradients of the lines in figure 2 differ for pairs of wards (control and intervention) within hospitals, the feedback intervention is having an effect and the larger this divergence, the stronger the effect. The significance of these effects was addressed using difference in difference regression. The data were fitted using standard ordinary least squares regression, and the

Table 2 Performance on medication safety indicators at baseline

Indicator	Hospital	Total patients	Mean patients* (SD)	% error (SD)
Allergy documentation	Hospital A	390	20.53 (3.79)	19.3 (9.3)
	Hospital B	240	16.00 (3.93)	18.0 (11.5)
	Total	630	18.53 (4.43)	18.7 (10.2)
Drug omissions	Hospital A	336	16.80 (3.86)	46.7 (15.6)
	Hospital B	238	15.87 (3.72)	62.4 (13.9)
	Total	574	16.40 (3.78)	53.4 (16.6)
ID wristbands	Hospital A	264	12.57 (3.63)	30.9 (20.9)
	Hospital B	218	14.53 (3.66)	8.5 (8.0)
	Total	482	13.39 (3.72)	22.1 (20.2)
POD locker contents	Hospital A	397	18.90 (5.43)	29.9 (14.2)
	Hospital B	237	15.80 (3.78)	66.9 (18.2)
	Total	634	17.61 (5.00)	45.3 (24.3)

*Mean number of patients per ward, per data collection visit. POD, patient's own drugs.

