

**Evaluating the Labour Government's English NHS health system reforms:
the 2008 Darzi reforms**

Nicholas Mays

Professor of Health Policy

Department of Health Services Research and Policy

Faculty of Public Health and Policy

London School of Hygiene and Tropical Medicine

15-17 Tavistock Place

London WC1H 9SH

Nicholas.Mays@lshtm.ac.uk

July 2013

Abstract

Starting in 2002, the UK Labour government of 1997–2010 introduced a series of changes to the National Health Service (NHS) in England designed to increase individual patients' choices of the place of elective hospital care and encourage competition among public and private providers of elective hospital services for NHS-funded patients. In 2006, the Department of Health initiated the Health Reform Evaluation Programme (HREP) to assess the impact of the changes. In June 2008, the White Paper, *High quality care for all*, was published. It represented the Government's desire to focus the next phase of health care system reform in England as much on the quality of care as on improving its responsiveness and efficiency.

The 2008 White Paper led to the commissioning of a further wave of evaluative research under the auspices of HREP, as follows:

- an evaluation of the implementation and outcomes of care planning for people with long term conditions;
- an evaluation of the personal health budget pilots;
- an evaluation of the implementation and outcomes of the commissioning for quality and innovation (CQUIN) framework; and
- an evaluation of cultural and behavioural change in the NHS focused on ensuring high quality care for all.

This Supplement includes papers from each project.

The evaluations present a mixed picture of the impact and success of the 2008 reforms. All the studies identify some limitations of the policies in the White Paper. The introduction of personal health budgets appears to have been the least problematic and, depending on assumptions, likely to be cost-effective for the sorts of patients involved in the pilot. For the rest of the changes, impacts ranged from little or none (CQUIN and care planning for people with chronic conditions) to patchy and highly variable (instilling a culture of quality in acute hospitals) in the three years following the publication of the White Paper. On the other hand,

each of the studies identifies important insights relevant to modifying and improving the policies. These findings have continuing relevance since both the 2008 White Paper's policies, and the issues they were focused on remedying, remain central to the current coalition Government's reform agenda.

Evaluating the Labour Government's Health System Reforms: the 2008 Darzi reforms

The Health Reform Evaluation Programme and Labour's English NHS market reforms

Starting in 2002, the United Kingdom Labour government of 1997-2010 introduced a series of changes to the National Health Service (NHS) in England designed to increase individual patients' choices of the place of elective hospital care and encourage competition among public and private providers of elective hospital services for NHS-funded patients. In 2006, the Department of Health initiated a Health Reform Evaluation Programme (HREP) to assess the impact of these changes. The main findings from the HREP projects commissioned to evaluate the market-related reforms were reported in a previous *Journal of Health Services Research & Policy* supplement published in January 2012¹ and in a book published by the King's Fund.² Taken together, the HREP studies showed that the changes had had the effects that proponents had predicted, but the effects were mostly modest.³ Most of the undesirable impacts feared by critics appeared not to have materialized to any discernible extent, at least by early in 2010 when the projects in this wave of HREP came to an end. Labour's market appeared to have generated stronger incentives for quality and efficiency than its 1990s predecessor with no obvious detriment to equity of access.

The market-related reforms that began in the early 2000s had been a response to the perceived policy imperatives of the time, namely, to increase output, efficiency and quality in elective care. While these aspects of NHS performance will always be important, it became increasingly apparent to policy makers from the mid-2000s that any further NHS changes needed to contribute more directly to improving the ability of an increasingly constrained Service to support the needs of people with multiple, long term (i.e. chronic) conditions. Competition between suppliers of discrete health services to meet the demands of patient choice may have been appropriate when the focus was on increasing the volume and responsiveness of providers of elective surgery, but would have little to offer when the challenge was how to provide better services for people with long term conditions whose needs required joining up the efforts of diverse providers.

For example, the so called ‘payment by results’ (PbR) diagnosis-related group (DRG)-style payment system for NHS hospital care was increasingly seen as frustrating the effective management of people with complex, long term conditions by reimbursing individual providers for episodes of care rather paying for integrated ‘care pathways’ (i.e. ‘seamless’ patterns of care that span primary, community and acute sectors). The solution was seen as the introduction of more ‘bundled’ payments for whole pathways of care to reward providers for delivering more integrated forms of care as well as ‘unbundling’ some payments (e.g. to isolate the diagnostic element so that it could be provided without the need for a hospital visit) so as to allow services to be tendered individually or to allow new pathways to be created for tendering.

There was also a desire on the part of the Government to make it plain that NHS reform in England was as much about improving the quality of the Service as about improving its responsiveness and efficiency using business management techniques, patient choice and supplier competition. Consistent with this, while the NHS Plan of 2000 that had led to the market reforms of the earlier 2000s had been inspired by the then Secretary of State, career politician Alan Milburn,⁴ the NHS Review of 2007-08 was led hands-on by Ara Darzi, a prominent academic surgeon and clinical innovator from Imperial College medical school in London. Darzi was more than an expert adviser, as he was given a life peerage and appointed a health Minister in the House of Lord’s in order to lead the Review from within Government.

