Debate & Analysis

Lessons from evaluation of the NHS white paper Our Health, Our Care, Our Say

INTRODUCTION

The NHS white paper Our Health, Our Care, Our Say was published by the previous UK Labour government in January 2006, describing a new strategic direction for health and social care in the community. This had four main goals: (a) better prevention and earlier intervention for improved health, independence, and wellbeing; (b) more choice and a stronger voice for individuals and communities; (c) tackling inequalities and improving access to services; and (d) more support for people with long-term needs.

The context for these policy goals was the need, experienced by all developed countries, to remodel their healthcare systems to reflect the changing needs of their populations. In particular, the focus of health care is increasingly to support people to manage long-term health conditions at home and to reduce the number of admissions to hospital. Helping people to get more convenient and faster access to health care, providing them with more information to enable them to care for themselves, and integrating health and social care systems, are all strategies to meet this need.

To achieve the policy goals, the white paper promoted a range of initiatives. The government produced a structured framework for ensuing implementation of these initiatives and tracking progress made towards the policy goals. A series of evaluations of the initiatives were commissioned, including formal programme evaluation, demonstration sites, pilot projects, and formative evaluation. This emphasis on evaluation reflected a commitment, which was increasingly evident from the late 1990s onward, to base policy on evidence about what works.

What has been learned from this programme of evaluation about the extent to which these initiatives are achieving the policy goals of the white paper? We were commissioned by the Department of Health to review the evaluations of 10 initiatives that were specifically promoted by the white paper. These included high-profile developments such as the Improving Access to Psychological Therapies Programme (IAPT), self-referral to physiotherapy, NHS LifeCheck, and a range of other initiatives in health and social care: the full list of initiatives is shown in Box 1.

An extensive mapping exercise was undertaken in an attempt to identify evaluations that were being conducted, or had been conducted, in England of any of the specified initiatives. We were interested not only in major nationally commissioned evaluations but also in studies conducted locally, typically commissioned by primary care trusts.

Principal investigators of evaluations identified by the mapping exercise were asked to complete an online questionnaire describing the aims, context, methods, progress, or findings and dissemination of their evaluation, as well as details about their funding and potential conflicts of interest. They were asked for copies of protocols, interim and final reports, and published papers, where available. Twenty-one evaluations were selected for in-depth analysis using a case study approach, including major nationally commissioned evaluations of 10 initiatives. For each case study, the principal investigator leading the evaluation was interviewed about how and why the evaluation was undertaken and strengths and limitations of the methodological approach chosen.

Using all the available information from the above sources, guidelines developed by the Clinical Appraisal Skills Programme (CASP) were used to assess the strengths and weaknesses of each evaluation, and structured summaries were constructed of the evidence available from all sources about the success of each of the white paper initiatives or ways in which they could be improved. Lessons were also drawn from across the different evaluations about the extent to which the initiatives were helping to meet the white paper goals.

STRENGTH OF EVIDENCE AVAILABLE ABOUT SPECIFIC INITIATIVES

Box 1 briefly describes each initiative and the methodological approach that was taken in each of the nationally commissioned evaluations. It is notable that these evaluations varied widely in terms of their aims, approach, and scale. The appraisal of the evidence available from the national evaluations, based on the information provided in the online questionnaire and the analysis of protocols, papers, and reports is summarised in Box 2. These conclusions may not necessarily correspond with the findings reported by those conducting the evaluations. Table 1 summarises the qualitative assessment about whether the evaluations provide evidence about the benefits and costs of each of the initiatives.

PROGRESS TOWARDS THE GOALS OF OUR HEALTH, OUR CARE, OUR SAY

Better prevention and earlier intervention

Several initiatives were clearly designed to support this strategic aim and there were a number of examples to suggest that the initiatives might be successful. For example, evaluation of the NHS LifeCheck programme found that most young people were positive about the potential of LifeChecks to impact on knowledge, attitudes, and behaviour, and the Partnerships for Older People Projects (POPPs) demonstrated improvements in how older people perceived their quality of life, as well as suggesting a reduction in emergency hospital admissions. However, none of the evaluations so far reported has demonstrated that this earlier intervention is associated with improved health outcomes for patients. In addition, providing earlier intervention can increase health service costs, and few of the evaluations provided robust evidence about costs.