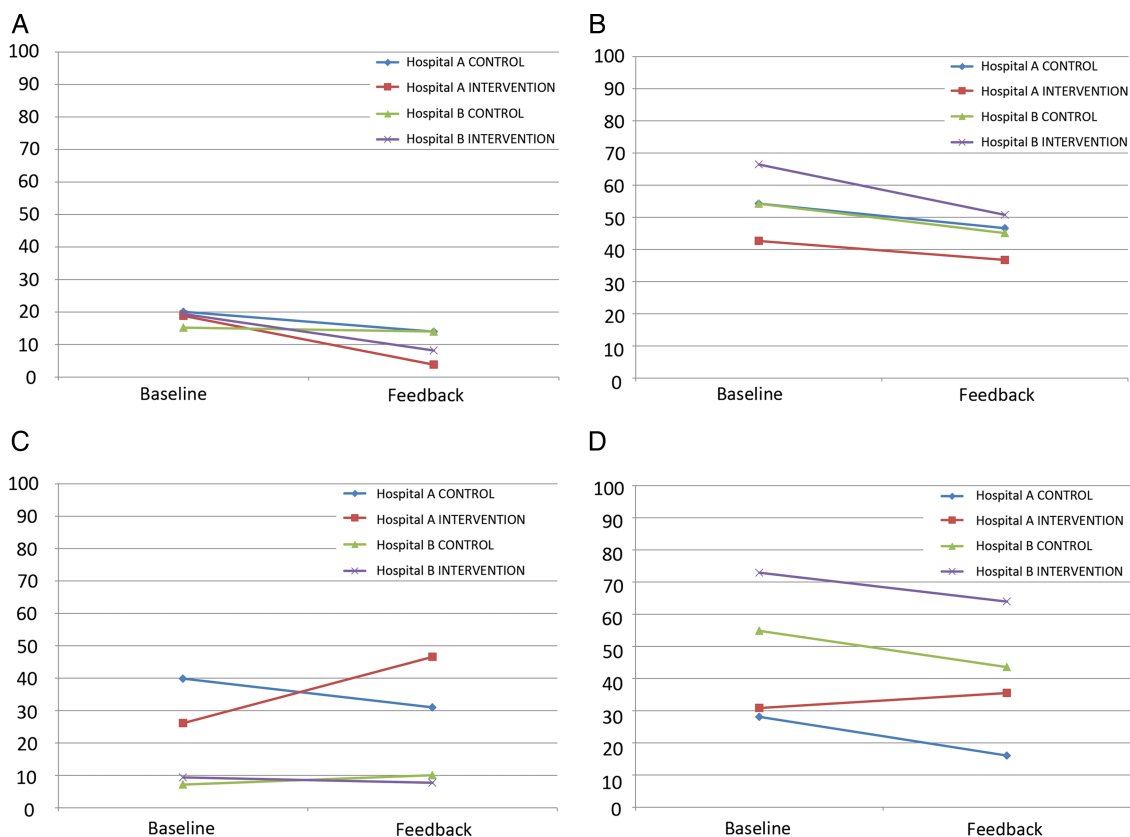


Figure 2 (A) Percentage of patients with incomplete allergy documentation. (B) Percentage of patients who experienced a drug omission in the preceding 24 h. (C) Percentage of patients who did not have an appropriately coloured ID wristband (D) Percentage of patients with inappropriate patient's own drugs locker contents.

effect on the intervention wards is captured by the difference in the difference. That is the difference between the control and treatment wards over the pretreatment and post-treatment periods. This difference in difference effect is reported as unstandardised β coefficient for each of the performance measures in table 3 alongside the R-squared (R^2), which represents the proportion of variance in the performance measure that is explained by the regression model. The statistical significance of the intervention effect is also indicated.

These models explain only a modest proportion of the variation in the data. This suggests other factors might have been influential or that patterns observed may represent random variation. Two significant effects were found: in Hospital A, performance on the ID wristbands and patient's own drugs (PODs) lockers indicators improved in the control ward while deteriorating in intervention wards (illustrated in figure 2C, and 2D, respectively). No statistically significant improvements were associated with presenting scorecard data to staff.

These data offer two messages: first, on all selected indicators, many patients were exposed to medication safety risks at baseline, and second, scorecard feedback did not significantly improve this situation in either hospital.

Qualitative analysis

In this section, we examine factors that might explain these results. First, we cover some practical issues that may have limited the effect of scorecard feedback. Second, we examine how staff perceived the content of the scorecard and approach to feedback.

Feedback challenges

Scorecard data were not always presented to staff. In both hospitals, meetings were sometimes cancelled due to competing priorities or practical issues. The most pronounced example occurred in a Hospital A intervention ward:

...there was a weekly juniors' meeting specifically for the [Hospital A ward] team. It seems to have gone into

Table 3 Summary of difference in differences analyses

Measure	Hospital A		Hospital B	
	R^2	β	R^2	β
Allergy documentation	0.480	-8.908	0.179	-10.013
Drug omissions	0.184	1.739	0.162	-6.574
ID wristbands	0.201	29.304*	0.025	-4.555
POD lockers	0.240	16.694*	0.167	2.318

β represents the difference in difference effect.

*Signifies $p < 0.05$.

POD, patient's own drugs.

abeyance [...] largely because of lack of space (Consultant, Hospital A)

Ward managers and consultants were identified as key to sharing information with nurses and junior doctors, respectively. However, interviewees suggested the degree to which clinical leaders engage in such processes might vary:

...you would just hope that the ward manager would cascade information. But one couldn't guarantee that (Pharmacist, Hospital A)

I was engaged in the conversation, and found it quite enjoyable, but I looked over at [consultant] and I know s/he was keen to get back out there (Consultant, Hospital B)

Some staff also suggested that translating messages into practice might take more time than the study period permitted.

Maybe it just takes longer time than the intervention period to actually get that message through (Consultant, Hospital B)

Other changes took place over the course of the study. These included a reorganisation of the Hospital A ward structure during the baseline phase. Broader issues, such as staffing and financial pressures, were frequently cited by interviewees. Taken together, these data indicate that competing priorities, wider resource issues and variable engagement of ward leadership may have limited the effect of scorecard feedback. Such factors are likely to represent more widely felt obstacles to leading quality and safety improvement initiatives.⁴⁴

Staff experiences of scorecard feedback

Nurses were receptive to scorecard feedback, finding it relevant to their practice:

...some of the things that the scorecard was talking about, it was like a reinforcement (Nurse, Hospital A)

It brought things forward that we already know as best practice [...], 'Oh yes, I should get on top of that!' (Nurse, Hospital B)

In both hospitals, nurses paid particular attention to feedback on POD locker contents and reported use of new routines to address the issue:

When you're cleaning the bed, make sure you tell them to open the POD locker. Once the POD locker is opened then you know there's medicine, take it out and leave it open, at least there is nothing in there when the next patient comes in (Nurse, Hospital A)

Now, when a patient gets discharged I always open up the cabinet and make sure everything is cleared out (Senior staff nurse, Hospital B)

These data suggest that nurses found scorecard feedback useful and sought to translate it into practice. This reflects scorecards' potential to support

quality improvement when a measure is perceived as relevant and performance is seen as suboptimal. However, given the levels of error presented, it is interesting to note that the overall response demonstrated relatively little urgency. This was also a common theme in doctors' responses to scorecard feedback:

So... just useful information to know but I'm not sure if it necessarily altered how we did things. (Junior Doctor, Hospital A)

In Hospital B, doctors frequently referred to prescribing measures as useful, reflecting staff prioritisation of issues perceived as relevant to their own practice:

The prescribing stuff is useful, particularly for the junior staff—reiterating the need to prescribe differently (Doctor, Hospital B)

Well, it has made me more mindful of not prescribing trade names (Consultant, Hospital B)

Junior doctors frequently referred to learning about medication safety directly from ward pharmacists, while medical consultants emphasised uniprofessional approaches to learning typified by the 'apprenticeship' model, where doctors reflect on individual practice under the guidance of more experienced peers.^{45–50}

Scorecards and traffic lights are very familiar now [...] actually it doesn't really fit that well in terms of [the] medical model of working. (Consultant, Hospital A)

I think that using the clinical supervisor or the education supervisor would be a more effective way of getting that kind of targeted feedback to me as an individual (Consultant, Hospital A)

Our data suggest nurses engaged more with scorecard feedback than doctors. These differences may reflect a contrast in how well the scorecard approach fitted with doctors' and nurses' established approaches to improvement. This 'lack of fit' may have reduced the extent to which doctors engaged with the scorecard intervention. However, another potential factor is the relative likelihood of sanction:

If a nurse commits a serious medicines error, they're removed from their responsibilities. If a doctor commits a serious medicines error, it doesn't happen: there's no censure at all (Consultant, Hospital B)

Considering the overall composition of the scorecard revealed a potential source of tension. While staff engaged with 'their' indicators, some stated that indicators associated with other professions were not relevant. However, the potential for a joint scorecard to encourage staff to take a wider view of their role in medication safety was noted:

in retrospect we would have been better to do that jointly with the nursing staff because [...] for the nurses

to hear the doctors say, 'Well, that's not my job,' might have been quite useful, because they could have said, 'Well, actually, it would be really helpful if you would look at that aspect' (Consultant, Hospital A)

I think it's helped nurses [...]to go up to the doctors and say, "You've written the brand name," [...] or if their writing's not clear: I think it's making us look more enabled (Deputy ward manager, Hospital B)

Thus, the relationship between the content of feedback and the forum in which it is to be presented is important. Where feedback is presented to a multiprofessional group, a mixed scorecard may prompt valuable discussion and negotiation of roles; where feedback is presented to a single profession, a mixed scorecard may include indicators perceived as irrelevant to their practice, and its effect diluted. Further, joint feedback sessions should be led with care to ensure discussions do not become an outlet for blame rather than learning.

DISCUSSION

Our data confirm previous research suggesting medication safety risks occur frequently.²⁻⁹ It also indicates that a notable proportion of these errors had the potential to cause harm. In relation to the impact of scorecard feedback, the quantitative analysis found no significant improvements. However, our qualitative data suggested staff were highly receptive to the data presented and importantly that staff reported that feedback prompted both process and culture changes. These findings support previous research that suggests presenting quality indicators may encourage engagement with patient safety issues, but with limited effect on actual performance.^{28 29} Our results also reflect a wider evidence base indicating that improvement initiatives may often have limited or no demonstrable effect on quality, and that this may often be associated with contextual factors.^{33 51}