The Darzi White Paper: *High quality care for all*

The June 2008 White Paper, *High quality care for all* (HQCFA) was presented as the third official stage in Labour’s strategy for reforming the English NHS.⁵ Stage 1 had been the large investment in increasing the capacity of the NHS that followed the NHS Plan of 2000 and stage 2, the reintroduction of supply side competition after 2002. *High quality care for all* argued that the English NHS had been improving for the previous decade under the Labour Government, particularly in terms of access to treatment and responsiveness (e.g. shorter waiting times for planned treatment), but that it now needed to emphasise improving three other equally important aspects of quality: patient safety; patient experience (i.e. the humanity of care); and the effectiveness of care. In the rest of this essay, following Darzi’s definition of quality, ‘quality’ is taken to include safety of care.

There was also to be a focus on reducing the well known clinical variations in how care was delivered across the country by deploying the best evidence of what constituted the most effective care. Part of this would involve giving people more information about the quality of the different sources of care available and continuing to offer them more scope to exercise choice between providers. However, the aspiration in the White Paper was to go further and enable people increasingly to shape the nature of the services they received more directly. The White Paper advocated the ‘personalisation’ of services for people with long term conditions, as a way of extending market thinking to long term condition care, just as the earlier market-oriented reforms had focused on responding to those needing quicker diagnosis and elective treatment. The commitment to ‘personalisation’ in the 2008 White Paper took a number of forms. For example, everyone with a long term condition was to have a dedicated care plan focused on shared decision making between the patient and a named professional. It would provide the basis for both NHS and non-NHS providers to organise services around the needs of individuals. None of this was entirely new to the NHS. In practice, as in a number of other areas, the Darzi White Paper simply reinforced the existing aspiration towards more ‘personalised’ care for people with long term conditions. Although care planning had long been a desired feature of systems for managing care of people with long term conditions, the Darzi Report gave it renewed impetus.

‘Personalisation’ was also to be extended, on a pilot basis, in the form of personal health budgets for people with selected long term conditions, based on the recent experience of piloting personal budgets for people requiring social care (e.g. paid help in their own homes with aspects of daily living). The aspiration was to offer individuals, and their families, greater control over their own health care, including permitting them to pay directly for care in specific circumstances from funds provided by the NHS and placed in their own bank accounts. The White Paper also introduced ways of increasing patients’ ability to choose a general practice and included a new ‘right to choice’ to be set out in the first ‘NHS Constitution’, a new national statement of patients’ and staff rights and responsibilities in relation to the NHS in England.

The White Paper also included a range of measures designed to signal that the NHS should become much more of a health promoting Service, including a requirement that every primary care trust (PCT), the then principal commissioners of NHS services, should commission comprehensive well-being and preventive services in partnership with local government. This can now be seen as the start of the process that led to the decision to shift

all public health responsibilities from the NHS to local authorities that took place in April 2013 following the Health and Social Care Act of 2012.

It is interesting to note that the White Paper additionally proposed a ‘Coalition for Better Health’, which was to be a set of voluntary agreements between the Government, business and Third sectors focused on increasing the availability of healthier food, the level of physical activity in the population and the level of investment by employers in a healthier workforce. This proposal echoes almost exactly the plans that Andrew Lansley was developing with the private sector at about the same time while still in opposition that were to lead to the controversial Public Health Responsibility Deal introduced in March 2011 by the successor Coalition Government.⁶

Despite the presence of proposals relating to public health and health promotion outside the NHS, the core of the White Paper lay in the commitment to put ‘quality at the heart of the NHS’ for all patients. Consistent with this, there was a wide-ranging series of proposals focused on health care delivery, including:

- improvements in patient safety and hospital infection control regimes, including new enforcement powers for the external quality regulator;
- an ‘independent’ quality standards and clinical priority setting role for the National Institute for Health and Clinical Excellence (NICE), including identifying new quality indicators to use to pay general practices under the Quality and Outcomes Framework (QOF) element in the national general practice contract;
- a new National Quality Board in the Department of Health to advise Ministers, supported by Quality Boards in each region, and more local quality improvement activity, including in primary and ambulatory care, with a key role for NHS trust medical directors in quality improvement;
- systematic measurement and publication of information on the quality of care from the ‘frontline’ upwards, including the mandatory use of patient reported outcome measures (PROMs) for some common elective surgical procedures;
- all providers to publish annual Quality Accounts for their local populations in the same way that they published financial accounts;
- adapting hospital payment methods to recognise the quality of care rather than just activity, through the Commissioning for Quality and Innovation (CQUIN) scheme in

which the national, average cost PBR tariff would be adjusted upwards where ‘best practice’ was provided in areas of service needing improvement;

- A range of initiatives relating to encouraging and embedding a culture of innovation in the NHS such as funds and prizes, more horizon scanning and new industry-NHS-university partnerships;
- A commitment to strengthen the involvement of front-line clinicians in leading and managing organisations so that decision-making at every level in the NHS would involve clinical staff.