Where evaluations explored cost-effectiveness, these analyses were sometimes based on models incorporating limited data about actual costs and a wide range of assumptions. In addition, there were concerns about the take-up of some early-intervention services. For example, take-up of LifeChecks, self-referral to physiotherapy, and assistive technologies in Whole System Demonstrator sites were all lower than anticipated.

More choice and a stronger voice for individuals and communities

Several initiatives were designed to provide patients with greater choice in how services are delivered, including information prescriptions, self-referral to physiotherapy, and the IAPT programme. A clear example is the individual budgets initiative, which was designed to give individuals more control by allowing them to purchase the care they felt...
they needed. The evaluation provides some support for this initiative.26,27 People receiving individual budgets were more likely to feel in control of their lives, how they accessed support, and how their care was delivered. The evaluation of individual budgets suggested that the benefits were greater for some groups of people (such as those with mental health problems) than for others (such as older people).

The evaluation of information prescriptions also demonstrated that this initiative gave people greater control [Box 2].28 But this evaluation also suggested that the benefits varied for different types of people (for example, those in poor health or living in disadvantaged areas were less likely to benefit). Information prescriptions led to more discussion and longer consultations with patients, but it was unclear whether the extra costs this incurs are offset by reductions in subsequent consultations or improvements in patient outcomes.

**Box 1. Evaluation of initiatives arising from Our Health, Our Care, Our Say**

<table>
<thead>
<tr>
<th>Initiative</th>
<th>Description of initiative</th>
<th>Description of evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Care closer to home demonstration sites</td>
<td>New services in community, aiming to reduce demand on hospitals. Included GPs with special interests, community-based consultant clinics, telephone support. Five demonstration sites in each of six specialities.</td>
<td>Aim: to describe organisation and implementation, impact on access, quality, and costs. Design: interviews with providers and commissioners, postal survey of patients using new services and a limited number of control patients. Compared costs of new services in six sites against national tariff.</td>
</tr>
<tr>
<td>Improving Access to Psychological Therapies</td>
<td>Two demonstration sites based on different approaches to significantly expanding availability of psychological therapies: high volume low intensity or case management.</td>
<td>Aim: to assess organisational implications of the new approaches and whether they are cost-effective and acceptable to patients. Design: cohort study of costs and outcomes for patients in demonstration sites compared with control sites and national datasets; mixed methods study of system impacts; questionnaire and qualitative study of patient experience.</td>
</tr>
<tr>
<td>Individual budgets</td>
<td>Service users allocated budgets according to social care needs that can be used to purchase care, equipment, housing, and employment support.</td>
<td>Aim: to assess implementation, cost-effectiveness, and user experience. Design: randomised controlled trial plus interviews with service users, carers, and staff.</td>
</tr>
<tr>
<td>Information prescriptions</td>
<td>Individually personalised information in relation to needs. Twenty pilot sites established, providing different forms of information prescription for different types of patients.</td>
<td>Aim: to assess effectiveness and impact on patients and services and to gather learning about implementation. Design: qualitative research with staff; survey of patients, carers, and staff; collection of data about activity and estimates of resources used; action learning events with pilot sites.</td>
</tr>
<tr>
<td>New types of workers</td>
<td>Twenty-eight pilot sites developing a range of new care roles intended to address policy objectives such as patient-centred care, improved access, supporting care at home.</td>
<td>Aim: to support pilot sites to develop new roles and management systems. Design: documentary analysis of pilot proposals and reports; interviews and focus groups with managers, workers, and people using services.</td>
</tr>
<tr>
<td>NHS LifeCheck</td>
<td>Self-assessment tools to help people identify their health behaviour and make changes. Developed around three lifestages: parents of babies; teenagers; mid-life. Pilot sites established for each lifestage.</td>
<td>Aim: to gain feedback about improving the tools and to explore their acceptability to potential user groups. Design: interviews and focus groups with potential users and staff, a survey of young people, and analysis of use of the website for Teen LifeCheck.</td>
</tr>
<tr>
<td>Partnerships for Older People Projects</td>
<td>Twenty-nine pilot sites were established to develop and evaluate innovative partnerships between health, social care, and third sector agencies to promote health and independence of older people.</td>
<td>Aim: to develop explanatory framework to understand most effective approach. Design: documentary analysis; collection of activity data from pilot site; interviews and focus groups with service users; a survey of people before and after accessing a partnership; comparison with routine data about emergency data from non-matched control sites.</td>
</tr>
<tr>
<td>Self-referral to physiotherapy</td>
<td>Six pilot sites allowed patients to refer themselves for physiotherapy rather than needing referral from a doctor.</td>
<td>Aim: to evaluate impact in terms of waiting times, changes in activity and identity uptake by different groups of people. Design: historical and prospective data about activity and waiting times, a minimum data set about patients following introduction of self-referral, and feedback forms from GPs and physiotherapists.</td>
</tr>
<tr>
<td>Social enterprise pathfinders</td>
<td>Twenty-six pathfinder projects to develop organisations providing a range of community health and social services on a social enterprise model.</td>
<td>Aim: to assess success in meeting social enterprise pathfinder goals. Design: mixed methods including interviews, focus groups, workshops, telephone survey of pathfinders, and collection of data on costs. Mainly formative approach.</td>
</tr>
<tr>
<td>Whole System Demonstrator sites</td>
<td>Three primary care trusts designated as demonstrator sites. Implemented integrated health and social care through system redesign. Focus on people with long-term conditions or complex needs supported through assistive technologies.</td>
<td>Aim: To assess impact on service use, patient outcomes, cost-effectiveness, patient, carer and providers experiences, factors associated with successful implementation. Design: large cluster randomised controlled trial with nested qualitative research on users and providers experiences.</td>
</tr>
</tbody>
</table>

