EVALUATION OF THE ADVANCING QUALITY PAY FOR PERFORMANCE PROGRAMME IN THE NHS NORTH WEST

Report October 2014

Ruth McDonald, Manchester Business School and Centre for Primary Care, University of Manchester[1]
Ruth Boaden, Manchester Business School, University of Manchester
Martin Roland, Institute of Public Health, University of Cambridge
Søren Rud Kristensen, Manchester Centre for Health Economics, University of Manchester
Rachel Meacock, Manchester Centre for Health Economics, University of Manchester
Yiu-Shing Lau, Manchester Centre for Health Economics, University of Manchester
Tom Mason, Manchester Centre for Health Economics, University of Manchester
Alex J Turner, Manchester Centre for Health Economics, University of Manchester
Matt Sutton, Manchester Centre for Health Economics, University of Manchester

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

[1] Corresponding author. ruth.mcdonald@mbs.ac.uk, Manchester Business School & Centre for Primary Care, Room F2, MBS East, Booth Street East, M15 6PB, UK.

Acknowledgements

We acknowledge the contribution of our dear colleague Dr Helen Lester who died. Helen was a valuable and productive member of the team from the outset of the evaluation until her death in March 2013.

We are also extremely grateful to all of the people who gave their time generously to assist in our study.

This project was funded by the NIHR HS&DR 08/1809/250 and we are grateful for this funding, without which the study would not have been possible.

Authors' contributions

Ruth McDonald was the principal investigator. She was involved in the design of the study, the collection of data in phases one and two, data analysis and contributing to the final report writing.

Ruth Boaden was involved in designing the research project, contributing to the analysis of qualitative data and to the writing of the final report.

Martin Roland was involved in designing the research project, contributing intellectual input with regard to analysis and interpretation of data as well as bringing a clinical perspective. He also contributed to the writing of the final report.

Søren Rud Kristensen was involved in the analysis of quantitative data and contributing to the writing of the final report.

Rachel Meacock was involved in the analysis of quantitative data and contributing to the writing of the final report.

Yiu-Shing Lau was involved in the analysis of quantitative data and contributing to the writing of the final report.

Tom Mason was involved in the analysis of quantitative data and contributing to the writing of the final report.

Alex J Turner was involved in the analysis of quantitative data and contributing to the writing of the final report.

Matt Sutton took lead responsibility for the quantitative components of the study. He was involved in designing the research project, analysing data and contributing to the writing of the final report.

Disclaimer

The views and opinions expressed herein are those of the authors and do not necessarily reflect those of the NIHR, the NHS or the Department of Health.

Important

A 'first look' scientific summary is created from the original author-supplied summary once the normal NIHR Journals Library peer and editorial review processes are complete. The summary has undergone full peer and editorial review as documented at NIHR Journals Library website and may undergo rewrite during the publication process. The order of authors was correct at editorial sign-off stage.

A final version (which has undergone a rigorous copy-edit and proofreading) will publish as part of a fuller account of the research in a forthcoming issue of the *Health Services and Delivery Research* journal.

Any queries about this 'first look' version of the scientific summary should be addressed to the NIHR Journals Library Editorial Office NIHRedit@soton.ac.uk.

The research reported in this 'first look' scientific summary was funded by the HS&DR programme or one of its predecessor programmes (NIHR Service Delivery and Organisation programme, or Health Services Research programme) as project number 08/1809/250. For more information visit http://www.nets.nihr.ac.uk/projects/hsdr/081809250

The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HS&DR editors have tried to ensure the accuracy of the authors' work and would like to thank the reviewers for their constructive comments however; they do not accept liability for damages or losses arising from material published in this scientific summary.

This 'first look' scientific summary 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.

Scientific Summary

Background

A wide variety of pay for performance (P4P) schemes has 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 is a voluntary programme which provides financial incentives for improvement in the quality of care provided to NHS patients. It has been implemented in the North West 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 Advancing Quality programme evaluation was undertaken over five years, from 1st April 2009.

Objectives

The study objectives were to

a. Identify the impact of Advancing Quality on key stakeholders (provider organisations, commissioners and patients) and clinical practice

b. Assess the cost-effectiveness of Advancing Quality

c. Identify key factors which assist or impede the successful implementation of Advancing Quality

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 Advancing Quality 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 the incentives had an impact on mortality in two ways: a between-region difference-in-differences analysis comparing changes in mortality over time between the North West 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 and the rest of England with the changes in mortality over time between the North West and the rest of England for the non-incentivised conditions. In addition, a cost effectiveness analysis of Advancing Quality based on the first 18 months of the programme was also undertaken.

This quantitative analysis was combined with qualitative data collection and analysis aimed at shedding light on 'how' and 'why' these impacts occurred. During the first three years of our five year evaluation we conducted interviews (n=391) with relevant NHS staff and observations (n=52) of meetings and events. During the final two years, we interviewed at least one member of staff from each participating provider organisation and eleven commissioners.