The breadth of the White Paper and the sheer number of proposals is overwhelming. It part this was because at least some of the initiatives included were already in train or had already been signalled in other ways. However, in other respects, the level of ambition was a reflection of Lord Darzi’s insight that deep cultural change at all levels of the NHS would be needed if the entire system were to become organised around the principle of high quality care. As a result, the range of proposals and the nature of the change envisaged meant that it was challenging to frame and commission a programme of evaluation that would coherently cover the main themes and initiatives in the White Paper, many of which are still the focus of NHS policy in 2013.

Evaluating the Darzi White Paper: the commissioned projects

The invitation to tender for evaluative research on the Darzi reforms was published in April 2009 and stated that, ‘At its core will be an assessment of the extent to which mindsets are changing and resulting in behavioural and cultural change.’ This required a very broad approach to research looking at how the NHS was altering as a whole. However, the evaluation programme was also planned to include projects looking at specific policy initiatives set out in HQCFA with a view to informing the modification of either the content of policy or processes of implementation (‘formative’ evaluation) in the short to medium term, as well as providing summative findings. In the event, it did not prove possible to commission research on all the chosen aspects of the White Paper, but the following projects were funded:

- an evaluation of the implementation and outcomes of care planning for people with long term conditions, led by Peter Bower;

- an evaluation of the personal health budget pilots, led by Julien Forder;
- an evaluation of the implementation and outcomes of the CQUIN framework, led by Ruth McDonald; and
- an evaluation of cultural and behavioural change in the NHS focused on ensuring high quality care for all, led by Michael West.

No proposals were funded looking at the different aspects of the new NHS Quality Framework, such as Quality Accounts and the collection of PROMs data from patients before and after common elective surgical procedures, though some attempts to evaluate these have been undertaken⁷⁻¹¹. An evaluation of a national programme of integrated care pilots led jointly by Ernst and Young and RAND Europe had been separately commissioned before April 2009.¹²

There are papers in this supplement from the projects led by Bower, Forder, McDonald and West.

There were a number of proposals in HQCFA that continued the process of change that had been going on in primary care since the early 2000s, such as making it easier for patients to change their general practice and widening the range of different types of primary and first contact care available in the NHS. Though the Health Reform Evaluation Programme did not include a project looking specifically at changes in primary care, this supplement contains a paper by Sheaff from a project funded by the National Institute for Health Research (NIHR) that surveys trends affecting primary care in the period from 2002 when Labour's market reforms began, through the Darzi reforms to the period of the Coalition Government after 2010.

Findings from studies evaluating the Darzi White Paper

Cultural and behavioural change in relation to quality of care

Perhaps the most ambitious project relating to the core goals of the White Paper of 2008, was that by Michael West and colleagues from which the paper by McKee et al. in this supplement is drawn. The project used interviews, surveys and ethnographic case studies to assess the state of quality practice in NHS acute hospitals in England before and after the Darzi White Paper.¹³ The study began after the failure of care at Mid-Staffordshire NHS

Foundation Trust, 2005-08, had become public knowledge and was undertaken in parallel to the various inquiries into the Mid-Staffs case by the Healthcare Commission published in 2009, by Robert Francis QC in 2010 and again in 2013.¹⁴⁻¹⁶ In contrast to these analyses, focused on a single poor performer, West and colleagues attempted to provide a wider picture of the English NHS in relation to many of the same issues. The picture they paint suggests that many of the failings identified at Mid-Staffs had the potential to have occurred elsewhere in the NHS.¹⁷

They identified a near universal desire to provide the best possible quality of care and found many outstanding examples of staff commitment, innovative approaches to improvement and excellent management approaches to quality and safety. However, the stated values and commitments did not always translate into good practice and positive action. As a result, the researchers found variable quality of care in that not all patients were treated with care, courtesy, civility and compassion. There were examples of vulnerable patients being neglected, staff preoccupied with administrative functions that eroded their time for patient care, and pressures of workload translating into lapses in the humanity of care very similar to those found at Mid-Staffs though these did not necessarily affect safety and effectiveness of care. Poor interfaces between teams were seen as posing particular risks to the quality of care.

Using the NHS national staff survey data, West and his colleagues found that the strongest predictor of hospital mortality was the percentage of staff working in well structured teams. Yet only 40% of NHS staff were found to be working in that way. Working in poorly structured teams is associated with higher hospital mortality, more errors that could harm staff or patients, and higher levels of injuries to staff. Good staff support and management were directly related to better patient experience, and safer and more effective care, yet were found to be highly variable across the NHS.

Shifting attention from the clinical to the organisational level, the study showed how local reorganisations and alterations in responsibility for quality at the national level had resulted in a perceived lack of leadership among the staff directly involved in patient care. Confusion and overlapping responsibilities for leading, monitoring and improving quality among multiple national level organisations was creating challenges for local NHS organisations

which faced multiple lines of accountability and misaligned performance measures. The researchers also observed disagreements between senior managerial staff and clinical staff on the nature of quality issues, and the appropriate solutions to these, which undermined a sense of shared purpose.

