



UNIVERSITY OF LEEDS

This is a repository copy of *The effect of external inspections on safety in acute hospitals in the National Health Service in England: A controlled interrupted time-series analysis.*

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/145007/>

Version: Accepted Version

Article:

Castro-Avila, A, Bloor, K and Thompson, C orcid.org/0000-0002-9369-1204 (2019) The effect of external inspections on safety in acute hospitals in the National Health Service in England: A controlled interrupted time-series analysis. *Journal of health services research & policy*, 24 (3). pp. 182-190. ISSN 1355-8196

<https://doi.org/10.1177/1355819619837288>

(c) The Author(s) 2019. Castro-Avila, A, Bloor, K and Thompson, C (2019) The effect of external inspections on safety in acute hospitals in the National Health Service in England: A controlled interrupted time-series analysis. *Journal of health services research & policy*, 24 (3). pp. 182-190. Reprinted by permission of SAGE Publications.

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Quantitative empirical research

Corresponding Author:

Ana C Castro-Avila, Department of Health Sciences, University of York, Heslington, York, YO10 5DD, United Kingdom

Email: ana.castro@york.ac.uk

The effect of external inspections on safety in acute hospitals in the National Health Service in England: a controlled interrupted time-series analysis.

Castro-Avila AC^{1,2},

¹ Research fellow, Department of Health Sciences, University of York, York, United Kingdom.

² Lecturer, Carrera de Kinesiología, Clínica Alemana Universidad del Desarrollo, Santiago, Chile.

Bloor KE³, Thompson CA⁴.

³ Professor of Health Economics and Policy, Department of Health Sciences, University of York, York, United Kingdom.

⁴ Professor of Applied Health Research, Faculty of Medicine and Health, University of Leeds, Leeds, United Kingdom.

Abstract

Objectives: To evaluate the effect of Care Quality Commission (CQC) external inspections of acute trusts on adverse event rates in the English National Health Service (NHS).

Methods: Interrupted time series analysis including all acute NHS trusts in England (n=155) using two control groups (new versus historical inspection regime and trusts not inspected). Multilevel random-coefficients modelling of 1) rates of falls with harm and 2) pressure ulcers, from April 2012 to June 2016, was undertaken using the new, resource-intensive regime of CQC inspections as an intervention. Data used in the model included dates and type of inspection, patient safety indicators, demographic characteristics and financial risk of hospitals.

Results: In one year, CQC inspected 66 acute trusts (42% of all English trusts) using their new regime and 46 (30%) using their previous one. Prior to inspections being announced, rates of falls with harm and pressure ulcers were improving in both intervention and control hospitals. The announcement of an inspection did not affect either indicator. After inspections, rates of falls with harm improved more slowly and pressure ulcers rates no longer improved for trusts inspected using both regimes.

Conclusions: Neither form of external inspection was associated with positive, clinically significant effects on adverse event rates. Any improvement happening before the announced CQC inspections slowed after the inspection.

Keywords: Quality of health care, external inspection, accreditation, interrupted time series analysis.

Introduction

The English National Health Service (NHS), like all health systems, has unwarranted variations in quality and clinical outcomes,¹ and the Care Quality Commission (CQC) exists to regulate and improve quality.² CQC ratings are a form of inspection-based accreditation, which is a common approach to quality improvement,³ although evidence on its effectiveness is inconsistent and of low quality.⁴ CQC sets standards, monitors compliance, undertakes on-site inspections and has enforcement powers.⁵ It can request compliance actions, modify conditions of registration, issue warning notices, recommend trusts are put under ‘special measures’ and even prosecute Trust Boards.⁶

Economic theory suggests that in a market, organizations compete for ‘customers’. In the NHS, competition and provider numbers are often limited, so the main incentive to provide quality health care is reputational concern.^{7,8} Any approach that relies on performance measures carries the risk that trusts will engage in gaming behaviours to gain perceived advantage.⁸

Aside from gaming, trusts might ‘improve’ following an inspection through various mechanisms.⁹ Pre-inspection, trusts can self-assess their level of compliance and obtain information from trusts previously inspected. Any remedial action undertaken can increase awareness of safety and increase reporting of adverse events. During inspections, regulators may suggest improvements, and post-visit, alongside regulator-directed changes, organizations can make changes based on lessons learned. Stakeholders (for example, commissioners and patients’ groups) can influence organizations to improve and the publication of inspection results may lead to further changes. These potential CQC improvement mechanisms have been explored qualitatively but the lack of a system-wide quantitative examination of CQC inspections means their effects remain uncertain.⁴

CQC’s perceived lack of sensitivity to detecting or avoiding unacceptable failures of care¹⁰ led to modification of the inspection regime in 2013.¹¹ Historically, the CQC’s teams of 2-5 inspectors, visiting a trust over 2-3 days, focused on whether a trust met

28 essential standards, required improvement or required enforcement actions to meet the standards.¹² The new regime¹³ outlines legally enforceable fundamental standards of care, uses ‘intelligent monitoring’ of over 150 indicators, uses more inspectors for longer, rates core services and the whole trust, allows for unannounced visits post-inspection and permits joint inspection (with the financial regulator, NHS Improvement). The new regime uses more resources, but the effect on processes and outcomes has been not quantitatively evaluated.¹⁴

We sought to explore the effect of external inspections of acute trusts by the CQC on rates of adverse events in the English NHS.