Tackling inequalities and improving access to services

Several initiatives appeared to have improved access to care. For example, the self-referral to physiotherapy scheme clearly made it easier for people to access care, and was popular with patients, but did not appear to lead to any overall increase in demand for physiotherapy.29 The ‘care closer to home’ initiative was also popular with patients, who appreciated the availability of more local services.16 Evaluation of the IAPT programme suggested that it has achieved to some extent the aim of increasing access to care for mental health problems. The NHS Teen LifeCheck programme achieved a high level of awareness in the most vulnerable groups, and the most positive responses with regard to the Early Years LifeCheck came from younger and less-experienced
Box 2. Summary of evidence from evaluations of the white paper initiatives

**Care closer to home demonstration sites**
Redesign of services demonstrates potential to deliver services closer to home. Improved convenience and shorter waiting times for patients. Developments were often driven by local circumstances or the enthusiasm of local champions. Lack of evidence about whether quality of care and patient outcomes are equivalent in ‘closer to home’ and hospital sites. May be cheaper for PCT commissioners, but impact on total NHS costs is complex and not yet fully explored.

**Improving Access to Psychological Therapies**
Evaluation experienced substantial challenges, making it hard to draw clear conclusions. Improved access to care; probably similarly effective to other therapy services but probably not more cost-effective than usual care.

**Individual budgets**
People in receipt of budgets, and their carers, reported greater independence and control over how care is provided. Individual budgets were slightly more cost-effective for some (but not all) groups of people. Implementation of budgets has important implications for staffing, training, and funding streams.

**Information prescriptions**
Evaluation provided useful information about improving implementation: important since establishing systems proved complex. The diversity of models of ‘information prescription’ made evaluation difficult. Evidence about benefits for users is limited. Most users felt more confident about their condition, but this was less true for those with the greatest needs. Carers and health professionals were also positive. Information prescriptions require considerable resources. Health and quality-of-life benefits have not been assessed.

**New types of workers**
Pilot projects, mainly designed to develop new services. Useful for identifying issues that need to be addressed, such as training and workforce needs. No tightly defined objectives so evidence of impact hard to assess. Evaluation was brief and limited.

**NHS LifeCheck**
Formative evaluation has been conducted to improve uptake and perceived usefulness of LifeCheck (separately for Baby LifeCheck, Teen LifeCheck, and Mid-life LifeCheck) but published information about methods and findings is limited. No evaluation yet about the costs or benefits of LifeCheck.

**Partnerships for Older People Projects (POPPs)**
Have improved awareness of, access to and engagement with services among older people. This appears to have increased independence and quality of life, although this conclusion is based on limited data. POPPs appear to be cost-effective, largely due to a reduction in emergency hospital admissions, in comparison with areas without POPPs.

**Self-referral to physiotherapy**
Popular with patients, this improves accessibility of services. Concerns about an unsustainable increase in demand appear unfounded, if provision of services is adequate at baseline. No robust evidence available about many of the aims: the impact on patient health outcomes, return to work, waiting times, GP workload, or NHS costs.