Few Trust Boards had clear, agreed upon and measurable, quality objectives. Clarity of aims relating to quality was not assisted by a multiplicity of poorly coordinated externally imposed targets, standards and incentives at that time in the NHS. A preoccupation with financial issues and performance targets was evident from surveys of Trust Boards and analysis of Board minutes, while a number of key scores in the National Staff Survey and Acute Inpatient Experience Survey, which had improved steadily in the years up to 2009, stagnated or deteriorated in the ensuing years.

Furthermore, there was wide variation in NHS organisations' ability to understand, monitor and act upon their quality performance. Some organisations struggled to collect and interpret valid data on the quality of the care they provided. Others were using data collection as a means of reassurance rather than to detect and act on problems. In contrast, more proactive organisations used data to challenge themselves and provided an environment in which quality concerns were in the consciousness of staff at every level. Poor organisational and information systems sometimes meant that staff struggled to deliver high quality care and disempowered staff from initiating improvements.

There were rich examples of innovations at the senior and clinical levels of NHS organisations, including clinical, administrative and managerial, service delivery and technological innovations. However, there was a decline in the amount of innovation over time at both Trust Board and clinical levels. Furthermore, relatively few innovations by Boards related to quality; most related to crude measures of efficiency and productivity. Taken in the round, the study suggests that it was taking time for the various initiatives set out in HQCFA in 2008 to have a detectable positive impact.

The paper by McKee and colleagues in this supplement¹⁸ focuses specifically on the type of leadership perceived by a large number of key participants at all levels in the system as likely to be the most effective for ensuring good quality of care and guarding against quality failures such as those seen at Mid-Staffordshire NHS Foundation Trust. Participants saw a

crucial role for leadership at all levels, and distinguished between two main types: traditional, hierarchical, ‘concentrated’ leadership associated with those in explicit managerial positions; and ‘distributed’ leadership involving people with particular skills and abilities operating throughout the system at all levels, irrespective of their formal job descriptions. While they saw an important role for distributed leadership, particularly where it led to strong coalitions between managers and clinicians, they were concerned lest an excessive reliance on distributed leadership lead to confusion and a vacuum of authority at national level. They argued that hierarchical leadership was needed to complement distributed leadership.

This is reinforced elsewhere in the findings from the project. West et al. concluded that clear national level direction-setting is crucial to maintain a coherent focus on quality improvement in acute hospitals, and that responsibility and accountability for quality should not be dispersed to avoid creating confusion for service providers.¹³ They argue that within NHS organisations, high quality leadership that sets clear, challenging objectives is needed. Trust Boards must make quality of care their over-riding focus, setting clear objectives and shaping organisational environments that facilitate improvement, and committing themselves to innovating for quality. This supports the current proposals for each Board to have a chief quality officer or director responsible for quality of care throughout the organisation.

The implementation and outcomes of care planning for people with long term conditions

While the project of West and colleagues focused on the systems, culture and behaviour supporting quality, primarily in hospitals providing NHS services, the evaluation led by Bower of the commitment to universal care planning for people with long term conditions set out in the HQCFA White Paper focused at the micro-level of care planning for individuals with long term conditions. Among other things, care plans are about linking the support and services provided by multiple providers and professionals across different organisations in contrast to West et al.’s concentration on hospitals alone.

A paper by Burt and colleagues¹⁹ from the same evaluation published in the previous HREP Supplement in January 2012 provided baseline information on the prevalence of care planning and care plans in England in 2009/10, the year after the Darzi report, among patients reporting a long term condition. It showed that 84% of people reported having had a care planning discussion of some sort in the previous 12 months, but only 12% of these patients reported being told that they had a ‘care plan’. Only 40% of patients with a long term

condition reported ‘definite’ benefits from care planning discussions. The most important factor associated with care planning, and care plans, was reported to be good interpersonal care. These findings indicated that the policy goal that everyone with a long term condition should have a written care plan related to a structured, comprehensive process of care planning was far from being realised. While care planning discussions were common, key elements such as goal setting and action planning were rarely present.²⁰ Indeed, patients were unfamiliar with the term ‘care planning’. It may be that this reflected professional uncertainty about the benefits of formal care plans. In the absence of specific incentives for the adoption of care planning and production of formal care plans, it seems scarcely surprising that take up was modest.

The remainder of the evaluation attempted to tease out what the benefits of care plans might be and who might be most likely to obtain them. It did this by comparing two groups of patients who were similar in demographic and clinical characteristics, but who had different levels of exposure to written care plans and care planning practices (‘high’ and ‘low’ care planning groups). There was no sign that the level and nature of care planning found in England at the time had any demonstrable impact on patients’ health outcomes after six or 12 months of follow up.²¹

The paper in this Supplement by Bower and colleagues from that evaluation²² focuses on the sub-group with multiple long term conditions who, it was hypothesised would report poorer experience of care and less ability to self-manage their conditions than people with a single long term condition. The implication was that this sub-group would pose particular problems for effective care planning and would require specific attention, but might benefit particularly from structured care planning. In fact, Bower and colleagues show that people with multiple long term conditions do not appear to be at greater risk of reporting poorer experience of care or a lower ability to self-manage their conditions, even including those with four or more conditions, possible depression, or a combination.²² They conclude that if there are particularly vulnerable sub-groups among people with long term conditions, they cannot be identified simply from counts of their conditions. Inevitably, the response rate for this part of Bower’s study was low, in line with other similar studies in this sort of population. Despite this, the findings suggest, perhaps encouragingly, that the English NHS was able, in this period, to provide care for people with multiple long term conditions which was at least no

worse than that for patients with single conditions, even if structured care planning did not seem to be a major contributor to this result.