Methods

Study design and setting

We undertook a controlled interrupted time-series analysis using data on adverse events prospectively and routinely reported by all English acute NHS trusts (n=155) from April 2012 to June 2016. Five trusts had missing data and were excluded. The interruption point was a CQC inspection between September 2013 and September 2014. In that period, a new inspection regime was introduced and applied to 66 acute NHS Trusts. Since CQC performs comprehensive inspections (at least) once every three years,⁵ some trusts were inspected under the old regime (n=46), whilst others were not inspected (n=38), which permitted us to use two non-equivalent controls. The announcement of the inspection occurring, 16 to 20 weeks before a visit, was treated as a further intervention point.

Outcome measures

We focused on two adverse events which are proxies for quality of nursing care¹⁵. Falls with harm are unplanned or unintentional descent to the floor requiring at least first aid, minor treatment or extra observation or medication¹⁶. Pressure ulcers are those that developed 72 hours or more after patient admission and are categorized using the European Pressure Ulcer Scale.¹⁶ Data on these outcomes are collected

monthly in a single-day census by trained staff using the NHS Safety Thermometer (NHS ST),¹⁷ a national quality improvement initiative that rewards trusts for collecting high(er) quality safety data.¹⁸ Falls with harm and pressure ulcers are indicators of quality used in CQC documentation and inspections, but are not used in CQC ‘intelligent monitoring’ algorithms to target inspections.⁵

Statistical analysis

We used a multilevel random-coefficients linear model for estimating changes in the level and trend in rates of falls with harm and pressure ulcers after each interruption.

The multilevel model estimates intervention effects for an average trust and for all individual trusts. To determine the pre-inspection change rate, time was used as an explanatory variable. A dummy variable, indicating a level change in adverse events rates, and an interaction term (i.e. group*time*post-intervention time), indicating a change in slope, were included in the model. Time was centred on the inspection month. In the case of not-inspected hospitals, an interruption was introduced in November 2013 (i.e. the month with the highest number of inspections). Outcome variables were logarithmically transformed, and coefficients were exponentiated to express them in the original scale. The assumptions of uncorrelated model residuals and normally distributed random effects were met.

We included descriptive data for each acute NHS trust (see Table 1) to minimize confounding. Because of changes in the commissioning system, the period between October 2012 and April 2013 was introduced as a dummy variable. Confounding variables were included one by one in the adjusted model and the most parsimonious model is reported based on the Lowest Bayesian Information Criterion.

Table 1 Descriptive information used for each acute NHS trust and its sources.

Descriptive information about trusts	Data Source
Number of beds	CQC inspection report
Geographical location	CQC inspection report
Trust size (Teaching, large, medium, small, specialist)	CQC inspection report
Foundation Trust status	CQC inspection report
Number of trust sites	CQC website
Catchment population	CQC inspection report
Reporting culture rating (2016)	CQC website
z-score for potential underreporting of safety incidents (2013)	Reported by National Reporting and Learning System (NRLS) in CQC website
Dates and number of inspections since September 2012	CQC website
Type of inspection	CQC inspection report
Ratings given during the first inspection under the new regime	CQC website
Financial performance during quarter of inspection	Monitor or Trust Development Authority
Mergers	Monitor
Enforcement actions (special measures)	Monitor
Deprivation levels of the area served by the main commissioner	NHS Digital for deprivation indexes and annual accounts for main commissioner
Monthly numbers of staff	NHS Digital

To mitigate possible time-series error term correlation, we introduced auto-correlation structures from 1 to 12 and compared different models using a likelihood ratio test, reporting the most parsimonious model. A cosine function in the pressure ulcer model accounts for seasonal changes.¹⁹ To test robustness of findings, we estimated models for the whole sample and for a sub-sample matched on foundation trust status, reporting culture, and special measures status (See Appendix 1).

To assess clinical/policy significance (as well as statistical significance) we used 95% confidence intervals for rate of change before and after announcements and inspections, and change in levels the month after each interruption, with p-values <0.05 considered statistically significant. Analyses were performed in Stata SE/14.0 (StataCorp LP. 2015. College Station, Texas).

Results

Missing NHS ST data meant five trusts were excluded from the analysis. The 38 uninspected trusts had a median time to an inspection before or after the observation period of five (interquartile range 2-7) months. Trusts inspected under the new regime were less likely to be Foundation Trusts (53% versus 71%, $p=0.02$), more likely to be under special measures (28% versus 1%, $p<0.001$) and more likely to be rated as poor or with significant concerns about reporting culture (67% versus 42%, $p=0.002$) (Table 2). Figure 1 shows the distribution of inspections over the observation period. The mean rates of falls with harm and pressure ulcers in the observation period are shown in Figure 2.

Table 2 Characteristics of acute NHS Trusts included in the comparison of old and new regime of CQC's inspections.

