A qualitative and quantitative evaluation of the Advancing Quality pay-for-performance programme in the NHS North West

Ruth McDonald,[1]* Ruth Boaden,[2] Martin Roland,[3] Søren Rud Kristensen,[4] Rachel Meacock,[4] Yiu-Shing Lau,[4] Tom Mason,[4] Alex J Turner[4] and Matt Sutton[4]

[1]Manchester Business School and Centre for Primary Care, University of Manchester, Manchester, UK
[2]Manchester Business School, University of Manchester, Manchester, UK
[3]Institute of Public Health, University of Cambridge, Cambridge, UK
[4]Manchester Centre for Health Economics, University of Manchester, Manchester, UK

*Corresponding author

Declared competing interests of authors: Dr Martin Roland has received personal fees from Advice to Brazilian Ministry of Health on pay for performance and personal fees from Advice to Singaporean Ministry of Health on pay for performance, outside the submitted work.

Disclaimer: The views and opinions expressed herein are those of the authors and do not necessarily reflect those of the National Institute for Health Research, the NHS or the Department of Health.

Published May 2015
DOI: 10.3310/hsdr03230

Scientific summary

Evaluation of the Advancing Quality pay-for-performance programme

Health Services and Delivery Research 2015; Vol. 3: No. 23
DOI: 10.3310/hsdr03230

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

A wide variety of pay-for-performance (P4P) schemes have been developed for health-care providers. Such schemes are being increasingly adopted internationally with the aim of improving care quality. However, increased adoption of P4P is occurring despite a scant evidence base.

Advancing Quality (AQ) is a voluntary programme which provides financial incentives to health-care providers for improvement in the quality of care provided to NHS patients. It has been implemented in the North West region of England since 2008. Initially, quality of care was measured by clinical process and outcome measures in five clinical areas – acute myocardial infarction, heart failure, coronary artery bypass graft, pneumonia and hip and knee replacement. Subsequently, the programme expanded to include additional clinical areas, but these do not form part of this evaluation.

The AQ programme evaluation was undertaken over 5 years from 1 April 2009.

Objectives

The study objectives were to:

(a) identify the impact of AQ on key stakeholders (provider organisations, commissioners and patients) and clinical practice
(b) assess the cost-effectiveness of AQ
(c) identify key factors that assist or impede the successful implementation of AQ
(d) provide lessons for the wider implementation of P4P schemes across the NHS as a whole.

Methods

The study used a combination of qualitative and quantitative methods. We assessed the impact of AQ in quantitative terms using national data on mortality, readmissions and length of stay from Hospital Episode Statistics. This component helped us understand what happened. We tested whether or not the incentives had an impact on mortality using two methods: a between-region difference-in-differences analysis comparing changes in mortality over time between the North West region of England and the rest of England for the incentivised conditions and a triple-difference analysis comparing the changes in mortality over time between the incentivised conditions in the North West region and the rest of England with the changes in mortality over time between the North West region and the rest of England for the non-incentivised conditions. In addition, a cost-effectiveness analysis of AQ based on the first 18 months of the programme was also undertaken.

This quantitative analysis was combined with qualitative data collection and analysis aimed to shed light on how and why these impacts occurred. During the first 3 years of our 5-year evaluation we conducted interviews ($n = 391$) with relevant NHS staff and observations ($n = 52$) of meetings and events. During the final 2 years, we interviewed at least one member of staff from each participating provider organisation and 11 commissioners.