Our qualitative analysis revealed a number of contextual factors, practical and cultural, that may mediate the effect of scorecard feedback. Practical factors identified by staff included the capacity for staff to engage with feedback and the capacity to translate feedback into practice, for example, having time to design and embed new processes: these are established obstacles to quality improvement work.^{16 31 33 44} Significant cultural factors included how staff respond to data indicating poor performance, and professions' different styles of engaging with quality improvement activity. The lack of urgency in staff response to the feedback was notable. This may reflect a wider cultural issue of 'normalised deviance', where frequent errors are perceived as unsurprising or even acceptable. Such normalisation has been identified as a potentially significant risk to patient safety.⁵¹⁻⁵³ This may relate to what Francis described as 'a culture of habituation and passivity' in

his report into poor quality care at Mid Staffordshire NHS Foundation Trust.⁵⁴ Our data support previous research indicating that different professions engage differently with quality and safety improvement activity.^{35 45-48 55-57} Doctors' and nurses' responses also support previous research indicating that staff engage with performance indicators that are seen as relevant to practice and associated with quality and safety outcomes.¹⁶ However, a corollary of this is that 'irrelevant' indicators may dilute a scorecard's credibility, unless an objective of scorecard feedback is to encourage shared understanding of medication safety across professions, facilitated through feedback to multiprofessional meetings.

This study has some limitations. First, that our intervention fitted poorly with doctors' approach to quality and safety improvement suggests a key assumption of the study—existence of a 'scorecard culture'—was not borne out. Widening our collaborative process to include a wider range of front-line staff may have highlighted this and encouraged reflection on how best to communicate with different professions. Second, the interventions did not take place concurrently in Hospitals A and B. The intervening 18-month period may have seen changes in local and national contexts. However, interviewees indicated that the contextual issues faced by the participating hospitals were quite consistent. Third, the scorecards used in Hospitals A and B differed and this may have had an effect on the results. For example, the difference between Hospital A's and Hospital B's performance on POD locker contents at baseline may reflect the additional criteria applied for 'appropriate labelling' in Hospital B. Given the potential influence of scorecard composition, the fact that the scorecards used in Hospitals A and B differed in terms of number and nature of indicators selected may be important. The presence of other indicators on scorecards may have influenced staff response, that is, a higher number of indicators may reduce the effect of specific indicators. However, the differences in timing and indicators used reflect this study's sequential approach. This approach permitted learning across the two interventions, reflecting recommendations for developing effective quality and safety improvement interventions.^{16 17} Fourth, as intervention and control wards experienced the same data collection process, the potential effects of being audited should also be acknowledged. Control ward staff may have changed behaviour because of an awareness that they were being measured; they may even have become aware of specific indicators being measured (eg, the visibility of POD locker and ID wristband inspections) and adjusted their behaviour accordingly.⁵⁸ Fifth, in terms of capturing the effect of the scorecard, research indicates that the intervention and associated data collection period was too short.¹⁶ This may have limited the extent to which staff had time to make and embed

changes, and for these to translate into improved medication safety. Finally, while interviews commenced after completion of the feedback phase, the quantitative analysis had not been completed. There may be value in incorporating quantitative findings, that is, whether the intervention had a significant effect into evaluation interviews.

Turning to future work, we found the mixed methods approach valuable in understanding the effect of scorecard feedback. However, we propose amending this approach to build on the learning and limitations identified above. Future research might address feedback approach and duration, scorecard design and evaluation methods. First, parallel implementation in two or more hospitals, using consistent wards, would allow clearer understanding of how similar wards in different organisational settings engage with quality and safety indicators. Second, using the same scorecard across all participating wards (developed in collaboration with all participating hospitals) would permit a stronger focus on how tool composition interacts with varying organisational settings. Testing the relationship between indicators and audience at other organisational levels, for example, hospital boards' interactions with quality dashboards, may provide valuable insights on how best to support leadership for quality improvement.^{22–24} Third, greater use of observational methods would allow a clearer understanding of processes at work in improvement work of this kind, including the extent to which control ward staff are influenced by factors such as the act of measurement and 'contamination' from wards that receive feedback.^{59 60} Key foci would include feedback sessions and the embedding of any changes identified in response to feedback. Finally, a longer follow-up period would be valuable. This would allow staff sufficient time to incorporate findings into process redesign, and for these changes to be adopted by staff. It would permit exploration of staff's initial responses to interventions and their views on the nature and sustainability of changes brought about in the longer term.