	Old regime n=46	New regime n=66	No inspection n=38	p-value
Foundation Trust	31 (67%)	35 (53%)	29 (76%)	0.047
Type of trust				
Large acute trust	14 (30.4%)	18 (27%)	9 (24%)	0.80
Medium acute trust	14 (30.4%)	19 (29%)	12 (32%)	
Small acute trust	7 (15.2%)	13 (19.7%)	6 (16%)	
Acute teaching trust	5 (11%)	13 (19.7%)	7 (18%)	
Acute specialist trust	6 (13%)	3 (4.6%)	4 (10%)	
NHS England region				
North	16 (35%)	17 (26%)	16 (42%)	0.48
Midlands and East	15 (33%)	18 (27%)	10 (26%)	
South	9 (19%)	18 (27%)	9 (24%)	
London	6 (13%)	13 (20%)	3 (8%)	
Beds, median (IQR)	658.5 (461, 1002)	745 (562, 1020)	742 (491, 1024)	0.68
Population in thousands, median (IQR)	465 (325.7, 610)	460 (350, 600)	450 (320, 600)	0.98
Special measures	0 (0%)	18 (27%)	1 (3%)	<0.001
Overall rating after first inspection under new regime				
Outstanding	3 (7%)	2 (3%)	3 (8%)	0.85
Good	12 (27%)	16 (24%)	11 (30%)	
Requires improvement	26 (59%)	41 (62%)	21 (57%)	
Inadequate	3 (7%)	7 (11%)	2 (5%)	
Merged trusts	2 (4%)	8 (12%)	3 (8%)	0.35
Number of CQC inspections, median (IQR)	4 (3, 5)	4 (3, 5)	3 (2, 4)	0.004
Months since previous inspection, median (IQR)	10 (7, 13)	11 (8, 14)	-	0.39
CQC reporting culture rating 2016				
Poor	6 (13%)	17 (25.7%)	2 (5%)	0.028
Significant Concerns	14 (30%)	27 (41%)	13 (34%)	
Good	22 (48%)	21 (31.8%)	19 (50%)	
Outstanding	4 (9%)	1 (1.5%)	4 (11%)	

* Values are mean \pm SD, P₅₀ (P₂₅-P₇₅) or n (%).