The overall conclusion of the evaluation was that while care planning may have some potential to contribute to care for people with long term conditions, there is scant evidence that the sort of care planning that currently occurs in the English NHS makes a measurable difference to patient outcomes. This is consistent with other evaluations, in part, because care planning has many aspects and is thus difficult to evaluate as a single intervention. The researchers argue that rather than promoting care planning in general in the face of professional scepticism, in future it might be more sensible to try to identify sub-groups of people with long term conditions that are more likely to benefit appreciably.²¹ They suggest that care planning and a formal care plan might particularly benefit patients in need of end of life care and those managing a personal health budget. In the latter group, the care plan would have the advantage of linking directly to the resources needed to bring about service change which would be under the control of patients themselves. In this regard, the evaluation of the personal health budget pilot programme was able to shed some light on this and related issues since people with personal health budgets in the pilot were encouraged to develop a care plan showing how their budgets were to be used to meet their needs.

Evaluation of the personal health budget (PHB) pilot programme

If care planning has been promoted as a means to encourage dialogue and joint deliberation between patients and professionals on packages of care, personal health budgets are associated with a more consumerist model in which patients themselves choose and exercise control over how their health care is managed and delivered. Building on the experience of personal budgets for social care, the pilot programme launched in 2009 allowed patients with a range of long-term conditions to decide to take: a notional budget (held by the commissioner, but where the patient is aware of the available services and their costs); a budget managed by a third party; or a direct payment into their bank account. The evaluation, from which the papers by Jones et al.²³ and Davidson et al.²⁴ in this supplement are drawn, focused on 20 out of 64 pilot sites and on patients with chronic obstructive pulmonary disease (COPD), diabetes, long-term neurological conditions, mental health problems and stroke, plus patients eligible for NHS Continuing Healthcare, and their respective carers. The evaluation consisted of a controlled trial that compared outcomes in equal numbers of people (n=1000) offered personal health budgets with those receiving

conventional care, in terms of health status, health-related quality of life, care-related quality of life (covering dimensions such as personal care, social participation, usual activities, anxiety), life satisfaction, psychological well-being, mortality, service use and cost. Some of the participants were randomised between personal budget and conventional care, but this was not possible in all the pilot sites. Sub-samples of budget holders and carers were interviewed in depth twice during the pilot about their experiences of being offered, receiving and using a budget.²⁵ Staff in the 20 evaluation PCTs were also interviewed at month 3 and month 24 of the evaluation about their experience of implementation of personal health budgets.

People with personal health budgets had significantly better care-related quality of life and psychological well-being, but no better health status, or health-related quality of life after 12 months. There was no significant difference in mortality or costs between the two groups. Personal health budgets were estimated to be cost-effective at current NHS thresholds in terms of care-related quality of life net benefits.²³

Although improved care planning was not the main aim of the personal health budget pilots, the study suggests that care planning in the context of personal health budgets may well be worthwhile for some patients with long term conditions. This broadly supports the Government's decision to extend personal health budgets beyond the pilot programme as an option available to all people in receipt of NHS Continuing Healthcare from April 2014, despite the fact that the decision was first announced about a year before the end of data collection in the evaluation.

The in-depth interviews with budget-holders and carers corroborated some of the positive outcome results from the quantitative analysis, but are probably even more useful for the insights they provide into how personal health budgets can best be implemented in future and for whom.²⁴ For example, there were some practical difficulties with defining what should go into the budget and setting its level; getting information and advice on possible uses of the budget; obtaining approval for expenditure; and the supply of some PCT-procured items. An important finding was that people recently diagnosed and/or with a sudden onset condition appeared to have greater difficulty in identifying outcomes that could benefit from the use of a personal health budget. It seems that the more experience people have of their long term conditions and associated needs, the more they can use a personal health budget to improve their well-being. On the other hand, budget-holders and carers from all patient groups

reported that improvements in health and well-being extended well beyond the specific condition for which the budget had originally been offered.

Evaluation of the Commissioning for Quality and Innovation (CQUIN) scheme

While the Darzi White Paper made much of the importance of improving intrinsic motivation among health care professions with its emphasis on embedding a new culture of high quality care throughout the NHS, there were also proposals in HQCFA that resonated more directly with Labour's previous market-related reforms and focused more explicitly on altering the extrinsic motivation of health care providers by offering new financial incentives. One such was the CQUIN payment scheme designed to shape commissioner-provider negotiations and to respond to the criticism of the PbR prospective payment system for hospitals that it rewarded activity regardless of its contribution to outcomes.