CONCLUSIONS

Medication errors occur frequently at ward level. Presenting evidence of this in the form of a scorecard has limited effect on staff performance. Established factors, including organisational processes and professional cultures, remain influential mediators of change. The persistence of these meso-level issues, relatively unchanged from past research in this area, suggests that efforts to prioritise and improve medication safety across the English NHS are yet to gain traction.^{3 10 15}

Our findings indicate that presenting evidence-based performance indicators has the potential to influence staff behaviour. However, when applying such interventions it is important to allow sufficient time for

such changes to translate into overall performance improvements. Further, there exist several practical and cultural contextual factors that may limit the effectiveness of feedback, and which should be considered when developing the improvement strategy. When selecting performance indicators, the extent to which they fit the task and the setting should be assessed, and the feedback approach should be considered carefully. Further research on scorecard composition and implementation, drawing on the lessons above, may contribute to clearer understanding of how performance indicators might be used effectively in a range of healthcare settings.

Acknowledgements The authors thank all the staff of the two study hospitals who contributed to this research. We also thank members of the NIHR King's PSSQ Research Centre and the Department of Applied Health Research, UCL, for their helpful comments on earlier drafts of the article. Finally, we thank two anonymous referees and the deputy editor for their comments on previous versions of this article.

Contributors AR and ST made substantial contributions to conception and design, acquisition of data and analysis and interpretation of data. GCa, RT and AO made substantial contributions to conception and design, acquisition of data and interpretation of data. GCo made substantial contributions to design, and analysis and interpretation of data. NF made substantial contributions to conception and design, and analysis and interpretation of data. All authors made substantial contributions to drafting the article or revising it critically for important intellectual content; and final approval of the version to be published.

Funding The NIHR King's Patient Safety and Service Quality Research Centre (King's PSSQRC) was part of the National Institute for Health Research (NIHR) and was funded by the Department of Health from 2007–2012.

Disclaimer This report presents independent research commissioned by the NIHR. The views expressed in this report are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Competing interests None.

Ethics approval This research was approved by King's College Hospital NHS research ethics committee (ref 09/H0808/78).

Provenance and peer review Not commissioned; externally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/3.0/>

REFERENCES

- 1 Bates DW, Larizgoitia I, Prasopa-Plaizier N, *et al*. Global priorities for patient safety research. *BMJ* 2009;338:1242–4.
- 2 Audit Commission. *A spoonful of sugar*. London: Audit Commission, 2001.
- 3 National Patient Safety Agency. *Safety in doses*. London: National Patient Safety Agency, 2009.
- 4 Kohn LT, Corrigan J, Donaldson MS. *To Err Is Human: Building a Safer Health System*. Washington, DC: National Academy Press, 2000.