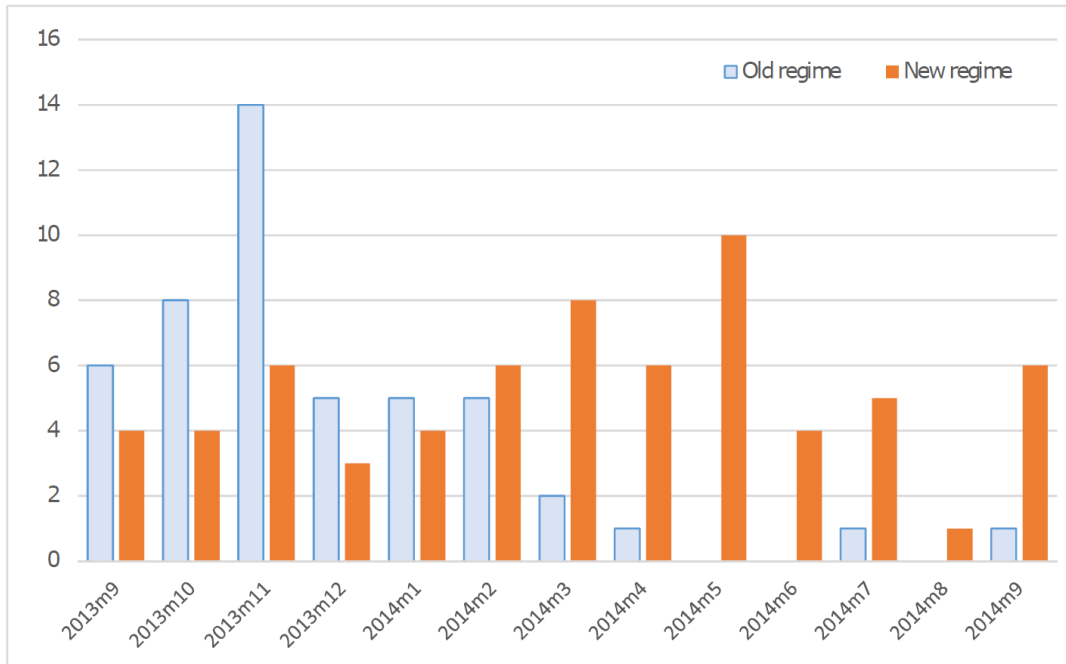


Figure 1 Frequency distribution of old and new CQC regime inspections

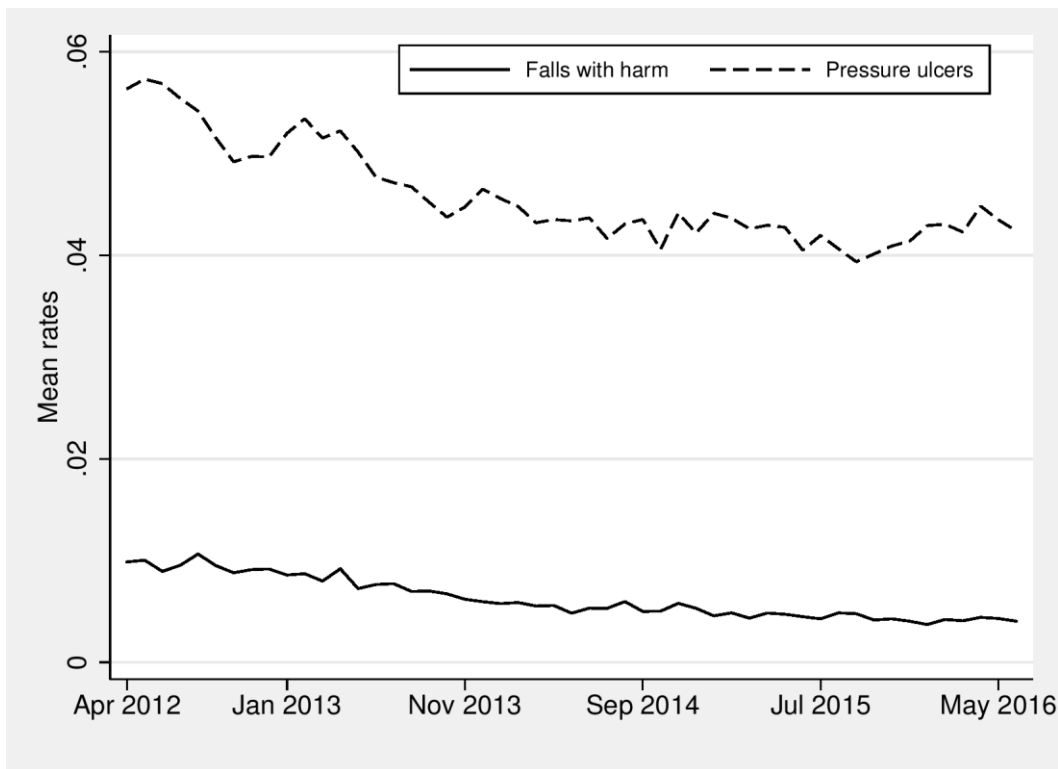


Figure 2 Mean rates of falls with harm and pressure ulcers during the observation period.

In the following section, figures show the trend for an average trust in each group for the three periods under study. Tables show the absolute values for the slopes and level changes observed. The model coefficients are reported in the Appendix 2.

Pressure Ulcers

The mean (95% CI) prevalence of pressure ulcers in April 2012 was 575 (470 to 679) for the historical inspection trusts, 590 (500 to 679) for the new inspection regimen trusts and 503 (407 to 600) for the uninspected group per 10,000 patients/month.

Rates of pressure ulcers were improving before CQC inspection in all three groups of trusts (Table 3, second column). Announcing the inspection had no immediate effect in the new regime group and significantly decreased rates in the historical regime group (-4.6 [-34.9 to 25.8] vs -47.4 [-81.9 to -12.9] pressure ulcers per 10,000 patients/month, see Table 3, third column). The rate of improvement remained similar to the pre-inspection period (Table 3, fifth column). Post-inspection, no effect was seen on the level or trend for any of the three groups (Figure 3). No significant changes were observed in the matched subsample (see Appendix 1).

Falls with harm

The mean (95% CI) prevalence of falls with harm in April 2012 was 90 (57 to 124) for the historical regime trusts, 108 (75 to 142) for new regime trusts and 90 (59 to 120) for uninspected trusts per 10,000 patients/month.

Rates of falls with harm were improving pre-inspection for all groups (Table 3, second column). Announcing the inspection was followed by non-significant changes in trend in all groups (Table 3, fifth column). The rate of falls with harm in new regime trusts was significantly lower than the one for the historically inspected trusts (-7.5 [-15.2 to 0.15] vs 8.8 [-2.9 to 20.4] falls per 10,000 patients) post-inspection announcement (Table 3, third column).

CQC inspections were associated with a slower improvement in rates of falls with harm in all groups (Figure 4 and Table 3, sixth column). No significant changes were observed in the matched subsample (see Appendix 1).