Results

Our assessment of impact found that AQ was associated with significant reductions in patient mortality during the first 18 months of the programme (Sutton M, Nikolova S, Boaden R, Lester H, McDonald R, Roland M. Reduced mortality with hospital pay for performance in England. *N Engl J Med* 2012;**367**:1821–8). Risk-adjusted mortality rates for all three of the conditions we studied (pneumonia, heart failure and myocardial infarction) decreased over the study period in both the North West region and the rest of England. The reduction in mortality for incentivised conditions was greater in the North West region than in the rest of England, reducing from 21.9% to 20.1% in the North West region and from 20.2% to 19.3% in the rest of England. Compared with non-incentivised conditions within the North West region (within-region difference-in-differences analysis), there was a significant reduction in overall mortality for incentivised conditions of 0.9 percentage points [95% confidence interval (CI) 0.1 to 1.7 percentage points], comprising a statistically significant reduction in pneumonia and non-significant reductions in the other two conditions. Comparing mortality for the incentivised conditions with mortality for the same conditions in other regions, there was again a significant reduction in overall mortality in the North West region of 0.9 percentage points (95% CI 0.4 to 1.4 percentage points), which was also made up of individually significant reductions in pneumonia and non-significant reductions in the other two conditions. Combining these two suggested an overall reduction in mortality of 1.3 percentage points in the North West region (95% CI 0.4 to 2.1 percentage points), with a similar pattern for the individual conditions. The reduction in mortality over the 18-month period studied for non-incentivised conditions was not significantly different between the North West region and the rest of England.

Based on the first 18 months, we found AQ to be a cost-effective use of resources. The total cost of the AQ programme was just over £13M over the initial 18-month period, with only £5M of this consisting of the financial incentives. The ongoing running costs of the scheme exceeded the bonus payments, making up the majority of the costs at just over £7M. We estimated a gain of 6700 quality-adjusted life-years (QALYs) as a result of the reduction in mortality for the programme as a whole. At a QALY value of £20,000, this equals an estimated health gain worth £134M. Our estimates suggest that AQ also resulted in a reduction of 22,700 bed-days in the first 18 months. This is equivalent to a £5M reduction in costs.

The average performance reported by the participating hospitals on all of the quality measures improved in the first 18 months and improved further in the following 24 months, particularly for heart failure and pneumonia. Some of the process quality measures were significantly associated with better health outcomes at a trust level but the magnitudes of the estimated coefficients were too large to represent clinically plausible direct consequences of these process measures. The findings suggest that these financial incentives only weakly led to improved patient outcomes through their direct effects on the process measures that were incentivised.

Advancing Quality appears to have also led to improved patient outcomes by inducing positive spillover effects in terms of wider improvements in care quality across unmeasured dimensions and improvements in care for all patients. Our qualitative data provide support for this explanation, highlighting developments at sites (e.g. recruitment of specialist nurses to join up gaps in care and maintain a sustained focus on patients as they moved through the hospital) to improve care quality for patients in AQ clinical areas. They also suggest that clinician compliance with data-recording requirements varied between clinicians and across sites. Performance on process measures reflects what is recorded as opposed to the care that was delivered, and failure to record care delivery in a systematic fashion was a persistent problem. This further complicates the issue of quantifying relationships between performance on process measures and the relevant outcomes.

When we looked over the longer-term period from 18 to 42 months, risk-adjusted mortality rates continued to decrease in both the North West region and the rest of England for both incentivised and non-incentivised conditions. The between-region difference-in-differences analysis showed that risk-adjusted mortality for the incentivised conditions fell by 2.3 percentage points in the rest of England

© Queen's Printer and Controller of HMSO 2015. This work was produced by McDonald *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

and by 1.8 percentage points in the North West region. This reduction in the rest of England was significantly larger (0.7 percentage points, 95% CI 0.3 to 1.2 percentage points) than in the North West region, and was concentrated in pneumonia (1.1 percentage points, 95% CI 0.4 to 1.8 percentage points). However, the reductions in mortality were also larger for the non-incentivised conditions in the North West region than in the rest of England between these periods (1.2 percentage points more, 95% CI 0.4 to 2.0 percentage points).

We considered various explanations for the smaller reduction in mortality for the incentivised conditions in the North West region in the long term (i.e. at 42 months) compared with the rest of England. The first is the possibility that the scheme became less effective with the change in incentive structure, as the AQ programme switched from a tournament scheme with bonuses to a scheme involving penalties for failure to reach quality benchmarks. The continued improvement in performance on incentivised process measures in the AQ hospitals suggests that the incentives may still have been effective, but we have no data from control hospitals for these measures. Moreover, as described previously, we did not find a significant relationship between performance on process measures and outcomes.