- 5 Lisby M, Nielsen LP, Mainz J. Errors in the medication process: frequency, type, and potential clinical consequences. *Int J Qual Health Care* 2005;17:15–22.
- 6 Dean B, Schachter M, Vincent C, *et al.* Prescribing errors in hospital inpatients: their incidence and clinical significance. *Qual Saf Health Care* 2002;11:340–4.
- 7 Barber N, Rawlins M, Franklin BD. Reducing prescribing error: competence, control, and culture. *Qual Saf Health Care* 2003;12:i29–32.
- 8 Franklin BD, Vincent C, Schachter M, *et al.* The incidence of prescribing errors in hospital inpatients. *Drug Saf* 2005;28:891–900.
- 9 Dean B, Schachter M, Vincent C, *et al.* Causes of prescribing errors in hospital inpatients: a prospective study. *The Lancet* 2002;359:1373–8.
- 10 National Patient Safety Agency. *Safety in doses*. London: National Patient Safety Agency, 2007.
- 11 Flemons WW, McRae G. Reporting, learning and the culture of safety. *Healthc Q* 2012;15:12–7.
- 12 Hillestad R, Bigelow J, Bower A, *et al.* Can electronic medical record systems transform health care? Potential health benefits, savings, and costs. *Health Aff* 2005;24:1103–17.
- 13 Upton D, Taxis K, Dunsavage J, *et al.* The evolving role of the pharmacist—how can pharmacists drive medication error prevention? *J Infus Nurs* 2009;32:257–67.
- 14 Wailoo A, Roberts J, Brazier J, *et al.* Efficiency, equity, and NICE clinical guidelines. *BMJ* 2004;328:536–7.
- 15 National Patient Safety Agency. *Reducing harm from omitted and delayed medicines in hospital*. London: National Patient Safety Agency, 2010.
- 16 de Vos M, Graafmans W, Kooistra M, *et al.* Using quality indicators to improve hospital care: a review of the literature. *Int J Qual Health Care* 2009;21:119–29.
- 17 Freeman T. Using performance indicators to improve health care quality in the public sector: a review of the literature. *Health Serv Manage Res* 2002;15:126–37.
- 18 Chassin MR, Loeb JM, Schmaltz SP, *et al.* Accountability measures—using measurement to promote quality improvement. *N Engl J Med* 2010;363:683–8.
- 19 Barach P, Johnson J. Understanding the complexity of redesigning care around the clinical microsystem. *Qual Saf Health Care* 2006;15:i10–i6.
- 20 Palmer RH. Using clinical performance measures to drive quality improvement. *Total Qual Manag* 1997;8:305–12.
- 21 Meyer GS, Nelson EC, Pryor DB, *et al.* More quality measures versus measuring what matters: a call for balance and parsimony. *BMJ Qual Saf* 2012;21:964–8.
- 22 Baker GR, Denis J-L, Pomey M-P, *et al.* Designing effective governance for quality and safety in Canadian healthcare. *Healthc Q* 2010;13:38–45.
- 23 Jha AK, Epstein AM. Hospital governance and the quality of care. *Health Aff* 2010;29:182–7.
- 24 Jha AK, Epstein AM. A survey of board chairs of English hospitals shows greater attention To quality of care than among heir US counterparts. *Health Aff* 2013;32:677–85. doi:10.1377/hlthaff.2012.1060
- 25 Atkinson H. Strategy implementation: a role for the balanced scorecard? *Manag Decis* 2006;44:1441–60.
- 26 de Vos ML, van der Veer SN, Graafmans WC, *et al.* Process evaluation of a tailored multifaceted feedback program to improve the quality of intensive care by using quality indicators. *BMJ Qual Saf* 2013;22:233–41.
- 27 Gardner B, Whittington C, McAteer J, *et al.* Using theory to synthesise evidence from behaviour change interventions: the example of audit and feedback. *Soc Sci Med* 2010;70:1618–25.
- 28 Hysong SJ. Meta-analysis: audit and feedback features impact effectiveness on care quality. *Med Care* 2009;47:356–63.
- 29 Ivers N, Jamtvedt G, Flottorp S, *et al.* Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane Libr* 2012;6:CD000259.
- 30 Flottorp SA, Jamtvedt G, Gibis B, *et al.* *Using audit and feedback to health professionals to improve the quality and safety of health care*. Copenhagen: World Health Organization, 2010.
- 31 Dixon-Woods M, McNicol S, Martin G. Ten challenges in improving quality in healthcare: lessons from the Health Foundation's programme evaluations and relevant literature. *BMJ Qual Saf* 2012;21:876–84. doi:10.1136/bmjqs-2011-000760
- 32 Greenhalgh T, Robert G, Bate P, *et al.* How to spread good ideas: a systematic review of the literature on diffusion, dissemination and sustainability of innovations in health service delivery and organisation. London: National Co-ordinating Centre for NHS Service Delivery and Organisation, 2004.
- 33 Kaplan HC, Brady PW, Dritz MC, *et al.* The influence of context on quality improvement success in health care: a systematic review of the literature. *Milbank Q* 2010;88:500–59.
- 34 Pettigrew A, Ferlie E, McKee L. *Shaping strategic change. Making change in large organizations: the case of the National Health Service*. London: Sage, 1992.
- 35 Greenfield D, Nugus P, Travaglia J, *et al.* Factors that shape the development of interprofessional improvement initiatives in health organisations. *BMJ Qual Saf* 2011;20:332–7.
- 36 Johnson RB, Onwuegbuzie AJ. Mixed methods research: a research paradigm whose time has come. *Educ Res* 2004;33:14–26.
- 37 Tashakkori A, Teddlie C. Foundations of mixed methods research: integrating quantitative and qualitative approaches in the social and behavioral sciences. Thousand Oaks, CA: Sage Publications Inc, 2008.
- 38 O' Cathain A, Murphy E, Nicholl J. Three techniques for integrating data in mixed methods studies. *BMJ* 2010;341:1147–50.
- 39 Lomas J. The in-between world of knowledge brokering. *BMJ* 2007;334:129–32.
- 40 Graham ID, Tetroe J. How to translate health research knowledge into effective healthcare action. *Healthc Q* 2007;10:20–2.
- 41 Ramsay A, Magnusson C, Fulop N. The relationship between external and local governance systems: the case of health care associated infections and medication errors in one NHS trust. *Qual Saf Health Care* 2010;19:1–8.
- 42 Ashenfelter O, Card D. Using the longitudinal structure of earnings to estimate the effect of training programs. *Rev Econ Stat* 1985;67:648–60.
- 43 Boyatzis RE. Transforming qualitative information: Thematic analysis and code development. Sage Publications Inc, 1998.
- 44 Braithwaite J, Hindle D, Finnegan TP, *et al.* How important are quality and safety for clinician managers? Evidence from triangulated studies. *Clin Governance Int J* 2004;9:34–41.
- 45 McDonald R, Waring J, Harrison S, *et al.* Rules and guidelines in clinical practice: a qualitative study in operating theatres of doctors' and nurses' views. *Qual Saf Health Care* 2005;14:290–4.