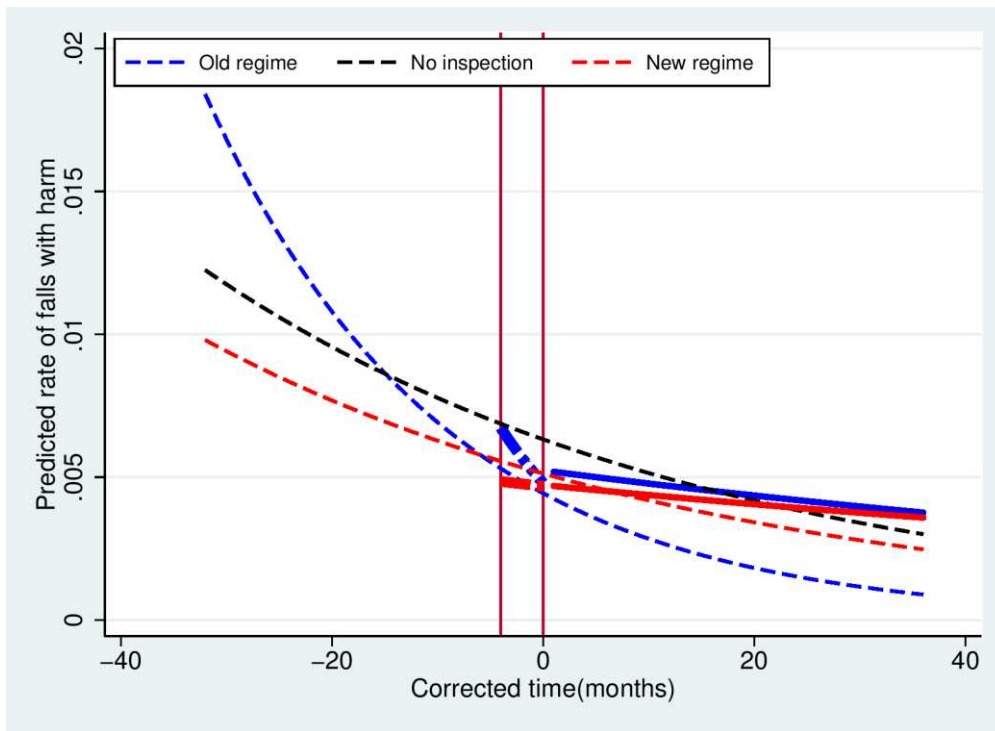


Figure 3 Predicted falls with harm.

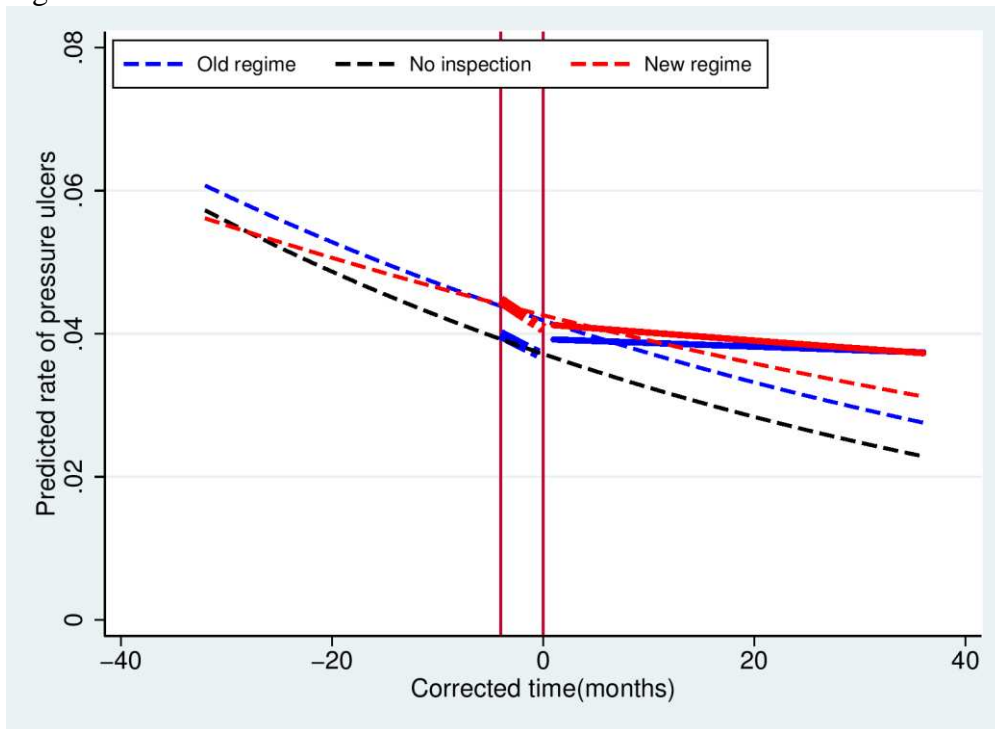


Figure 4 Predicted pressure ulcers.

Figure 3 and 4 show values for an average acute NHS trust before, after the announcement and after inspection by inspection group. Lines represent the mean rate of pressure ulcers obtained from the multilevel model for each period. The first vertical line signals the announcement, while the second the inspection. Dash lines show the trends observed during the pre inspection period, which were extended for the post announcement and post inspection periods. Long-dash lines represent the observed trends in the period between the announcement and the inspection. Solid lines show the trends after the inspection. Data for the hospitals inspected by the old regime are shown in blue, new regime in red and hospitals not inspected in black.

Table 3 Predicted changes in the slope and level of falls with harm and pressure ulcers before and after the announcement and inspection.

		Baseline rates	Pre-inspection slope	Change in level		Slope	
				Post-Announcement	Post-inspection	Post-announcement	Post-inspection
Pressure ulcers Trusts=150 Obs=7,291	New regime	561.4 (492.6, 640)	-4.8 (-8.9, -0.7)	-4.6 (-34.9, 25.8)	4.7 (-23, 32.3)	-9.9 (-23.7, 4)	-1.2 (-2.9, 0.5)
	Old regime	607.2 (504.3, 731.1)	-7 (-13.5, -0.6)*	-47.4 (-81.9, -12.9)	23 (-7.7, 53.7)	-7.6 (-22.5, 7.4)	-0.5 (-2.3, 1.3)
	No inspection	572.6 (464.1, 706.5)	-7.7 (-14.7, -0.6)	25.9 (-12.1, 63.9)	10.4 (-24, 44.8)	-12 (-29.6, 5.6)	-2.4 (-4.4, -0.4)
Falls with harm Trusts=149 Obs=6,255	New regime	98 (75.9, 126.4)	-2 (-3.4, -0.5)	-7.5 (-15.2, 0.15)*	0.09 (-6.8, 7)	-0.43 (-3.9, 3)	-0.4 (-0.7, -0.05)
	Old regime	184 (128.5, 263.4)	-8 (-13.4, -2.6)*	8.8 (-2.9, 20.4)	3.9 (-5, 12.8)	-5 (-10.9, 0.9)	-0.5 (-0.9, -0.08)
	No inspection	101.7 (77.4, 133.7)	-2.5 (-5.4, 0.4)	-7.2 (-20.1, 5.8)	-6.8 (-18.1, 4.5)	-0.3 (-6, 5.5)	-0.6 (-1, -0.1)