**Social Enterprise Pathfinders**
Formative evaluation described organisation and activities of pathfinder sites, and identified issues relating to commissioning and implementing/developing pathfinders. Intended benefits were broad and success criteria not defined. No substantive evidence about the benefits for service users has published.

**Whole System long-term condition demonstration sites**
Ambitious evaluation is underway, using a randomised controlled trial combined with qualitative and economic methods. No findings have yet been published.

Parents, as well as those from more deprived backgrounds and those lacking support networks.\(^\text{18}\)

More support for people with long-term needs
Initiatives to address this goal include information prescriptions, new types of workers, and the Whole System Demonstrator programme, but information from this evaluation is not yet available. Those people who took up the offer of an information prescription found it helpful.\(^\text{12}\) Evaluation of the new types of workers’ initiative provided examples of how workers operating within new roles were able to support people with long-term conditions to enable them to live at home.\(^\text{15–17}\) The limited evidence available about the IAPT programme showed that people who use these services generally have positive experiences of them (although only 6% of the eligible population were referred).\(^\text{18}\)

There was some evidence that initiatives such as the POPP\(^\text{9}\) and care closer to home demonstration sites\(^\text{14}\) may help to prevent hospital admissions or secondary care referrals. These findings are potentially important, but, given their resource implications for the NHS, they need replicating using stronger research designs with concurrent randomised or matched control groups.

**TRADE-OFFS BETWEEN COMPETING PRIORITIES**
When considering the broad sweep of evidence arising from the various evaluations of white paper initiatives, it is clear that there is some evidence that these initiatives are helping to achieve each of the policy goals set out in the white paper.\(^\text{1}\) The evidence is strongest in relation to improving access to care and less strong in relation to better prevention and greater support for people with long-term needs.

There may be trade-offs to be made between the policy objectives set out by the white paper. For example, evaluation of several initiatives demonstrated that use of services and their apparent benefits varied for different groups of the population.\(^\text{10–13}\) Evaluation of information prescriptions suggested that those with the greatest needs were least likely to benefit.\(^\text{12}\) Improved access to services could therefore potentially increase rather than decrease inequalities, by improving access most for those with least needs.

Improving access to services is also not
The extent to which governments are committed to evidence-based policy can be assessed by how far they insist that new initiatives are implemented in a way that allows for meaningful evaluation, and how the results of evaluation are used in the policy process, rather by the amount of money spent on evaluation-related activity.”

### Table 1. Evaluation of costs and benefits of initiatives

<table>
<thead>
<tr>
<th>Initiative</th>
<th>Robust evidence from evaluation on:</th>
<th>Health or care outcomes</th>
<th>NHS or social care costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Care closer to home demonstration sites</td>
<td>×</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improving Access to Psychological Therapies</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Individual budgets</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Information prescription</td>
<td>Limited</td>
<td>Limited</td>
<td></td>
</tr>
<tr>
<td>New types of workers pilots</td>
<td>Very limited</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NHS LifeCheck</td>
<td>×</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partnerships for older people</td>
<td>Very limited</td>
<td>Limited</td>
<td></td>
</tr>
<tr>
<td>Self-referral to physiotherapy</td>
<td>×</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social enterprise pathfinders</td>
<td>×</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whole System long-term condition demonstration sites</td>
<td>✓ (Robust evidence)</td>
<td>✓ (Robust evidence)</td>
<td></td>
</tr>
</tbody>
</table>

*Robust evidence = comparison between those receiving the initiative and a concurrent, comparable, control group. Data about health or social care outcomes based on reliable empirical data collected directly from patients/clients; data about costs based on economic analysis of directly collected data about resource utilisation and costs. Control areas may have important differences from intervention areas. ✓ = robust evidence available. × = no robust evidence. ( ) = Planned, but results not yet available.

LIMITATIONS OF THE EVIDENCE

The total investment in evaluation of these white paper initiatives has been considerable. At a national level, several evaluations had budgets of over £100 000, with some considerably more expensive than that. Alongside this investment, many local primary care trusts were conducting local evaluations of the same initiatives. This commitment to evaluation is commendable and appropriate, given that the costs to the NHS and social care services of establishing these new services will have been many times greater. For example, the IAPT programme alone represents an investment of over £173 million. However, there were clear limitations to the evidence arising from this investment in evaluation, and it is questionable whether the money spent on it was used in the best way.