A second possible explanation is that there was a positive spillover effect from the adopting region (i.e. the North West region) to other regions. The early results of AQ had been widely disseminated in England and two other regions had adopted a form of AQ programme incentives. These regions showed a greater reduction in mortality in the long term compared with other control regions that did not incentivise the AQ indicators, although the reduction was statistically significant only for acute myocardial infarction.

We also found limited evidence for positive spillover effects within the AQ hospitals, as the patients with non-incentivised conditions that were treated by specialists who also treated patients with incentivised conditions experienced the largest reductions in mortality in the long term.

A number of factors appeared to contribute to the success (as measured by improving performance on process measures and mortality at 18 months) of the scheme. These include in-person collaborative learning events, dedicated infrastructure support, financial rewards to invest in additional staff and a combination of competition to spur improvement and collaboration to facilitate learning. Additionally, programme participants were able to contribute to shaping the programme as it evolved, enhancing legitimacy and buy-in.

At the same time, there were a number of barriers to implementation. In the context of heavy workloads and competing priorities, frontline staff did not always adhere to AQ requirements. Furthermore, data collection was burdensome in a context in which AQ was not part of existing electronic patient information systems. AQ did not become institutionalised and embedded into routine behaviours. Instead, there was a reliance on core AQ staff to cajole and persuade other staff members, which often resulted in going around obstacles rather than resolving enduring problems. Although there were some common themes in the approach taken (in particular, the employment of specialist nurses), more generally, hospitals implemented AQ using a range of activities tailored to and developed in their local context. This suggests that there was no one blueprint for implementing AQ in each site.

In terms of impact on commissioners, input from staff in commissioning organisations was relatively limited in the first year of AQ. Although some commissioner staff had begun to engage with AQ by year 2, the subsequent reorganisation of NHS commissioning functions during the study period meant that input from commissioners was limited or non-existent for most of the study period.

The AQ scheme design incorporated features of what the literature identifies as good practice. It did not involve penalties and it rewarded relative, as well as absolute, performance. The fact that participation was on a voluntary basis and was universal (i.e. all 24 eligible organisations took part) appeared to add to the legitimacy of the AQ programme. Additionally, the competitive nature of the scheme did not crowd out knowledge sharing and collaboration more generally. However, our findings, which highlight

implementation challenges and a failure to embed change in routine practice, suggest that, although scheme design is important, there are other aspects relating to implementation that require attention if financial incentive schemes are to fulfil and maintain their potential.

Conclusions

Based on the first 18 months, AQ was a relatively cost-effective intervention. The findings after 42 months are open to several interpretations. Our failure to find a relationship between process and outcome measures at 18 months suggests that there were positive effects beyond the changes in the specific AQ measures. An alternative interpretation, however, is that short-term improvements were not sustained and that the observed improvements in mortality in the non-incentivised conditions within hospitals participating in AQ were unrelated to the programme.

The first explanation is supported by changes to care delivery identified by our evaluation. It may be that there were further positive spillover effects in quality of care both from participating to non-participating hospitals and from incentivised to non-incentivised conditions in the participating hospitals. We found some modest evidence for both of these hypotheses. However, we did not explicitly focus on non-incentivised conditions. Furthermore, because we collected qualitative data from a large number of sites ($n = 24$), we were unable to conduct detailed, in-depth research to explore these issues comprehensively.

Further research to investigate the relationship between AQ and changes in incentivised and non-incentivised conditions would shed light on this area. Linked to this, research exploring changes in rest-of-England sites would also add to our knowledge.

The study highlights the importance of considering costs beyond the incentive payments of financial incentive programmes intended to improve care quality. It also suggests that competition did not inhibit collaboration, with providers keen to share learning within the AQ community of practice. Instead, cohesive network relationships appeared to support the social enforcement of anticompetitive norms. In-person collaborative learning events were an important part of building and sustaining such relationships.