Values presented in number of events per 10,000 patients/month. Figures were calculated using model coefficients to present the absolute instead of relative value. For example, the pre-inspection slope for hospitals inspected by the new regime is calculated as $\exp(\beta_{0j} + \beta_{1j} * \text{corrected time}_{ij} + \beta_{6j} * \text{new inspection}_j + \beta_{7j} * \text{corrected time new inspection}_{ij})$

The estimations of the pre-inspection slope comprise a median (min-max) of 16 (4 - 26) data points per trust. Estimations for the post-announcement phase comprise four data points per trust, while for the post-inspection period a median (min-max) of 31 (21-34) data points were used. Only two trusts had fewer than ten observations for the pre-inspection period.

* Significant values at 0.05 level

Model for pressure ulcers was adjusted for specialist trusts, while model for falls with harm was adjusted for type of trust.

Discussion

CQC inspections were not associated with changes in the rates of two markers of quality: falls with harm and pressure ulcers. Any improvements made prior to the CQC inspection were diminished after inspection in an average trust. Historical and new regimes had similar effects, as did trusts not inspected at all.

Strengths and limitations

In the absence of a definitive trial, controlled interrupted times-series is a robust design to assess the impact of health policy.²⁰ Our analyses included most English acute trusts and a reliable time-series before and after inspections. Intervention points differed between trusts in our observation period (12 months), reducing confounding from historical events. The lack of effect in our matched – by potential confounders – group comparison increases confidence in our estimations.

As with all research, there are some limitations. Firstly, statistical phenomena could explain adverse events trends. With high rates, any intervention might reduce them due to regression to the mean. Low rates may remain stable, as any possibility of improvement is reduced due to the floor effect.²¹ Post-inspection trend changes could be an artefact of the interruption point chosen. Since adverse events vary randomly, modelling could artificially change levels or trends.

The NHS ST has some strengths: it is publicly available, conducted monthly, nationally, uses consistent definitions and standards, facilitates real-time reporting and rapid trust responses.¹⁷ However, it relies on self-reporting; consequently, variable data entry skills, flexible interpretations of operational definitions and variations in patient acuity within and between trusts can make it less reliable, and prone to ‘gaming’.¹⁸ We are unaware of any published reliability assessment of the NHS ST, and so incorporating random variation into the estimation of effects was impossible.

Our choice to use NHS ST data for research purposes was pragmatic: hospitals themselves use this dataset to assess their safety levels and the effectiveness of safety interventions. CQC inspections and feedback are part of the approach trusts use to improve safety and quality. It follows, then, that trusts carrying out improvements will look for an impact on their NHS ST and adverse event rates.

Our statistical model estimates each hospital's trajectory compared with the average performance of inspected trusts. Confidence intervals suggest variability within each group. Intrinsic features of NHS trusts may explain where inspections have an effect and where they do not. In 2013, trusts took over responsibility (from two NHS agencies) for the NHS ST. Individual trusts' implementation of NHS ST may have differed in line with competing priorities and commitment to patient safety, diluting a national, coordinated effort and effects. Further subgroup and retrospective (qualitative) analysis of implementation could explore this possibility.

Events surrounding or during inspections may change the safety behaviour of NHS trusts and the outcomes seen. Examples of potential 'external biases' include: i) NHS financial constraints that could have limited resources dedicated to patient safety; ii) changes in clinical commissioning – whilst 17 (11%) of the trusts were inspected before clinical commissioning was fully implemented in 2013, this significant reform affected organizations for some time post implementation; and iii) several high-profile quality improvement initiatives^{10, 13} that happened alongside our time-series. Such external factors influence managerial and clinical priorities and behaviours and may contribute to our findings; however, any potential for bias was spread between all trusts.

Finally, the type of CQC inspection that trusts experienced was not random.²² Trusts inspected by the new regime were more likely to be in special measures. Any trends in this group may be partially explained by regression to the mean for the regulatory intervention experienced.

Strengths and limitations in relation to other studies

Falls with harm and pressure ulcers are nursing-sensitive outcomes¹⁵ influenced by the quantity and quality of nursing care.²³ The sensitivity of these indicators to CQC-style external inspections is equivocal. Implementing standard safety practices requested by accreditation agencies can be simultaneously associated with lower pressure ulcer rates²⁴ and an increase in level and trend of rate of falls after an accreditation visit.²⁵

Evidence regarding CQC and its predecessors is primarily qualitative.^{2, 14} This is the first study quantitatively assessing impact of CQC inspections of acute trusts. Walshe et al.¹⁴ evaluated the CQC's new inspection regime in its pilot phase, finding that whilst acute trusts saw inspection as a catalyst for resuming 'on-hold' improvement activities, inspections were time and resource-consuming.