Few evaluations sought to demonstrate the impact of initiatives in terms of improved health outcomes. Although the nature of these initiatives means that it is difficult to demonstrate impact on disease outcomes within the short term, it would be possible to assess the impact on health status, on quality of life, and on intermediate outcomes (such as blood pressure control).

Many of the evaluations reviewed were largely formative, describing the processes being undertaken within each initiative and ways in which the initiative might be improved. Many described patients’ perceptions of potential benefit rather than actual benefit. Where evaluation of outcomes was attempted, this often did not make use of the strongest research designs for this purpose. Some evaluations were essentially descriptive, with no meaningful comparison with alternatives. Others used observational methods to compare areas with and without new services, which is a weak design if there are other differences between the areas. This was particularly the case in these evaluations, as many of the pilot sites and demonstration sites were chosen for specific reasons or were self-selected, and control areas were likely to have important differences.

The interviews conducted with principal investigators identified many good operational reasons why researchers used these observational methods rather than stronger experimental methods (for example, prospectively allocating areas or individuals to receive the new services or conventional services, ideally by random...
allocation). These weaker designs were generally chosen, not because researchers were unaware of their limitations, but because of the constraints they faced when designing the research. For example, the implementation of initiatives had often already started before evaluation had begun, and there were often imperatives to produce findings at an early stage, before the impact of initiatives could be meaningfully assessed. In other examples, meaningful evaluation was undermined because participating sites were chosen for a variety of reasons that made them unrepresentative, making it difficult or impossible to identify control sites. Although the difficulties of conducting randomised controlled trials of these initiatives are not to be underestimated, these difficulties would not have been insurmountable. The benefits of having stronger evidence of effectiveness would have outweighed the costs of gaining this evidence, given the total national investment in these initiatives.

Few of the evaluations directly collected data about costs at a service-user level. Instead, where costs were considered at all, most evaluations had to make use of routinely collected data or reference costs, along with models based on a wide range of assumptions. Several evaluation reports described the difficulties experienced in obtaining data about the costs of either new or existing services. It is difficult to conduct meaningful evaluation of new initiatives when the information available about the costs of providing current services is so poor.

It is argued that cost-effectiveness is always the most important measure of the success of a new healthcare initiative. Even if a new initiative provides health benefits, if this is achieved at greater cost, then the extra investment in resources needs to be considered in comparison with other potential uses of those resources that may have offered greater benefits. This approach is well recognised and encouraged by government in relation to the appraisal of new drugs and technologies. It needs to be more widely applied to the introduction of other broader policy initiatives within health care.

CONCLUSION

An unusually extensive programme of evaluation was conducted of initiatives arising from the white paper, using a wide range of methodological approaches. There is some evidence of success in addressing the policy aims of the white paper, particularly innovations to improve access to care and to help people feel greater control over their health and health care.

However, all of the evaluations that have so far reported have major limitations. Despite a rhetorical commitment to evaluation, which to some extent has translated into several substantially funded projects, this was compromised by an inability or unwillingness to consider evaluation as an integral part of the implementation of these initiatives and to take account of the findings in making decisions about development of services. As a consequence, much of the considerable investment in evaluation at both national and local levels was not as productive as it might have been. A more systematic approach to evaluation of initiatives in health and social care is needed, with more use of direct comparisons with individuals or areas not receiving the new service, and much better assessment of objective benefits in relation to costs. The extent to which governments are committed to evidence-based policy can be assessed by how far they insist that new initiatives are implemented in a way that allows for meaningful evaluation, and how the results of evaluation are used in the policy process, rather by the amount of money spent on evaluation-related activity.

Chris Salisbury
Professor of Primary Health Care, Academic Unit of Primary Health Care, School of Social and Community Medicine, University of Bristol, Bristol.

Kate Stewart
Research Fellow, Academic Unit of Primary Health Care, School of Social and Community Medicine, University of Bristol, Bristol.

Sarah Purdy
Consultant Senior Lecturer, Academic Unit of Primary Health Care, School of Social and Community Medicine, University of Bristol, Bristol.

Helen Thorp
Research Assistant, Academic Unit of Primary Health Care, School of Social and Community Medicine, University of Bristol, Bristol.

Ailsa Cameron
Senior Lecturer and Head, Centre for Health and Social Care, School for Policy Studies, University of Bristol, Bristol.

Rachel Lart
Senior Lecturer, Centre for Health and Social Care, School for Policy Studies, University of Bristol, Bristol.
REFERENCES