We found no evidence of changes in care resulting from AQ being institutionalised. Instead, modifications to practice were generally not systematised and behaviour change was still largely reliant on prompting by particular individuals. The success of AQ seems to have been a result of persistent and focused individuals working to remind staff and to plug gaps in data collection and/or care pathways. Furthermore, far from being everybody's business and part of organisation-wide change, AQ was delivered in a context in which many staff were unaware of its existence. Further research should be undertaken to determine the explanation for the findings.

Funding

The National Institute for Health Research Health Services and Delivery Research programme.

© Queen's Printer and Controller of HMSO 2015. This work was produced by McDonald *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Health Services and Delivery Research

ISSN 2050-4349 (Print)

ISSN 2050-4357 (Online)

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: nihredit@southampton.ac.uk

The full HS&DR archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hsdr. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Services and Delivery Research* journal

Reports are published in *Health Services and Delivery Research* (HS&DR) if (1) they have resulted from work for the HS&DR programme or programmes which preceded the HS&DR programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

HS&DR programme

The Health Services and Delivery Research (HS&DR) programme, part of the National Institute for Health Research (NIHR), was established to fund a broad range of research. It combines the strengths and contributions of two previous NIHR research programmes: the Health Services Research (HSR) programme and the Service Delivery and Organisation (SDO) programme, which were merged in January 2012.

The HS&DR programme aims to produce rigorous and relevant evidence on the quality, access and organisation of health services including costs and outcomes, as well as research on implementation. The programme will enhance the strategic focus on research that matters to the NHS and is keen to support ambitious evaluative research to improve health services.

For more information about the HS&DR programme please visit the website: http://www.nets.nihr.ac.uk/programmes/hsdr

This report

The research reported in this issue of the journal was funded by the HS&DR programme or one of its preceding programmes as project number 08/1809/250. The contractual start date was in March 2009. The final report began editorial review in May 2014 and was accepted for publication in November 2014. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HS&DR editors and production house have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the final report document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HS&DR programme or the Department of Health. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HS&DR programme or the Department of Health.

© Queen's Printer and Controller of HMSO 2015. This work was produced by McDonald *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

Health Services and Delivery Research Editor-in-Chief

Professor Ray Fitzpatrick Professor of Public Health and Primary Care, University of Oxford, UK

NIHR Journals Library Editor-in-Chief

Professor Tom Walley Director, NIHR Evaluation, Trials and Studies and Director of the HTA Programme, UK

NIHR Journals Library Editors

Professor Ken Stein Chair of HTA Editorial Board and Professor of Public Health, University of Exeter Medical School, UK

Professor Andree Le May Chair of NIHR Journals Library Editorial Group (EME, HS&DR, PGfAR, PHR journals)

Dr Martin Ashton-Key Consultant in Public Health Medicine/Consultant Advisor, NETSCC, UK

Professor Matthias Beck Chair in Public Sector Management and Subject Leader (Management Group), Queen's University Management School, Queen's University Belfast, UK

Professor Aileen Clarke Professor of Public Health and Health Services Research, Warwick Medical School, University of Warwick, UK

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Peter Davidson Director of NETSCC, HTA, UK

Ms Tara Lamont Scientific Advisor, NETSCC, UK

Professor Elaine McColl Director, Newcastle Clinical Trials Unit, Institute of Health and Society, Newcastle University, UK

Professor William McGuire Professor of Child Health, Hull York Medical School, University of York, UK

Professor Geoffrey Meads Professor of Health Sciences Research, Faculty of Education, University of Winchester, UK

Professor John Powell Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK

Professor James Raftery Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

Dr Rob Riemsma Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

Professor Helen Roberts Professor of Child Health Research, UCL Institute of Child Health, UK

Professor Helen Snooks Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

Please visit the website for a list of members of the NIHR Journals Library Board:
www.journalslibrary.nihr.ac.uk/about/editors

Editorial contact: nihredit@southampton.ac.uk