- 46 Parker D, Lawton R. Judging the use of clinical protocols by fellow professionals. *Soc Sci Med* 2000;51:669–77.
- 47 Davies H, Powell A, Rushmer R. Why don't clinicians engage with quality improvement? *J Health Serv Res Policy* 2007;12:129–30.
- 48 Braithwaite J, Westbrook M, Travaglia J. Attitudes toward the large-scale implementation of an incident reporting system. *Int J Qual Health Care* 2008;20:184–91.
- 49 Hoff T, Jameson L, Hannan E, *et al.* A review of the literature examining linkages between organizational factors, medical errors, and patient safety. *Med Care Res Rev* 2004;61:3–37.
- 50 Rodriguez-Paz JM, Kennedy M, Salas E, *et al.* Beyond “see one, do one, teach one”: toward a different training paradigm. *Postgrad Med J* 2009;85:244–9.
- 51 Dixon-Woods M. Why is patient safety so hard? A selective review of ethnographic studies. *J Health Serv Res Policy* 2010;15:11.
- 52 Hughes C, Travaglia JF, Braithwaite J. Bad stars or guiding lights? Learning from disasters to improve patient safety. *Qual Saf Health Care* 2010;19:332–6.
- 53 Reiman T, Pietikäinen E, Oedewald P. Multilayered approach to patient safety culture. *Qual Saf Health Care* 2010;19:1–5.
- 54 Francis R. Report of the Mid Staffordshire NHS Foundation Trust public inquiry. London: Stationery Office, 2013.
- 55 Hockey PM, Marshall MN. Doctors and quality improvement. *JRSM* 2009;102:173–6.
- 56 Davies H, Powell A, Rushmer R. Healthcare professionals' views on clinician engagement in quality improvement. London: The Health Foundation, 2007.
- 57 Powell AE, Davies HTO. The struggle to improve patient care in the face of professional boundaries. *Soc Sci Med* 2012;75:807–14.
- 58 Wickström G, Bendix T. The “Hawthorne effect” — what did the original Hawthorne studies actually show? *Scand J Work Environ Health* 2000;26:363–7.
- 59 Waring JJ. Constructing and re-constructing narratives of patient safety. *Soc Sci Med* 2009;69:1722–31.
- 60 Dixon-Woods M, Bosk CL, Aveling EL, *et al.* Explaining Michigan: developing an ex post theory of a quality improvement program. *Milbank Q* 2011;89:167–205.