Results

Our assessment of impact found that Advancing Quality was associated with significant reductions in patient mortality during the first 18 months of the programme (Sutton et al, 2012). 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 and the rest of England. The reduction in mortality for incentivised conditions was greater in the North West than in the rest of England, reducing from 21.9% to 20.1% in the North West and from 20.2% to 19.3% in the rest of England. Compared to non-incentivised conditions within the North West (within-region difference-in-differences analysis), there was a significant reduction in overall mortality for incentivised conditions

of 0.9 percentage points (CI 0.1 to 1.7), comprising a statistically significant reduction for pneumonia and a non-significant reduction for 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 of 0.9 percentage points (CI 0.4 to 1.4), again made up of individually significant reductions for pneumonia and non-significant reductions for the other two conditions. Combining these two suggested an overall reduction in mortality of 1.3 percentage points in the North West (CI 0.4 to 2.1), 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 and the rest of England.

Based on the first 18 months, we found Advancing Quality to be a cost-effective use of resources. The total cost of the Advancing Quality programme was just over £13million over the initial 18 month period, with only £5million 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 £7million. We estimated a gain of 6,700 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 £134 million. Our estimates suggest that Advancing Quality also resulted in a reduction of 22,700 bed-days in the first 18 months. This is equivalent to a £5 million 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 to improve quality 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 spillovers 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 Advancing Quality 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 relevant outcomes.

When we looked over the longer term from 18 to 42 months, risk-adjusted mortality rates continued to decrease in both the North West and the rest of England, for both incentivised and non-incentivised conditions. The between-region difference-in-differences analyses showed that risk-adjusted mortality for the incentivised conditions fell by 2.3 percentage points in the rest of England and 1.8 percentage points in the North West. This reduction in the rest of England was significantly larger (0.7 percentage points; CI 0.3 to 1.2) than in the North West, and was concentrated in pneumonia (1.1 percentage points; CI 0.4 to 1.8). However, the reductions in mortality were also larger for the non-incentivised conditions in the North West than in the rest of England between these periods (1.2 percentage points more, 95% CI 0.4 to 2.0).

We considered various explanations for the smaller reduction in mortality for the incentivised conditions in the North West in the long-term (i.e. at 42 months) compared to the rest of England. The first is the possibility that the scheme became less effective with the change in incentive structure, as the Advancing Quality 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 Advancing Quality hospitals suggests that the incentives may still have been effective, but we have no data from control hospitals for these measures. Moreover, as described above, 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 from the adopting region (i.e. the North West) to other regions. The early results of AQ had been widely disseminated in England and two other regions had adopted a form of Advancing Quality incentives. These regions showed a greater reduction in mortality in the long term compared to other control regions which did not incentivise the Advancing Quality indicators, though the reduction was only statistically significant for acute myocardial infarction.

We also found limited evidence for positive spillover effects within the Advancing Quality 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 Advancing Quality requirements. Furthermore, data collection was burdensome in a context where Advancing Quality was not part of existing electronic patient information systems. Advancing Quality did not become institutionalised and embedded into routine behaviours. Instead, there was a reliance on core Advancing Quality staff to cajole and persuade, 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 Advancing Quality using a range of activities tailored to and developed in their local context. This suggests that there was no one 'blueprint' for implementing Advancing Quality in each site.

In terms of impact on commissioners, input from staff in commissioning organisations was relatively limited in the first year of Advancing Quality. Although some commissioner staff had begun to engage with Advancing Quality by year two, 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 Advancing Quality 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 Advancing Quality's legitimacy. 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, whilst scheme design is important, there are other aspects relating to implementation which require attention if financial incentive schemes are to fulfil and maintain their potential.

Conclusions

Based on the first 18 months, Advancing Quality 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 Advancing Quality 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 Advancing Quality were unrelated to Advancing Quality.

The first explanation is supported by changes to care delivery identified by our evaluation. It may be that there were further positive spillovers 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, since 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 Advancing Quality 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 Advancing Quality 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 Advancing Quality 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 Advancing Quality seems to have been due to 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, Advancing Quality was delivered in a context where many staff were unaware of its existence.

2179 words

Plain English Summary

Background
In 2008 a scheme was introduced offering the potential to earn financial rewards by improving quality for NHS patients. All 24 eligible hospitals in the North West participated.

What we did
We talked to people involved in the scheme and observed them in meetings related to the scheme. We also measured the impact of the scheme by looking at whether it had made a difference to the death rate from various conditions. These were adjusted for risk, and known as 'risk adjusted mortality'.

What we found
After the first 18 months of the scheme, we found that there was a reduction in risk adjusted mortality for three clinical conditions included in the scheme. Although there was a reduction elsewhere, the reduction in the North West was larger.

After the first 42 months of the scheme, we found that the fall in risk adjusted mortality was greater in the rest of England than the North West. However, there was a fall in risk adjusted mortality in the North West for some clinical conditions, not included in the incentive scheme, which was greater than the rest of England.

What can we conclude?
One interpretation is that short term improvements we found after 18 months were not sustained and that the observed improvements in mortality in the non-scheme conditions at 42 months were unrelated to the scheme. An alternative interpretation, however, is that the incentive scheme led to positive benefits in other clinical conditions in the same hospital.