Under CQUIN, a proportion of a provider's income is made conditional on achieving a set of measurable quality-related goals set each year. The intention is to improve safety, effectiveness and patient experience through innovative services. Unlike much policy-making under Labour, the Government explicitly decided to make the scheme almost entirely locally determined so that the goals and indicators would reflect local priorities and thus be more likely to be enthusiastically pursued than if a set of national indicators had been imposed. The overall maximum size of the financial incentive in CQUIN was increased from 0.5% of total contract value in 2009/10 to 1.5% in 2010/11 and 2.5% in 2012/13, and applies to acute hospital and, ambulance, community and learning disability service providers. For hospitals, the CQUIN approach was inspired by the Advancing Quality (AQ) initiative implemented in North West England's 24 NHS hospital trusts starting from 2008. There were encouraging early, informal reports from the North West and these proved influential in encouraging the setting up of CQUIN. AQ was based on the US Hospital Quality Incentive Demonstration (HQID), but with larger financial incentives and more emphasis on hospital quality improvement activities. Subsequent formal evaluation published in 2012, well after CQUIN had been instituted, showed that AQ was associated with a statistically and clinically significant reduction in mortality in the conditions included in the scheme (a 1.3% percentage reduction in the combined mortality for the three conditions studied).²⁶ However, the study did not compare the North West with trends in other regions of the country.

The CQUIN evaluation in the Health Reform Evaluation Programme was led by Ruth McDonald with colleagues (a number of whom were also involved in the previous AQ evaluation) and used a mix of qualitative (documentary analysis of schemes, observation of meetings in 12 sites and interviews) and quantitative (costs and impacts of schemes, e.g. on outcomes such as mortality) methods to look at the way in which local CQUIN schemes were developed (e.g. the choice of goals and indicators). Unlike the AQ evaluation, this allowed a more rigorous comparison of places in which providers had selected a CQUIN goal with those that had not.²⁷ In contrast to AQ, the results were described by the evaluators as ‘disappointing’, principally because of the way that the idea behind CQUIN was implemented through a series of divergent local schemes.

The paper by Kristensen and colleagues in this supplement²⁸ focuses on the extent to which the local design of CQUIN schemes enabled or hindered the realisation of the expectations set out in the HQCFA White Paper⁵ and related guidance (a focus on outcomes and processes with a clear link to quality, use of established indicators where possible and coverage of the three domains of quality – safety, effectiveness and patient experience, plus innovation). They show that the expectations of national policy makers were mostly not met in that the 337 locally negotiated schemes included in their analysis covered more than 100 clinical topics and relied on more than 3000 different indicators (in total, in 2010/11, 5000 indicators appeared in the local schemes). There was no sign of any trend towards greater standardisation of indicators across schemes. Furthermore, the indicators were mostly focused on processes of care and rarely covered all four of the quality domains. It appears that there was a mismatch between the expectations set nationally and the decision to encourage extensive local variation in schemes with little attempt to assure the quality of what came out of local development processes. For example, many locally agreed indicators were not only based on structure and processes of care rather than outcomes, but tended to lack precision and be based on weak evidence of likely effectiveness. It was also a struggle to engage local clinicians in CQUIN in contrast with schemes such as AQ and Best Practice Tariffs, the latter also introduced as part of the Darzi White Paper.

The contrast with the design and implementation of AQ in the North West is instructive, further reinforcing the knowledge that the details of the design of payment for performance schemes are extremely important for their (cost) effectiveness.²⁹ Although there was very extensive emphasis on strong local involvement in the implementation of AQ, the latter was

based on clearly identified clinical issues and a set of carefully selected, consistent indicators. The financial incentive programme covered five clinical areas only which were relevant to all participating organisations: acute myocardial infarction (AMI); heart failure; coronary artery bypass grafting (CABG); pneumonia; and hip and knee replacements. Furthermore, unlike CQUIN, the indicators did not change annually, allowing implementers to take a longer term approach to quality improvement. As a result, it was possible for organisations and staff to compare directly how their peers in the region were engaging on the same issues and with what results. The CQUIN assumption, that clinicians would only engage in incentivised quality improvement activities if they were entirely locally derived, does not seem to have been borne out in reality in AQ because of other features of its design and implementation.

Given the way that the CQUIN schemes were developed locally, as described by Kristensen et al.²⁸, it is hardly surprising that the evaluators struggled to identify any impact of the content of local CQUIN schemes on performance improvement with the exception of hip fracture. None of the other estimated effects was statistically significant. Non-significant quality effects were both negative and positive.²⁷

Building directly on their findings, McDonald and her colleagues make a number of telling recommendations in the final report of the evaluation for improving payment for performance schemes in future. In particular, they emphasise the need to recognise that while quality improvement priorities and/or goals can feasibly be identified locally, the technical design work (e.g. defining valid indicators, setting thresholds, determining the financial rewards and standardising reporting) should be undertaken separately by experts. They also recommend that schemes focus on a small number of sustained indicators linked to high impact changes rather than a large number of indicators covering a wide range of conditions that change frequently.