Changes seen in adverse event levels post-inspection announcement suggests trusts may change their behaviour to respond to the demands of the regulator – an anticipatory effect. A recent study examining gaming effects surrounding external inspections of cleanliness in English hospitals supports this premise.²⁶

The CQC produces a detailed report about overall hospital performance after each inspection. For serious concerns about performance, an enforcement action is issued. However, when improvements are requested without enforcement, hospitals may be unclear regarding how best to proceed, timescales for implementation, their own accountability and standards expected.¹⁴ Any potential CQC improvement influence can be diluted and delayed with unclear guidance.

CQC inspections allow trusts to validate and showcase their services' quality.²⁷ CQC visits may confirm what providers already know of their deficiencies as well as outstanding areas of care.^{14, 27} NHS providers asked if the benefits of CQC inspections justified their costs, gave mixed responses: 38% agreed and 38% disagreed. Many felt the 'same benefits could be achieved through a more streamlined approach'.²⁷

Meaning of the study

Although CQC inspections were not associated with changes on rates of two adverse events, this does not mean CQC did not improve overall care quality. The existence of the CQC and the possibility of being subject to a rigorous and detailed inspection at relatively short notice doubtless provides an incentive for maintaining hospital quality and thus managerial reputation.

Alternative explanations for the trend observed may include the diversion of clinical teams and managers away from clinical practice and towards implementation of improvements achievable in the time available (approximately 20 weeks). CQC requires detailed information in advance and holds focus groups with senior staff; all requiring trust administrative time.

A change in safety reporting behaviours may also offer some explanation. Sari et al. suggest that approximately 56% of falls and 23% of pressure ulcers are reported to routine systems.²⁸ The possibility exists, therefore, that increases in falls with harm and pressure ulcers observed post inspection reflect an increase in incident reporting instead of the number of events. If true, CQC inspections might be driving improvements in fundamental processes necessary for quality improvement that are not detected in a short time span.

Bevan and Fasolo²⁹ argue that various models of quality governance exist at various times. Whilst a combination of top-down hierarchical approaches, target setting and appeals to reputation were the most effective, reputation alone can still be a powerful lever.²⁹ Reputation-based governance works when performance information is ranked, transparent, understandable by the public and when performance is followed up.²⁹ Current CQC arrangements may encourage transient organizational change, but, arguably, lack the essential components for effective reputation-based governance. Reports are publicly available but not easily understandable, hospitals receive a rating but no ranking, and accessible tabular or graphical methods for comparing performance over time are hard to find.

Future research

Our results suggest that the expectation of future inspection(s) may influence acute trust behaviour, but has not led to positive lasting change. With the power to issue enforcement actions and monitor remotely, the CQC might consider reducing the burdensome (and resource-intensive) method of comprehensive inspections. The optimal ‘dose’ of inspection required for efficient scrutiny of trusts is unknown.

Bevan and Hood⁸ suggest that informing institutions how and when their performance will be measured is an open invitation to game the monitoring system.⁸ Adding randomness to oversight (and evaluating the effects) could increase certainty that observed improvement is real. Several attempts to improve external oversight of the NHS have been deployed historically: diverse types of rating systems, risk stratification methods, and varied approaches to on-site visits. The sense of dissonance between oversight activities and ‘everyday work’ in trusts remains, and with it a limited ability to influence improvement.²⁷

Oversight of quality and safety has been in place for nearly 20 years within the NHS.³⁰ Each regulator has been criticized and then replaced with ever more complex solutions promising to remedy the ills of previous structures and institutions. If increasing regulation has not delivered improvement, then it may be time to consider reducing the administrative burden of inspection and regulation, and monitoring the costs and effects of doing so.

Future research should focus on the effect of external oversight on other outcomes of care, such as risk-adjusted mortality, waiting times and emergency readmissions, to increase certainty about the effects of these systems. The NHS needs evidence of the cost-effectiveness of CQC external oversight. Finally, exploring the relationship between the CQC rating systems and specific outcome measures may help establish the validity of the instruments used.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organization for the submitted work; no financial relationships with any organizations that might have an interest in the submitted work in the previous three years; Professor Thompson reports that from August 2013 to July 2016 he was a non-executive director of Leeds York Partnership Foundation Trust which was fully inspected twice by CQC during that time.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. However, ACA is funded through a Becas Chile scholarship programme (CONICYT, Chile). CT and KB receive funding as part of the NIHR CLAHRC Yorkshire and Humber. The views expressed are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, or the Department of Health.