Trends in the provision of primary medical care, 2002-12

While the encouragement by successive governments of greater diversity among providers of NHS hospital services in England during the first decade of the 21st century has been widely discussed,³⁰ much less attention has been given to related trends in ambulatory care, particularly the changing nature of the provision of primary medical care in the NHS. In part, this may be because since the inception of the NHS in 1948, the vast majority of general practices have remained in private ownership on contract to the NHS whereas the reverse is true of the acute hospitals.

The final paper in the Supplement, by Sheaff, drawn from a NIHR-funded project outside the Health Reform Evaluation Programme, examines how the range of organisational forms and ownership types among primary care providers changed during the period of Labour's market-led reforms and identifies some of the implications of these trends.³¹ He shows that existing providers (mostly the traditional, privately owned partnership general practices) changed at the margins to become larger, with more salaried GPs and providing a wider range of services, in response to a greater choice of contractual forms with the NHS. The entry of new types of providers in NHS primary medical care was far more limited, though by 2011 there was at least one limited company providing care on contract in most primary care trust areas. The vast majority of primary medical care contracts (97%) remained with standard GP partnerships. New providers appeared to make greater use of nurses for case management than traditional general practices.

The effects of existing providers providing some different services and/or changing their identity, and entirely new types of providers entering the NHS on the cost and quality of care were hard to discern and there was no systematic bias in favour of 'new' versus 'old' providers. This was, in part, due to a lack of research because primary medical care was not specifically included within the Health Reform Evaluation Programme as the principal focus of market-led reforms in the mid-2000s when the Programme was initiated had been on the hospital sector and elective care in particular. Another difficulty in comparing performance between incomers and incumbents lay in the fact that new providers tended to enter either to provide new services for which there was no previous equivalent (e.g. walk-in centres) or to replace poorly performing providers.

Conclusions from the evaluations of elements of the 2008 Darzi reforms

The Health Reform Evaluation Programme endeavoured to fund evaluative research on most of the novel elements in the *High quality care for all* White Paper.⁵ In the event, it was only possible to commission high quality work in some of the priority areas identified in the research brief. Unsurprisingly perhaps, taken together, the evaluations present a mixed picture of the impact and success of the reforms. All the studies identify some limitations of the policies. The introduction of personal health budgets appears to have been the least problematic and, depending on assumptions, likely to be cost-effective for the sorts of

patients involved in the pilot. For the rest of the changes, impacts ranged from little or none (CQUIN and care planning for people with chronic conditions) to patchy and highly variable (instilling a culture of quality in acute hospitals) in the three years following the publication of the White Paper. On the other hand, each of the studies identifies important insights relevant to modifying and improving the policies. For instance, the CQUIN evaluation highlighted the importance for successful implementation of distinguishing between local control over the priority goals of payment for performance schemes and the technical knowledge required to design schemes in detail. Both the evaluation of care planning for people with long term conditions and the personal budget pilot evaluation identified sub-groups of the patient population more and less likely to benefit from the intervention, and thus ways of improving targeting and increasing the (cost) effectiveness of the policy initiatives. These findings have continuing relevance since both the White Paper's policies and the issues they were focused on remedying remain central to the current Government's reform agenda.

The papers in this Supplement show that good policy evaluations should not only attempt to provide summative conclusions about whether, and to what extent, a policy is 'working' as intended, but also identify the unanticipated consequences and the ways in which the policy can be modified to produce better results in future.

Acknowledgement

The Health Reform Evaluation Programme coordination and individual projects in the Programme were funded by the Department of Health through its Policy Research Programme. However, any views expressed are those of the researchers, not the Department of Health.

References

1. Supplement. *J Health Serv Res Policy* 2012; 17 (suppl 1): 1-71
2. Mays N, Dixon A, Jones L, eds. *Understanding New Labour's market reforms of the English NHS*. London: King's Fund, 2011.
3. Mays N and Tan S. Evaluating Labour's market reforms, 2002-10. *J Health Serv Res Policy* 2012; 17(Suppl 1); 1-6 doi:10.1258/jhsrp.2011.011162
4. Department of Health. *The NHS Plan – a plan for investment, a plan for reform*. CM4818-I. London: The Stationery Office, 2000.
5. Department of Health. *High quality care for all: NHS Next Stage Review Final Report*. Cm 7432. London: The Stationery Office, 2008.
6. Department of Health. *The Public Health Responsibility Deal*. London: Department of Health, 2011.
7. MHP Health Mandate. *Accounting for quality: an analysis of the impact of quality accounts in the NHS*. London: MHP Health Mandate, 2010. Available at: www.mhpc.com/sites/default/files/docs/Accounting%20for%20quality.pdf (accessed 11 January 2012)
8. West D. Quality Accounts: trusts dodge scrutiny by omitting vital measures. *Health Serv J* 2010; 120(Jul 29): 12–3
9. Foot C, Raleigh V, Ross S, Lyscom T. *How do quality accounts measure up? Findings from the first year*. London: King's Fund, 2011. Available at: www.kingsfund.org.uk/document.rm?id=8886 (accessed 11 January 2012)
10. Black N, Nossiter J. Reporting national clinical audits in quality accounts. *British Journal of Healthcare Management* 2011; 17: 75–9
11. Ritchie B, Mays N. Quality accounts and quality monitoring by NHS boards. *British Journal of Healthcare Management* 2012; 18: 87-94
12. RAND Europe and Ernst and Young LLP. *National evaluation of the Department of Health's integrated care pilots*. Cambridge: RAND Europe, 2012.
13. West M, Baker R, Dawson J, Dixon Woods M, Lilford R, Martin G, McKee L, Murtagh M, Wilkie P. (2013) *Quality and safety in the NHS: evaluating progress, problems and promise*. Final report to the Department of Health. Lancaster: University of Lancaster, 2013. <http://www.lums.lancs.ac.uk/research/centres/hr/nhs-quality/>