References

1. Illingworth J. Is the NHS getting safer? 2015. The Health Foundation.
2. Sutherland K and Leatherman S. Regulation and quality improvement: a review of the evidence. Health Foundation, 2006.
3. World Health Organization. Quality and accreditation in health care services: a global review. 2003. Geneva.
4. Flodgren G, Gonçalves-Bradley DC and Pomey MP. External inspection of compliance with standards for improved healthcare outcomes. Cochrane Database of Systematic Reviews 2016. DOI: 10.1002/14651858.CD008992.pub3.
5. Care Quality Commission. How CQC regulates NHS and independent acute hospitals. Provider handbook. 2015. Newcastle upon Tyne: Care Quality Commission.
6. Care Quality Commission. Enforcement policy. 2015. Newcastle: Care Quality Commission.
7. Laffont J-J and Tirole J. A theory of incentives in procurement and regulation. MIT press, 1993.
8. Bevan G and Hood C. What's measured is what matters: targets and gaming in the English public health care system. Public administration 2006; 84: 517-538.
9. Robertson R, Walshe K, Boyd A, et al. Q03. Does regulation improve care quality? A framework for understanding the impact of regulation, applied to the Care Quality Commission's regulatory model. In: Health Services Research UK Conference Nottingham, 2017.
10. Francis R. Report of the Mid Staffordshire NHS Foundation Trust public inquiry: executive summary. The Stationery Office, 2013.
11. Care Quality Commission. Chief Inspector of Hospitals announces inspection plans. 2013.
12. Care Quality Commission. Essential standards of quality and safety: guidance about compliance. 2010. London.
13. Keogh B. Review into the quality of care and treatment provided by 14 hospital trusts in England: overview report. NHS, 2013.
14. Walshe K, Boyd A, Addicott R, et al. Evaluating the Care Quality Commission's *acute hospital regulatory model: final report*. 2014. Manchester Business School and The King's Fund.
15. Butler M, Collins R, Drennan J, et al. Hospital nurse staffing models and patient and staff-related outcomes. Cochrane Database of Systematic Reviews 2011. DOI: 10.1002/14651858.CD007019.pub2.
16. Madsen J. NHS Safety Thermometer: Patient Harms and Harm Free Care. 2014. England: Health and Social Care Information Centre.
17. Power M, Stewart K and Brotherton A. What is the NHS Safety Thermometer? Clinical Risk 2012; 18: 163-169. DOI: doi:10.1258/cr.2012.012038.
18. 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: 287-297. DOI: 10.1093/intqhc/mzu043.
19. Lopez Bernal J, Cummins S and Gasparrini A. Interrupted time series regression for the evaluation of public health interventions: a tutorial. International Journal of Epidemiology 2017; 46: 348-355. DOI: 10.1093/ije/dyw098.

20. Kontopantelis E, Doran T, Springate DA, et al. Regression based quasi-experimental approach when randomisation is not an option: interrupted time series analysis. *BMJ: British Medical Journal* 2015; 350. DOI: 10.1136/bmj.h2750.
21. Hessling RM, Schmidt TJ and Traxel NM. Floor effect. In: Lewis-Beck M, Bryman A and Futing Liao T, (eds.). *The SAGE Encyclopedia of Social Science Research Methods*. Thousand Oaks, California 2004.
22. Griffiths A, Beaussier AL, Demeritt D, et al. Intelligent Monitoring? Assessing the ability of the Care Quality Commission's statistical surveillance tool to predict quality and prioritise NHS hospital inspections. *BMJ quality & safety* 2017; 26: 120-130. DOI: 10.1136/bmjqs-2015-004687.
23. Griffiths P, Ball J, Drennan J, et al. Nurse staffing and patient outcomes: Strengths and limitations of the evidence to inform policy and practice. A review and discussion paper based on evidence reviewed for the National Institute for Health and Care Excellence Safe Staffing guideline development. *International Journal of Nursing Studies* 2016; 63: 213-225. DOI: <http://dx.doi.org/10.1016/j.ijnurstu.2016.03.012>.
24. Thornlow DK and Merwin E. Managing to improve quality: the relationship between accreditation standards, safety practices, and patient outcomes. *Health care management review* 2009; 34: 262-272. 2009/07/25. DOI: 10.1097/HMR.0b013e3181a16bce.
25. Devkaran S and O'Farrell PN. The impact of hospital accreditation on quality measures: an interrupted time series analysis. *BMC health services research* 2015; 15: 137. 2015/04/19. DOI: 10.1186/s12913-015-0784-5.
26. Toffolutti V, McKee M and Stuckler D. Evidence Points To 'Gaming' At Hospitals Subject To National Health Service Cleanliness Inspections. *Health Affairs* 2017; 36: 355-361. DOI: 10.1377/hlthaff.2016.1217.
27. NHS Providers. *Regulation Survey Report 2017: Impact of regulation in a shifting environment*. 2017. NHS Providers.
28. Sari AB-A, Sheldon TA, Cracknell A, et al. Sensitivity of routine system for reporting patient safety incidents in an NHS hospital: retrospective patient case note review. *BMJ (Clinical research ed)* 2006. DOI: 10.1136/bmj.39031.507153.AE.
29. Bevan G and Fasolo B. Models of governance of public services: empirical and behavioural analysis of 'econs' and 'humans'. In: Angus and Oliver (eds) *Behavioural public policy*. Cambridge, UK: Cambridge University Press, 2013, pp.38-62.
30. Bevan G. Regulation and system management. In: Mays N, Dixon A and Jones L (eds) *Understanding New Labour's market reforms of the English NHS*. King's Fund, 2011, pp.208.