14. Healthcare Commission. *Investigation into Mid Staffordshire Foundation Trust*. London: Healthcare Commission, 2009.
15. The Mid Staffordshire NHS Foundation Trust Inquiry (Chair: Robert Francis QC). *Independent Inquiry into care provided by Mid Staffordshire NHS Foundation Trust January 2005 – March 2009. Volumes I and II*. HC375-I/II. London: The Stationery Office, 2010. <http://www.midstaffspublicinquiry.com/report>
16. The Mid Staffordshire NHS Foundation Trust Public Inquiry (Chair: Robert Francis QC) (2013) *Report of the Mid Staffordshire NHS Foundation Trust Public Inquiry*. HC 947. London: The Stationery Office, 2013. <http://www.midstaffspublicinquiry.com/report>
17. Dixon-Woods M, Baker R, Charles K, et al. Culture and behaviour in the English National Health Service: overview of lessons from a large multi-method study. *BMJ Quality & Safety* 2013, in press
18. McKee L, Charles K, Dixon-Woods M, et al. 'New' and distributed leadership in quality and safety in health care, or 'old' and hierarchical? An interview study with strategic stakeholders. *J Health Serv Res Policy* 2013; 18(suppl 2): [pages]
19. Burt J, Roland M, Paddison C, et al. Prevalence and benefits of care plans and care planning for people with long-term conditions in England. *J Health Serv Res Policy* 2012; 17(Suppl 1): 64–71. doi:10.1258/jhsrp.2011.010172
20. Newbould J, Burt J, Bower P, et al. Experiences of care planning in England: interviews with patients with long term conditions. *BMC Family Practice* 2012; 13:71. doi: 10.1186/1471-2296-13-71
21. Bower P and the CAPITOL project team. *Care planning in the treatment of long term conditions: final report of the CAPITOL project*. Manchester: Centre for Primary Care, Institute of Population Health, University of Manchester, 2012 <http://www.population-health.manchester.ac.uk/primarycare/research/longterm/multimorbidity/research/index.aspx>
22. Bower P. Multi-morbidity and care for long term conditions in the English NHS: baseline data from a cohort study. *J Health Serv Res Policy* 2013; 18(suppl 2): [pages]
23. Jones K, Forder J, Caiels J, et al. Personalisation in the health care system: do personal health budgets impact on outcomes and cost? *J Health Serv Res Policy* 2013; 18(suppl 2): [pages]

24. Davidson J, Baxter K, Glendinning C, et al. Choosing health: qualitative evidence from the experience of personal health budget holders. *J Health Serv Res Policy* 2013; 18(suppl 2): [pages]
25. Forder J, Jones K, Glendinning C, et al. *Evaluation of the personal health budget pilot programme*. Discussion Paper 2840_2. Canterbury: Personal Social Services Research Unit, University of Kent, 2012. <https://www.phbe.org.uk>
26. Sutton M, Nikolova S, Boaden R, et al. Reduced mortality with hospital pay for performance in England. *New England Journal of Medicine* 2012; 367: 1821-8
27. McDonald R, Kristensen SR, Sutton M, et al. Evaluation of the Commissioning for Quality and Innovation Framework. Final report. Nottingham: University of Nottingham, 2013.
28. Kristensen S, McDonald R, Sutton M. Should pay-for-performance schemes be locally designed? Evidence from the Commissioning for Quality and Innovation (CQUIN) framework. *Journal of Health Services Research & Policy* 2013; 18(suppl 2): [pages]
29. Van Herck, P, De Smedt D, Lieven A, et al. Systematic review: effects, design choices, and context of pay-for-performance in health care. *BMC Health Services Research* 2010; 10: 247. doi:10.1186/1472-6963-10-247
30. Allen P. Provider diversity in the English NHS: a study of recent developments in four local health economies. *J Health Serv Res Policy* 2012; 17(Suppl 1): 23-30. doi: 10.1258/jhsrp.2011.011015
31. Sheaff R. Plural provision of primary care in England, 2002-10. *J Health Serv Res Policy* 2013; 18(suppl 2): [pages]